

UNITEXT for Physics

Massimo Falcioni
Angelo Vulpiani

Meccanica Statistica Elementare

I fondamenti

 Springer

UNITEXT for Physics

Series editors

Michele Cini, Roma, Italy
Attilio Ferrari, Torino, Italy
Stefano Forte, Milano, Italy
Inguscio Massimo, Firenze, Italy
G. Montagna, Pavia, Italy
Oreste Nicrosini, Pavia, Italy
Luca Peliti, Napoli, Italy
Alberto Rotondi, Pavia, Italy

For further volumes:

<http://www.springer.com/series/13351>

Massimo Falcioni · Angelo Vulpiani

Meccanica Statistica Elementare

I fondamenti

 Springer

Massimo Falcioni
Università la Sapienza di Roma
Roma
Italy

Angelo Vulpiani
Università la Sapienza di Roma
Roma
Italy

ISSN 2198-7882

ISBN 978-88-470-5652-7

DOI 10.1007/978-88-470-5653-4

ISSN 2198-7890 (electronic)

ISBN 978-88-470-5653-4 (eBook)

Springer Milan Heidelberg New York Dordrecht London

© Springer-Verlag Italia 2015

Quest'opera è protetta dalla legge sul diritto d'autore e la sua riproduzione è ammessa solo ed esclusivamente nei limiti stabiliti dalla stessa. Le fotocopie per uso personale possono essere effettuate nei limiti del 15 % di ciascun volume dietro pagamento alla SIAE del compenso previsto dall'art. 68. Le riproduzioni per uso non personale e/o oltre il limite del 15 % potranno avvenire solo a seguito di specifica autorizzazione rilasciata da AIDRO, Corso di Porta Romana n.108, Milano 20122, e-mail segreteria@aidro.org e sito web www.aidro.org. Tutti i diritti, in particolare quelli relativi alla traduzione, alla ristampa, all'utilizzo di illustrazioni e tabelle, alla citazione orale, alla trasmissione radiofonica o televisiva, alla registrazione su microfilm o in database, o alla riproduzione in qualsiasi altra forma (stampata o elettronica) rimangono riservati anche nel caso di utilizzo parziale. La violazione delle norme comporta le sanzioni previste dalla legge. L'utilizzo in questa pubblicazione di denominazioni generiche, nomi commerciali, marchi registrati ecc., anche se non specificatamente identificati, non implica che tali denominazioni o marchi non siano protetti dalle relative leggi e regolamenti.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Prefazione

La meccanica statistica è spesso vista, in particolare nei corsi universitari introduttivi, come una materia ancillare di supporto allo studio della struttura della materia. A causa di questa impostazione, e anche della necessità di arrivare velocemente a fornire agli studenti strumenti per affrontare calcoli specifici, le parti di tipo più concettuale (quelle che coinvolgono la teoria ergodica e le connessioni tra meccanica statistica e calcolo delle probabilità) vengono ignorate o, nel migliore dei casi, accennate velocemente all'inizio del corso.

La meccanica statistica spesso sembra una sorta di metodologia generale, con deboli legami con la fisica delle cose che si toccano. Lo studente più attento rimane, per quanto riguarda aspetti come l'ergodicità, gli insiemi statistici e l'irreversibilità, con la sensazione di qualcosa di sospeso e vago, argomenti che (così viene spesso detto) in fondo non sono veramente essenziali per andare avanti.

Il problema si ripete anche su aspetti meno formali. Anche in alcuni buoni libri manca una pur rapida discussione delle evidenze sperimentali sulla validità della meccanica statistica e del suo utilizzo per determinare le caratteristiche del mondo atomico, combinando esperimenti e teoria. Per nostra esperienza diretta possiamo dire con certezza che la stragrande maggioranza degli studenti termina il corso di laurea triennale in fisica senza avere un'idea di come sia stato possibile determinare il numero di Avogadro e le caratteristiche essenziali del potenziale di interazione tra le molecole.

Lo scopo di questo libretto è riempire queste lacune discutendo gli aspetti di tipo fondamentale, sia di carattere fisico che concettuale, della meccanica statistica; in particolare:

- chiarire la rilevanza fisica dell'ipotesi ergodica e la sua connessione con la probabilità e gli insiemi statistici;
- discutere l'importanza dei tanti gradi di libertà coinvolti negli oggetti macroscopici;
- mostrare come, con l'ausilio di semplici risultati di meccanica statistica, sia possibile determinare le principali proprietà fisiche del mondo atomico;
- chiarire i principali aspetti tecnici e concettuali dell'irreversibilità.

Lo schema del libro è il seguente:

- Il Capitolo iniziale introduce i problemi di fondo della connessione tra termodinamica e meccanica statistica, con qualche accenno al ruolo della probabilità, la dinamica e i tanti gradi di libertà.
- Nei Capitoli 2 e 3 viene presentata l'ipotesi ergodica, la sua connessione con gli insiemi statistici, i suoi limiti e una sua versione debole che è interessante per gli oggetti macroscopici.
- Il Capitolo 4 è dedicato all'irreversibilità, argomento affascinante e ancora controverso: si cerca di mostrare come molti aspetti oscuri scompaiano se si accetta di considerare sistemi macroscopici.
- Il Capitolo 5 discute come, con l'aiuto della meccanica statistica, sia possibile determinare le proprietà fondamentali dei potenziali efficaci di interazione tra le molecole, la loro massa e il numero di Avogadro. Una particolare attenzione è dedicata al moto browniano.
- L'ultima parte è costituita da una serie di Appendici che sono state introdotte per diversi motivi. Alcune contengono risultati tecnici, utili per la discussione generale, che abbiamo preferito inserire alla fine del libro, per non spezzare il discorso e per non appesantire la presentazione. Altre sono dei complementi tecnici o concettuali di interesse per i lettori più esigenti; abbiamo inoltre inserito brevi note storiche.

A scanso di equivoci, diciamo esplicitamente che il termine "elementare" nel titolo è da intendersi come fondamentale, non come semplice.

Su alcune questioni basilari, quali, ad esempio, il ruolo dell'ergodicità e l'origine dell'irreversibilità, esistono punti di vista diversi da quello qui esposto, che noi riteniamo essere il più soddisfacente dal punto di vista concettuale e il più aderente alle idee di L. Boltzmann.

Il libro può essere utilizzato come un testo complementare ai tanti buoni libri di meccanica statistica, i prerequisiti richiesti sono solo la matematica di base a livello universitario, conoscenze elementari di meccanica statistica e probabilità.

Desideriamo ringraziare Luca Peliti, che ci ha incoraggiato a scrivere questo libro un po' fuori dagli usuali schemi; Luca Cerino, Sergio Ciuchi e Andrea Puglisi, che hanno letto parti del manoscritto dandoci utili suggerimenti. Un particolare ringraziamento va ad Andrea Gnoli, per l'attenta lettura di una prima versione del libro e a Sergio Ciuchi, per averci fornito alcune figure.

Roma,
primavera 2014

Massimo Falcioni
Angelo Vulpiani

Indice

1	Il problema di fondo	1
1.1	I paradossi	1
1.2	Cercando di superare i paradossi	4
1.3	Un'analogia dalla teoria delle probabilità	7
1.3.1	Qualche commento sulle probabilità	8
1.4	Ritorno alla Meccanica Statistica	11
1.5	Ruolo della dinamica	16
1.6	Lecture consigliate	17
2	Dalla Meccanica alla Termodinamica	19
2.1	Ipotesi ergodica e densità microcanonica	19
2.2	Commenti e osservazioni sull'ergodicità	21
2.3	L'entropia	29
2.4	Lecture consigliate	34
3	Il Microcanonico e oltre	35
3.1	Osservazioni preliminari sui volumi e sull'entropia	36
3.2	Dall'insieme microcanonico a quello canonico	38
3.3	La distribuzione di Maxwell-Boltzmann	40
3.4	Gli <i>ensembles</i>	43
3.5	Ergodicità e simulazioni numeriche	46
3.5.1	La dinamica molecolare	47
3.5.2	L'algoritmo di Verlet	48
3.5.3	Il metodo Monte Carlo	49
3.6	Lecture consigliate	50
4	Il problema dell'irreversibilità macroscopica	51
4.1	Il problema dell'irreversibilità	51
4.2	L'equazione di Boltzmann	53
4.2.1	Due paradossi apparentemente insormontabili	57
4.2.2	La risposta di Boltzmann	59

4.3	Un modello probabilistico in cui tutto è chiaro	61
4.3.1	Caos molecolare nel modello di Ehrenfest	62
4.3.2	Tempi di ricorrenza	64
4.4	Ancora sull'irreversibilità	64
4.5	Considerazioni finali	66
4.6	Letture consigliate	68
5	Studiare il mondo atomico con la meccanica statistica	69
5.1	Meccanica statistica ed esperimenti	70
5.1.1	La distribuzione di Maxwell	70
5.1.2	La distribuzione di Fermi-Dirac	72
5.1.3	Il problema del potenziale di interazione	74
5.2	Contare gli atomi senza poterli vedere: il moto Browniano	75
5.2.1	Un po' di storia	75
5.2.2	La teoria di Einstein nella versione di Langevin	77
5.2.3	L'equazione di Langevin a tempi discreti	79
5.2.4	Gli esperimenti di Perrin	80
5.2.5	L'importanza storica e concettuale del moto Browniano	81
5.3	Le fluttuazioni sono piccole ma rivelabili	81
5.4	Le dimensioni delle molecole	84
5.4.1	Qualche calcolo semplice	84
5.4.2	Dalla termodinamica ai potenziali di interazione	86
5.4.3	Osservazioni conclusive	87
5.5	Letture consigliate	88
	Appendice A Osservazioni su esponenziali e potenze	89
A.1	Le funzioni	89
A.2	Gli integrali	92
A.3	Volumi	95
	Appendice B Legge dei grandi numeri, piccole e grandi fluttuazioni	97
B.1	Dalla legge binomiale al teorema del limite centrale	97
B.2	Oltre il limite centrale	99
B.3	Grandi deviazioni	101
	Appendice C Un altro modo di scrivere la densità nel microcanonico	105
	Appendice D La teoria di Einstein del moto Browniano	107
	Appendice E Relazione di Taylor-Kubo per il coefficiente di diffusione	111
	Appendice F Lo sviluppo del viriale per gas debolmente interagenti	113
	Appendice G Le catene di Markov in due parole	117
	Appendice H Tempi di ricorrenza	121

Appendice I Un tributo a Ludwig Boltzmann, il grande visionario 123

Indice analitico 131

Il problema di fondo

Sommario Nella seconda metà del *XIX* secolo tre grandi schemi teorici dominavano la fisica: la Meccanica, l'Elettromagnetismo e la Termodinamica.

La Meccanica regolava il moto dei corpi materiali; l'Elettromagnetismo dava le leggi che guidano i fenomeni elettrici e magnetici; la Termodinamica dettava i principi che sottostanno agli scambi energetici tra sistemi nei processi di raggiungimento dell'equilibrio termico. Quest'ultima teoria era stata sviluppata partendo da osservazioni ed esperimenti sui corpi materiali, ma aveva alla fine mostrato la grande generalità delle sue leggi quando, applicata allo studio delle proprietà della radiazione elettromagnetica in equilibrio all'interno di una cavità (problema del corpo nero), aveva correttamente predetto un andamento proporzionale alla quarta potenza della temperatura per la densità di energia del campo elettromagnetico (legge di Stefan e Boltzmann).

La Meccanica Statistica si pose l'obiettivo di ricondurre la Termodinamica alla Meccanica, partendo dall'*ipotesi atomistica*, cioè dall'assunzione che la materia sia costituita da unità microscopiche elementari, atomi e molecole, che seguono le leggi della meccanica. È opportuno ricordare che all'inizio del *XX* secolo la realtà degli atomi non aveva ancora ricevuto convincente evidenza sperimentale e l'ipotesi atomistica era tutt'altro che universalmente accettata. Al contrario, le prove della validità delle leggi della Termodinamica, anche al di fuori del campo in cui essa era stata sviluppata, facevano guardare con molto scetticismo al tentativo di unificazione tra le due teorie, portato avanti dalla Meccanica Statistica.

L'obiezione fondamentale nei riguardi del progetto si può ricondurre a una serie di incompatibilità tra i principi fondanti delle due teorie.

1.1 I paradossi

La Termodinamica è basata sull'osservazione che esistono *stati di equilibrio*: un insieme di corpi che possono scambiare tra loro energia, posti in contatto e mantenuti in sufficiente isolamento con l'esterno, raggiungono uno stato in cui scambi d'ener-

gia (e anche eventualmente flussi di materia o altro) cessano; le variabili con cui il sistema è caratterizzato non mostrano più dipendenza dal tempo. Questa osservazione, assunta a principio, introduce la nozione di irreversibilità in Termodinamica, affermando che esistono processi che non ammettono un'evoluzione temporale inversa: se un sistema ha raggiunto l'equilibrio in uno stato diverso dal suo stato iniziale, non ritorna spontaneamente in quest'ultimo. Più esattamente due eventi sono esclusi dalla Termodinamica:

- a) che il processo di rilassamento all'equilibrio abbia un inverso (a livello macroscopico); cioè che il sistema possa essere preparato in uno stato iniziale di equilibrio, in modo tale che spontaneamente evolva verso il vecchio stato iniziale, di non equilibrio;
- b) che il sistema, una volta raggiunto l'equilibrio, mantenuto in isolamento, al trascorrere del tempo torni nel suo stato iniziale, di non equilibrio.

Entrambi questi divieti appaiono ingiustificabili in meccanica. Si assuma il punto di vista atomistico e si supponga che un sistema isolato sia costituito da N particelle tra cui agiscono delle forze, \mathbf{F}_i , che dipendono solo dalle posizioni, \mathbf{x}_i , delle particelle.

\tilde{a}) La struttura delle equazioni di Newton

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} = \mathbf{F}_i(\mathbf{x}_1, \dots, \mathbf{x}_N) \quad (i = 1, 2, \dots, N) \quad (1.1)$$

è tale che, se $(\mathbf{x}_1(t), \dots, \mathbf{x}_N(t)) \equiv \{\mathbf{x}_i(t)\}$ è una soluzione permessa, descrivente una possibile evoluzione delle N particelle, allora anche $\{\mathbf{x}_i(-t)\}$, descrivente il moto inverso, con tutte le velocità invertite, è una soluzione possibile. Questo perché, se effettuiamo la trasformazione $t \rightarrow -t$ e $\mathbf{v}_i \rightarrow -\mathbf{v}_i$ mentre \mathbf{x}_i non cambia, l'Eq.(1.1), che dipende solo dalle derivate seconde rispetto al tempo, resta inalterata. Più precisamente: si consideri una condizione iniziale con particelle nelle posizioni $(\mathbf{x}_1(0), \dots, \mathbf{x}_N(0))$ con velocità $(\mathbf{v}_1(0), \dots, \mathbf{v}_N(0))$, e la si lasci evolvere, secondo l'Eq.(1.1), fino al tempo t . A questo tempo si invertano le velocità, e si prendano come nuove condizioni iniziali quelle con le particelle in $(\mathbf{x}_1(t), \dots, \mathbf{x}_N(t))$ con velocità $(-\mathbf{v}_1(t), \dots, -\mathbf{v}_N(t))$. L'evoluzione dalla nuova condizione, durante un altro intervallo di tempo t , riporterà le particelle del sistema nelle posizioni $(\mathbf{x}_1(0), \dots, \mathbf{x}_N(0))$ con velocità invertite $(-\mathbf{v}_1(0), \dots, -\mathbf{v}_N(0))$.

Questa proprietà delle soluzioni della (1.1) è in contrasto con il punto *a*). Si assuma che le N particelle descrivano due sistemi in contatto. Per fissare le idee, e per semplicità, qui e nel seguito del capitolo, consideriamo come sistema di esempio due gas che occupano volumi fissati e possono scambiare energia attraverso una parete divisoria: un gas *A* con N_A particelle e un gas *B* con N_B particelle ($N_A + N_B = N$). Indicando con $V(\{\mathbf{x}_i\})$ l'energia potenziale associata alle forze tra le molecole, l'energia totale del sistema

$$E_{tot} = \sum_{i=1}^N \frac{1}{2} m \mathbf{v}_i^2 + V(\{\mathbf{x}_i\}), \quad (1.2)$$

si può decomporre in questo modo:

$$E_{tot} = \sum_{i=1}^{N_A} \frac{1}{2} m \mathbf{v}_i^2 + V_A + \sum_{i=1}^{N_B} \frac{1}{2} m \mathbf{v}_i^2 + V_B + V_{AB} \equiv E_A + E_B + V_{AB}, \quad (1.3)$$

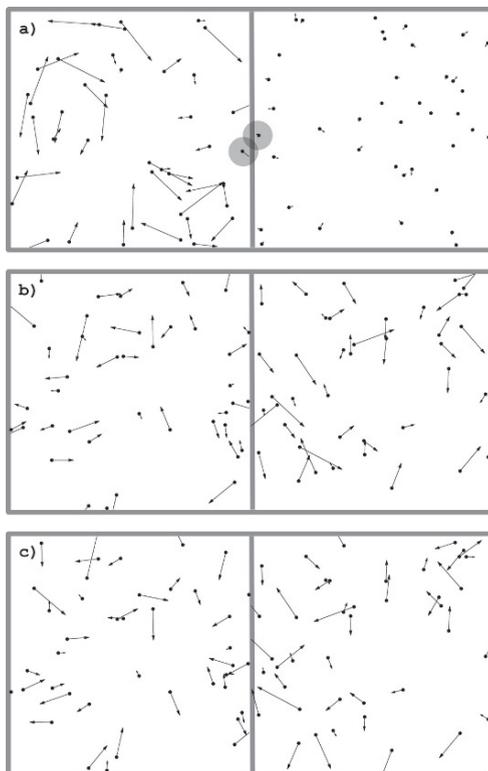
dove V_A (V_B) dipende solo dalle particelle del gas A (B) e V_{AB} indica il contributo all'energia potenziale dovuto alle molecole dei due gas che si trovano in prossimità della parete di separazione e che permettono lo scambio di energia tra i due sistemi. Tali molecole sono quelle che distano dalla parete per una quantità dell'ordine della portata delle interazioni intermolecolari e sono quindi relativamente poche rispetto al numero totale¹: anche se definito con un certo grado di arbitrarietà, quantitativamente V_{AB} è una piccola correzione a $E_A + E_B$. Si supponga che lo stato meccanico iniziale $\{\mathbf{x}_i(0)\}$, $\{\mathbf{v}_i(0)\}$ sia di non equilibrio, caratterizzato da una distribuzione di energia tra i due sistemi data da $E_A(0) + E_B(0) + V_{AB}(0) = E_{tot}$. Si assuma che al tempo t il sistema sia all'equilibrio termodinamico, caratterizzato dalla distribuzione energetica $E_A^{eq} + E_B^{eq} + V_{AB}^{eq} = E_A(t) + E_B(t) + V_{AB}(t) = E_{tot}$, dove, a meno di una correzione trascurabile (data da V_{AB}^{eq}), E_A^{eq} e E_B^{eq} sono le energie di equilibrio dal punto di vista termodinamico ($E_A^{eq} \neq E_A(0)$ e $E_B^{eq} \neq E_B(0)$). Si esegua ora l'inversione descritta sopra. Tenendo conto del fatto che le energie dipendono solo dal modulo delle velocità, e quindi anche il nuovo stato iniziale ha le energie di equilibrio, si capisce che si sta preparando uno stato in cui due sistemi inizialmente in equilibrio in un tempo t vanno fuori equilibrio ritornando nel vecchio stato iniziale, vedi Fig.1.1.

\tilde{b}) Un'importante proprietà delle equazioni meccaniche è relativa a sistemi il cui moto avviene in una regione finita dello spazio, come quelli di cui si occupa la Termodinamica. In questo caso vale il *Teorema di ricorrenza* di Poincaré: il sistema ritorna (infinite volte) arbitrariamente vicino a qualunque suo stato dinamico. Ciò implica la violazione da parte dei sistemi meccanici del divieto b) imposto dalla Termodinamica. Ritornando ai due gas introdotti sopra, una volta che sia stato raggiunto l'equilibrio (al tempo t) il teorema di Poincaré implica che, aspettando un tempo sufficiente, il sistema ripasserà arbitrariamente vicino allo stato iniziale di non equilibrio: in meccanica l'equilibrio non è per sempre.

Le incompatibilità di principio delle leggi meccaniche coi punti a) e b) sono note, rispettivamente, come *paradosso della reversibilità* (dovuto a J. Loschmidt) e *paradosso della ricorrenza* (dovuto a E. Zermelo).

¹ Per esempio, per un gas contenuto in un recipiente cubico di lato $L = 10$ cm, essendo la portata delle interazioni tra le molecole, r_0 , dell'ordine di $1\text{Å} = 10^{-10}$ m, la frazione di particelle contenute entro una distanza r_0 da una parete è r_0/L , di ordine 10^{-9} .

Figura 1.1 a) Due sistemi inizialmente non all'equilibrio evolvono verso lo stato di equilibrio. La velocità di una particella è proporzionale alla lunghezza della freccia che parte da essa. Inizialmente il gas contenuto nella parte sinistra del recipiente ha temperatura maggiore di quello nella parte destra. Le interazioni fra le particelle in vicinanza della superficie di contatto fra i due sistemi sono responsabili del raggiungimento dello stato di equilibrio. Due particelle ed il raggio d'azione delle forze di interazione (non in scala) sono mostrate in a). In b) è mostrato lo stato di equilibrio, in c) quello ottenuto da b) in vertendo *tutte* le velocità.



1.2 Cercando di superare i paradossi

Se si vogliono far coesistere due teorie che presentano contraddizioni derivanti dai loro principi, si deve necessariamente intervenire sui principi stessi. In questo caso il raccordo tra Meccanica e Termodinamica, prodotto dalla Meccanica Statistica, si ottiene a costo di una interpretazione meno rigida della nozione di equilibrio.

Un primo aspetto da notare è che, una volta accettato che un corpo macroscopico sia costituito da un numero N molto grande di molecole (dell'ordine del numero di Avogadro $N_0 \sim 10^{24}$), la meccanica di un insieme di N molecole fornisce l'evoluzione di $6N$ variabili dinamiche², che descrivono completamente gli stati microscopici del sistema. È utile studiare il moto nello spazio delle fasi usando come variabili le $3N$ componenti delle posizioni e le $3N$ componenti degli impulsi, che indichiamo rispettivamente con $\{\mathbf{q}_i\}$ e $\{\mathbf{p}_i\}$, la cui l'evoluzione è guidata dalle equazioni di

² In questo capitolo assumiamo che le molecole siano ben descritte dalle sole variabili del centro di massa ovvero che si possano considerare puntiformi.

Hamilton

$$\begin{aligned} d\mathbf{q}_i/dt &= \partial H/\partial \mathbf{p}_i = \mathbf{p}_i/m \\ d\mathbf{p}_i/dt &= -\partial H/\partial \mathbf{q}_i = -\partial V/\partial \mathbf{q}_i \end{aligned} \quad (1.4)$$

dove si è introdotta la funzione di Hamilton

$$H = \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{2m} + V(\{\mathbf{q}_j\}), \quad (1.5)$$

che, in questo caso, fornisce anche l'energia del sistema. È comodo usare una notazione compatta per indicare uno stato microscopico completo di un sistema di N particelle mediante la seguente definizione: $\mathbf{X} \equiv (\mathbf{q}_1, \dots, \mathbf{q}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$; \mathbf{X} è un vettore con $6N$ componenti reali.

Al contrario, se si studia lo stesso corpo dal punto di vista termodinamico, le variabili necessarie per una descrizione esauriente sono molto poche; nell'esempio del gas diviso in due parti di volume fissato, una sola variabile è sufficiente: l'energia di uno dei due gas, E_A o E_B , l'altra essendo fissata dalla conservazione dell'energia totale, potendosi trascurare il termine di interazione reciproca V_{AB} .

Quindi la meccanica fornisce una descrizione del sistema *enormemente*, ed *eccessivamente* più dettagliata (e ridondante) di quella termodinamica. Inoltre, mentre l'evoluzione meccanica descrive le variazioni dello stato di ogni singola molecola, la termodinamica si occupa delle variazioni di quantità *collettive*, come l'energia di un sistema, che istante per istante ricevono contributi da *tutte* le molecole che lo compongono. In questo senso il punto di vista meccanico si può chiamare *microscopico* e quello termodinamico *macroscopico*.

Lo studio del sistema dal punto di vista microscopico riguarda osservazioni sperimentali coinvolgenti una o poche molecole. Supponiamo di essere in grado di misurare l'energia di una data molecola del gas, la i -esima, di energia $e_i = e_i(t)$. Usando l'equazione di stato dei gas ideali monoatomici, si può stimare l'ordine di grandezza (che indichiamo con \tilde{e}) della variabile e_i a temperatura ambiente ($T = 300$ K):

$$\tilde{e} = \frac{3}{2} k_B T \approx 6.2 \cdot 10^{-21} \text{ J}. \quad (1.6)$$

dove k_B è la costante di Boltzmann. È chiaro che per la variabile e_i non è sensato parlare di equilibrio, come situazione in cui scompare la dipendenza temporale: a causa della presenza delle altre molecole, con cui la i -esima interagisce, il suo valore cambierà in continuazione, fluttuando intorno a \tilde{e} per quantità dello stesso ordine di grandezza³, subendo quindi variazioni di cui ci accorgiamo, essendo in grado di misurarle, per l'ipotesi fatta all'inizio.

³ Per esempio, nell'urto elastico bidimensionale tra una particella di massa m ed energia (cinetica) iniziale E e una particella di massa $m' = \gamma m$, inizialmente ferma, la variazione di energia della prima particella si scrive: $|\Delta E| = E (\cos\theta)^2 4\gamma/(1+\gamma)^2$, dove θ è l'angolo tra le direzioni finale e iniziale della particella incidente. Se le masse sono dello stesso ordine di grandezza si ha: $|\Delta E| = O(E)$. Con la notazione $O(x)$ intendiamo indicare una quantità dello stesso ordine di grandezza di x .

Dal punto di vista macroscopico si è invece interessati a misurare l'energia, per esempio, del sistema A , il cui valore, a meno di piccole correzioni dovute a V_{AB} , è dato dalla variabile $E_A = E_A(t)$, che come ordine di grandezza è stimabile in

$$\tilde{E}_A = N_A \tilde{e} \approx N_0 \tilde{e} \approx 6.2 \cdot 10^3 \text{ J}. \quad (1.7)$$

A questo punto è necessario chiarire due cose: l'incertezza sulla misura di E_A e l'ordine di grandezza della variazione che E_A subisce nel tempo.

Supponiamo che il sistema d'esempio sia in equilibrio termodinamico. Indicando con U_A il risultato di una misura di energia di A e con Δ la sua incertezza, possiamo scrivere che, a un certo tempo, assunto come iniziale, si ha

$$U_A - \Delta < E_A < U_A + \Delta, \quad (1.8)$$

dove U_A è una quantità di ordine \tilde{E}_A . Seguendo le leggi del moto (1.4-1.5) l'osservabile $E_A(t)$ fluttuerà intorno a un valore di ordine \tilde{E}_A . Assumiamo come valore di riferimento U_A , quello determinato al tempo iniziale, e indichiamo con $\delta(t)$ la differenza tra $E_A(t)$ e U_A , cioè poniamo

$$E_A(t) = U_A + \delta(t). \quad (1.9)$$

È chiaro che, perché sia possibile recuperare la nozione di equilibrio per la variabile E_A , è cruciale il confronto tra Δ e $\delta(t)$. Se tra i valori che assume E_A a tempi diversi ci sono differenze di ordine Δ , o più grandi, misure effettuate in tempi successivi lo rivelano e concludiamo che E_A varia nel tempo. Se la variazione di E_A durante l'evoluzione del sistema è invece tale che risulta $|\delta(t)| \ll \Delta$, l'apparato di misura non rivelerà differenze tra i valori di E_A , che apparirà costante.

Se esiste qualche ragione per aspettarsi che una osservabile macroscopica, come E_A , al trascorrere del tempo subisca fluttuazioni "molto piccole": $|\delta(t)| \ll \Delta$, avremmo dei motivi per parlare di regime d'equilibrio per l'osservabile, potendo affermare che, *entro l'incertezza sperimentale*, il valore dell'osservabile non muta nel tempo.

Sappiamo che una variabile di singola molecola, come e_i , subisce continui cambiamenti relativamente grandi, cioè tali che la differenza a tempi successivi $\delta e = e_i(t') - e_i(t)$ risulta una frazione finita dei valori di e_i in t o t' , per cui: $\delta e = c e_i(t)$ con c di ordine 1. C'è però una sostanziale differenza tra e_i ed E_A in quanto il valore che quest'ultima osservabile *macroscopica* assume su ogni singolo stato dinamico del sistema A , è determinato dal contributo di un numero enorme $N_A \approx N_0$ di quantità *microscopiche*. Per esempio, nell'approssimazione di gas ideale si ha esattamente

$$E_A = \sum_{i=1}^{N_A} e_i. \quad (1.10)$$

A questo punto, volendo confrontare le fluttuazioni di e_i ed E_A , vengono alla mente la legge dei grandi numeri e il teorema del limite centrale nella teoria delle probabilità.

1.3 Un'analogia dalla teoria delle probabilità

Le leggi probabilistiche cui si allude nella sezione precedente si possono presentare nel modo seguente. Si consideri un esperimento in cui può verificarsi l'evento a , con probabilità p_a , oppure l'evento complementare b , con probabilità $p_b = 1 - p_a$. Si supponga di ripetere l'esperimento N volte, contando quante volte si presentano gli eventi, n_a e n_b , e calcolando le frequenze di apparizione: $v_a = n_a/N$, $v_b = n_b/N$. Naturalmente si ha $n_a + n_b = N$ e $v_a + v_b = 1$. Se si prendono in esame tutte le possibili sequenze lunghe N che possono realizzarsi (sono 2^N), la legge dei grandi numeri afferma che la probabilità totale delle sequenze in cui $v_a \rightarrow p_a$, e quindi $v_b \rightarrow p_b$, nel limite $N \rightarrow \infty$ tende a 1. Dal teorema del limite centrale segue inoltre che, per grandi valori di N , la distribuzione di probabilità di v_a è ben approssimabile da una gaussiana, centrata in p_a con varianza $(p_a p_b)/N$.

Gli eventi a e b sono arbitrari (per esempio, pioverà o non pioverà), ma è comodo caratterizzarli con variabili numeriche. Supponiamo che all'evento a sia associato il valore y_a di una variabile y , e all'evento b il valore y_b ; il valore medio e la varianza della variabile y su un singolo esperimento sono quindi

$$\begin{aligned}\langle y \rangle &= p_a y_a + p_b y_b, \\ \sigma^2 &= \langle (y - \langle y \rangle)^2 \rangle = \langle y^2 \rangle - \langle y \rangle^2 = p_a p_b (y_a - y_b)^2.\end{aligned}\tag{1.11}$$

A questo punto possiamo definire su ogni sequenza lunga N la variabile $Y_N = \sum_{i=1}^N y^{(i)}$, avendo indicato con $y^{(i)}$ il valore di y realizzato nell' i -esima ripetizione dell'esperimento elementare. La legge dei grandi numeri ci permette allora di affermare che, per N molto grande, sulla classe di sequenze la cui probabilità complessiva tende a 1, la variabile somma ha valore

$$\bar{Y}_N = \bar{n}_a y_a + \bar{n}_b y_b \cong N(p_a y_a + p_b y_b) = N \langle y \rangle = \langle Y_N \rangle,\tag{1.12}$$

mentre il teorema del limite centrale assicura che la distribuzione di probabilità di Y_N , per N molto grande, è ben approssimata da una gaussiana centrata in $\langle Y_N \rangle \cong \bar{Y}_N$ con varianza

$$\sigma_N^2 = N \sigma^2.\tag{1.13}$$

Per una discussione su legge dei grandi numeri e teorema del limite centrale si veda la sezione Appendice B delle Appendici.

L'analogia col problema da cui siamo partiti è abbastanza evidente. Possiamo interpretare le sequenze lunghe N come gli stati di un gas di N molecole: la i -esima molecola ha energia $e_i = y^{(i)}$, il gas ha energia $E_A = Y_N$. Mentre il gas al trascorrere del tempo passa da uno stato a un altro (da una sequenza a un'altra) l'energia della i -esima molecola potrà passare da y_a a y_b o viceversa, subendo quindi una variazione dello stesso ordine di grandezza del suo valore iniziale; stiamo assumendo che i valori y_a siano y_b simili. Al contrario, l'energia del gas $E_A = \sum_{i=1}^N y^{(i)}$, se il passaggio, come è naturale assumere, avviene principalmente tra le sequenze con Y_N di maggior probabilità, subirà una variazione di ordine $\sigma \sqrt{N}$ intorno al valore

$\langle E_A \rangle = N \langle y \rangle$. Dato che $\langle y \rangle$ e σ sono quantità dello stesso ordine di grandezza, si ha che, se $N = O(N_0)$, la fluttuazione del valore di E_A , $|\delta| \approx \sigma \sqrt{N}$, risulta molto minore di $E_A \approx N \langle y \rangle$, e quindi di ogni ragionevole incertezza Δ sulla sua misura. Abbiamo trovato un buon motivo per aspettarci che si abbia $|\delta(t)| \ll \Delta$.

1.3.1 Qualche commento sulle probabilità

È opportuno tornare brevemente sul contenuto della legge dei grandi numeri e del teorema del limite centrale.

1) Le probabilità delle singole sequenze, essendo date da prodotti di N termini (p_a o p_b) minori di 1, sono tutte esponenzialmente piccole in N ; si può scrivere $p_N = \exp(-\alpha \times N)$, dove $\alpha(v_a, v_b | p_a, p_b)$ è una funzione positiva delle frequenze (Nella sezione Appendice B delle Appendici si trova una esposizione più dettagliata e quantitativa degli argomenti esposti in questo punto). Se non si è interessati a una particolare realizzazione della sequenza di esperimenti (cioè non interessa sapere nel dettaglio se si è verificato, per esempio, prima l'evento a , poi b , poi ancora b , ecc.) ma si vuol sapere solo quante volte è avvenuto a (e quindi b), allora si raggruppano le 2^N sequenze in classi, caratterizzate dai valori (n_a, n_b) oppure (v_a, v_b) , e si calcola la probabilità della classe, moltiplicando la probabilità di una singola sequenza appartenente alla classe per il numero di elementi della classe (dato che, in questo caso, tutte le sequenze di una classe hanno uguale probabilità). Esistono anche classi con pochi elementi, ma quelle che interessano, nel limite $N \rightarrow \infty$, sono quelle contenenti un numero esponenzialmente grande, in N , di sequenze per le quali si ha $\mathcal{N}_N \sim \exp(\sigma \times N)$, con $\sigma(v_a, v_b) > 0$. La legge dei grandi numeri segue dal fatto nel delicato bilanciamento tra una quantità esponenzialmente piccola (la probabilità di singola sequenza, p_N) e una esponenzialmente grande (la molteplicità della classe, \mathcal{N}_N), nel limite $N \rightarrow \infty$, il prodotto $p_N \times \mathcal{N}_N \sim \exp[-(\alpha - \sigma) \times N]$ resta diverso da zero, e quindi tende a 1, per una sola classe: quella con $v_a = p_a, v_b = p_b$, per cui si ha $\alpha = \sigma$ (si può verificare che $(\alpha - \sigma) \geq 0$). Avendo introdotto la variabile y , la classe la cui probabilità tende a 1 è anche caratterizzabile col valore assunto da Y_N sulle sue sequenze.

Si noti che la classe di alta probabilità, in generale, non è costituita dalle sequenze singolarmente più probabili, né coincide con quella più numerosa; come detto, è il bilanciamento tra molteplicità e probabilità di singola sequenza che genera la dominanza risultante. Per chiarire questo punto consideriamo il caso di variabili binarie indipendenti $a = 0$ e $b = 1$ con probabilità diverse da $1/2$ ad esempio $p_a = 0.9$ e $p_b = 0.1$.

Per ogni N la successione più probabile è chiaramente quella costituita da tutti zeri, cioè $0, 0, \dots, 0$, che ha probabilità $(0.9)^N = e^{\ln(0.9)N}$ e costituisce una classe a sé. Invece le successioni tipiche, cioè nella classe con probabilità circa uno per $N \gg 1$, in accordo con la legge dei grandi numeri, sono quelle in cui il numero di zeri è circa il 90%. Per ognuna di queste successioni la probabilità è $(0.9)^{0.9N} (0.1)^{0.1N} = e^{-hN} \ll e^{\ln(0.9)N}$ dove $h = -0.1 \ln(0.1) - 0.9 \ln(0.9) \simeq 0.324$

mentre $\ln(0.9) \simeq -0.105$. Inoltre per ogni N la molteplicità di una classe è data dal coefficiente binomiale $\binom{N}{n_a}$, che ha massimo per il valore di n_a più vicino a $N/2$; la classe più numerosa contiene quindi le sequenze con uguale percentuale di zeri e uni. Solo nel caso particolare di sequenze equiprobabili (se $p_a = p_b = 1/2$, e quindi $v_a = v_b = 1/2$) la classe di massima probabilità è anche quella di massima molteplicità.

2) Esiste un'interessante interpretazione geometrica di quanto esposto sopra. La classe dominante corrisponde a quella che occupa maggior volume tra tutte e, nel limite $N \gg 1$, il volume totale. Data una sequenza lunga N , la sua probabilità, prodotto di N termini p_a o p_b , si può pensare come il volume di un iperparallelepipedo, i cui lati hanno lunghezza p_a o p_b , in uno spazio di dimensione N . Assumiamo che sia $p_a \neq p_b$. Tutti gli iperparallelepipedi (cioè tutte le sequenze) occupano un volume pari a 1 (che è la probabilità totale delle sequenze) e riempiono esattamente un ipercubo di lato 1 in questo spazio (ogni lato dell'ipercubo va pensato suddiviso in due segmenti di lunghezze p_a e p_b). Le sequenze con ugual numero di lati lunghi p_a (e p_b) hanno ugual volume (probabilità) e appartengono alla stessa classe, avendo ugual valore della variabile v_a (e v_b). Il volume occupato da una classe (ovvero la sua probabilità) è quindi il prodotto tra la molteplicità della classe (\mathcal{A}_N) e il volume caratteristico dei suoi iperparallelepipedi (p_N). La suddivisione della probabilità totale delle sequenze nei contributi delle varie classi di sequenze equivale quindi alla suddivisione del volume dell'ipercubo nei volumi occupati dalle varie classi di iperparallelepipedi. Disponendo della variabile y , le classi possono caratterizzarsi, oltre che con la quantità n_a (o v_a), anche col valore della variabile Y_N sulle varie classi $Y_N = n_a y_a + n_b y_b = N(v_a y_a + v_b y_b)$. In questo schema, la legge dei grandi numeri afferma che, quando la dimensionalità dello spazio $N \rightarrow \infty$, il volume dell'ipercubo è dominato dal volume di un'unica classe. Ovvero, per $N \gg 1$ il volume dell'ipercubo è occupato quasi esclusivamente dagli iperparallelepipedi su cui $Y_N = N(p_a y_a + p_b y_b)$ e da quelli su cui Y_N differisce da questo valore per quantità di ordine \sqrt{N} .

Nel caso particolare $p_a = p_b = 1/2$ gli iperparallelepipedi degenerano in ipercubi di ugual volume e la classe di massimo volume (praticamente tutto il volume nel limite $N \rightarrow \infty$) diventa quella di massima molteplicità, in corrispondenza al valore $Y_N = N(y_a + y_b)/2$.

Queste proprietà di "saturazione" dei volumi da parte di una singola classe, che si manifesta negli spazi di alta dimensionalità e che equivale alla legge dei grandi numeri, è legata alle proprietà delle funzioni esponenziali di tipo $\exp(Nf)$, con $N \gg 1$ e f una quantità di ordine 1; proprietà che sono anche alla base dei metodi di integrazione alla Laplace. Esponenziali del genere compaiono facilmente nello studio di funzioni in spazi di dimensione N . Per esempio, nel caso interessante di una funzione definita come somma di contributi provenienti dalle singole componenti $F(x_1, x_2, \dots, x_N) = \sum_i f(x_i)$, se si considera una regione di volume finito e la si suddivide in sottoregioni classificate mediante i valori assunti dalla F (che ha qui il ruolo della Y_N), quello che succede tipicamente è che esiste una sottoregione di volume massimo che, nel limite $N \gg 1$, è esponenzialmente più grande delle altre. Una conseguenza inaspettata e sorprendente di ciò è il fatto che i volumi di iperso-

lidi convessi (come i cubi e le sfere) sono concentrati in una buccia sottilissima al di sotto della superficie esterna (per i dettagli vedere la sezione Appendice A delle Appendici).

3) Abbiamo discusso il caso in cui in ogni esperimento elementare siano possibili 2 risultati. Tutto quanto esposto sopra si estende facilmente al caso con M eventi (mutuamente esclusivi) a_1, a_2, \dots, a_M con probabilità rispettive p_1, p_2, \dots, p_M . Su ogni sequenza lunga N si contano le volte n_k in cui si è verificato il risultato a_k e si definisce la relativa frequenza $v_k = n_k/N$. La legge dei grandi numeri assicura che, nel limite $N \rightarrow \infty$, la probabilità complessiva delle sequenze in cui $v_k \rightarrow p_k$ tende a 1. Inoltre, se all'evento a_k è associato il valore y_k di una variabile numerica y , su ogni sequenza lunga N si può definire la variabile $Y_N = \sum_{i=1}^N y^{(i)}$, avendo ancora indicato con $y^{(i)}$ il valore di y realizzato nell' i -esima ripetizione dell'esperimento elementare. La legge dei grandi numeri permette allora di affermare che, per N molto grande, sulle sequenze la cui probabilità complessiva tende a 1, la variabile somma ha valore

$$\bar{Y}_N = \sum_k \bar{n}_k y_k \cong N \sum_k p_k y_k = N \langle y \rangle = \langle Y_N \rangle, \quad (1.14)$$

mentre il teorema del limite centrale garantisce che la distribuzione di probabilità di Y_N , in quel limite, è ben approssimata da una gaussiana centrata in $\langle Y_N \rangle$ con varianza

$$\sigma_N^2 = N \sigma^2 = N \sum_{k=1}^M p_k (y_k - \langle y \rangle)^2. \quad (1.15)$$

L'interpretazione basata sui volumi in questo caso significa: ogni sequenza lunga N è rappresentata da un iperparallelepipedo i cui lati possono avere una lunghezza scelta tra M ; tutte le sequenze con la stessa distribuzione di frequenze $\{v_k\}$ hanno la stessa probabilità e occupano ugual volume nell'ipercubo di volume unitario; nel limite $N \rightarrow \infty$ la classe con frequenze $\{v_k\} \rightarrow \{p_k\}$ tende a occupare tutto il volume. Ovvero per $N \gg 1$ il volume dell'ipercubo è occupato quasi esclusivamente dagli iperparallelepipedi su cui $Y_N = N \sum_k p_k y_k$ e da quelli su cui Y_N differisce da questo valore per quantità di ordine \sqrt{N} .

Questi risultati relativi a M possibilità per ogni evento elementare, rendono naturalmente ancora più significativa l'analogia con un gas, le cui molecole dispongono di un numero illimitato di stati energetici.

4) Precedentemente abbiamo introdotto l'idea di interpretare ogni sequenza lunga N come un possibile stato microscopico di un gas di N molecole. È utile allora supporre che sia data una regola che permetta di passare da una sequenza a un'altra, cosicché, continuando nell'analogia, possiamo disporre di una legge di evoluzione per gli stati del gas.

Supponiamo che il processo che regola la transizione tra le sequenze si svolga a tempi discreti, multipli di un intervallo fondamentale τ che poniamo uguale a 1. Indicando con $s^{(N)}$ una sequenza lunga N , la legge del moto produrrà una successione di sequenze: $\{s_1^{(N)}, s_2^{(N)}, \dots, s_r^{(N)}\}$, per ognuna delle quali è definita una

probabilità $Prob \{s_1^{(N)}, s_2^{(N)}, \dots, s_r^{(N)}\}$. Qui non interessa specificare, né la legge del moto (che, in generale sarà di tipo casuale), né come sono fatte, di conseguenza, queste probabilità. È però necessario, e ragionevole, imporre che non ci siano sequenze “irraggiungibili” e che ci si trovi in regime stazionario. La prima condizione richiede che da qualunque sequenza si parta si possa arrivare, prima o poi, a ognuna delle altre. La seconda significa che la probabilità di osservare una data successione $\{s_1^{(N)}, s_2^{(N)}, \dots, s_r^{(N)}\}$ non dipenda dal momento iniziale (che in questo esempio è stato posto a $t = 1$); deve quindi risultare $Prob \{s_{1+k}^{(N)}, s_{2+k}^{(N)}, \dots, s_{r+k}^{(N)}\} = Prob \{s_1^{(N)}, s_2^{(N)}, \dots, s_r^{(N)}\}$, per ogni k . Naturalmente per $r = 1$ si riottiene la probabilità di una singola sequenza lunga N , che abbiamo introdotto nel punto 1: $Prob \{s_1^{(N)}\} = p_N = \exp(-\alpha \times N)$.

Per un sistema che soddisfa queste condizioni, il cosiddetto *lemma di Kac* (per i dettagli vedere la sezione Appendice H delle Appendici), permette di determinare il tempo medio, $\langle \mathcal{T} \rangle$, che deve trascorrere affinché, partendo da una qualunque sequenza, $\tilde{s}^{(N)}$, l'evoluzione riporti il sistema in $\tilde{s}^{(N)}$. Si ha

$$\langle \mathcal{T} \rangle = \frac{\tau}{Prob \{\tilde{s}^{(N)}\}} = \tau e^{(\alpha \times N)}. \quad (1.16)$$

Il risultato (1.16), in cui abbiamo reintrodotta esplicitamente l'intervallo elementare τ , ha un contenuto intuitivo: poiché il numero delle sequenze è esponenzialmente alto (sono, nel caso generale, M^N) e la dinamica può visitarle tutte, con preferenze confrontabili date da p_N , ci si deve aspettare che, tipicamente, prima di ritornare in una data sequenza, il sistema visiti una buona parte di tutte le altre, impiegando quindi un tempo esponenzialmente grande in N .

Ritornando all'analogia con un gas di N molecole, il lemma di Kac (1.16) è molto interessante perché suggerisce che la critica alla Meccanica Statistica basata sul paradosso della ricorrenza, è in realtà irrilevante a livello macroscopico. È sicuramente vero che un qualunque sistema meccanico ritorna vicino a ogni suo stato; ma il tempo necessario perché ciò avvenga, e si possa eventualmente osservare un allontanamento spontaneo dall'equilibrio, risulta esponenzialmente grande in N , numero delle molecole. Si noti che, essendo $N \approx N_0 \cong 10^{24}$, qualunque sia il valore di τ , il tempo di ricorrenza $\langle \mathcal{T} \rangle = \tau \times O(10^{10^{24}})$ risulta decisamente “non fisico”. Per esempio, il più piccolo intervallo di tempo che si conosca in fisica è il tempo di Planck $t_P = \sqrt{\hbar G/c^5} \cong 5 \times 10^{-44}$ s; ma anche ponendo $\tau = t_P$ resta $\langle \mathcal{T} \rangle = O(10^{10^{24}})$ s, che si può confrontare con l'età dell'universo, che è circa 10^{17} s. L'allontanamento dall'equilibrio mediante il ritorno allo stato iniziale (alla base del paradosso della ricorrenza) non sembra un'impossibilità “di principio”, ma “di fatto”.

1.4 Ritorno alla Meccanica Statistica

Ritorniamo ai due gas del nostro esempio e riconsideriamo i paradossi alla luce dei risultati probabilistici sopra descritti.

Per prima cosa è opportuno introdurre in modo più esplicito una nozione di probabilità nello spazio delle fasi del sistema, usando la cosiddetta *ipotesi ergodica*.

Il sistema si muove in una regione limitata dello spazio delle fasi, quella in cui l'energia vale E_{tot} , che indichiamo con \mathcal{R} . Indichiamo con \mathcal{V} il volume che \mathcal{R} occupa nello spazio delle fasi.

Possiamo supporre che per le equazioni del moto (1.4) tutti gli stati nella regione con data E_{tot} siano equivalenti, cioè che il sistema durante l'evoluzione abbia accesso a tutta la regione \mathcal{R} , senza preferenze, e quindi *trascorra tempi uguali in regioni di uguale volume*: questo è uno dei modi di formulare l'*ipotesi ergodica*.

Dato che non esistono zone privilegiate né zone proibite dalla dinamica, dobbiamo assumere che, partendo da uno stato qualunque in \mathcal{R} , il sistema prima di ritornare vicino allo stato di partenza abbia esplorato una gran parte di \mathcal{R} . Se il sistema tornasse "troppo presto", e quindi troppo frequentemente, intorno a un certo stato, questo sarebbe uno stato diverso dagli altri e quindi, contrariamente all'ipotesi, "privilegiato". Il tempo necessario per il ritorno (il *tempo di ricorrenza* di un ciclo di Poincaré) dipende di volta in volta dallo stato esatto di partenza e da come è definita la vicinanza; indichiamo con \mathcal{T} il tempo tipico (opportunitamente definito) per questo moto attraverso \mathcal{R} . Possiamo dividere la regione \mathcal{R} , in tante regioni r_i di volume v_i , indicando con T_i la parte del tempo \mathcal{T} che il sistema trascorre in r_i . Deve essere $\sum_i r_i = \mathcal{R}$, $\sum_i v_i = \mathcal{V}$ e $\sum_i T_i = \mathcal{T}$. L'ipotesi ergodica permette di scrivere

$$\frac{T_i}{\mathcal{T}} = \frac{v_i}{\mathcal{V}}. \quad (1.17)$$

Adottando il punto di vista frequentistico della probabilità, interpretiamo la frazione di tempo trascorso in una certa regione, T_i/\mathcal{T} , come la probabilità, p_i , di trovare il sistema in quella regione durante l'evoluzione. Poniamo quindi

$$p_i = \frac{T_i}{\mathcal{T}} = \frac{v_i}{\mathcal{V}}. \quad (1.18)$$

Le regioni r_i sono scelte arbitrariamente e possono essere pensate "piccole" o anche infinitesime, in modo da disporre di probabilità relative a configurazioni microscopiche di un sistema, e poter eventualmente definire una densità di probabilità (densità di probabilità "microcanonica", vedi il Capitolo seguente). Quando invece si abbia interesse solo a proprietà macroscopiche di un sistema come nel caso del nostro esempio, è conveniente e sufficiente pensare a delle regioni anche "grandi" e alle relative probabilità.

Nello spazio delle fasi del sistema composto $A + B$, gli stati con energia totale E_{tot} occupano una regione di volume \mathcal{V} che si può dividere in n sottoregioni, distinte dalla diversa distribuzione dell'energia E_{tot} tra i due gas (questo è l'analogo della divisione dell'ipercubo di volume 1 nelle classi coi diversi valori di Y_N). La regione in cui si ha $E_{tot} \cong E_A^{(k)} + E_B^{(k)}$ ha volume \mathcal{V}_k . Deve essere $\sum_{k=1}^n \mathcal{V}_k = \mathcal{V}$, ma, essendo la dimensione dello spazio delle fasi pari a $6N$, con $N \approx N_0 \gg 1$, ci aspettiamo che in questo spazio di alta dimensionalità (come visto al punto 2 della

Sezione 1.3.1) esista una regione, che contrassegnamo con l'indice \hat{k} , il cui volume sovrasta enormemente quello delle altre, cosicché, per $N \rightarrow \infty$, si abbia

$$\frac{\mathcal{V}_{k \neq \hat{k}}}{\mathcal{V}_{\hat{k}}} \rightarrow 0 \quad \text{e} \quad \frac{\mathcal{V}_{\hat{k}}}{\mathcal{V}} \rightarrow 1. \quad (1.19)$$

Se $\mathcal{T}^{(k)}$ è il tempo che il sistema trascorre nella regione \mathcal{R}_k di volume \mathcal{V}_k , interpretiamo la frazione $\mathcal{T}^{(k)}/\mathcal{T}$ come la probabilità, $P^{(k)}$, di trovare il sistema nella regione di volume \mathcal{V}_k , in cui i gas A e B hanno energie $E_A^{(k)}$ ed $E_B^{(k)}$. Usando l'ipotesi ergodica, si ha:

$$P^{(k)} = \mathcal{T}^{(k)}/\mathcal{T} = \mathcal{V}_k/\mathcal{V}. \quad (1.20)$$

Esplicitamente: stiamo assumendo che, eseguendo misure di energia sul sistema, la probabilità di trovare i valori $E_A^{(k)}$ ed $E_B^{(k)}$ per i due sottosistemi, è pari alla frazione del volume dello spazio delle fasi occupato dagli stati microscopici con questi valori delle energie.

Se quella esposta sopra e condensata nelle relazioni (1.19) è una buona descrizione dello spazio delle fasi in cui si evolve il sistema, segue immediatamente che i due gas passeranno quasi tutto il loro tempo (tutto, se $N \rightarrow \infty$) nella regione di volume preponderante, $\mathcal{V}_{\hat{k}}$, in cui il gas A ha energia $E_A^{(\hat{k})} + \delta_A(t)$ e il gas B ha energia $E_B^{(\hat{k})} + \delta_B(t)$. In questa regione si ha $E_A^{(\hat{k})} = O(N_A)\bar{\epsilon}$, $E_B^{(\hat{k})} = O(N_B)\bar{\epsilon}$, $\delta_A(t) = O(\sqrt{N_A})\bar{\epsilon}$, $\delta_B(t) = O(\sqrt{N_B})\bar{\epsilon}$. Nella pratica sperimentale l'incertezza sulla misura dell'energia, Δ , è una frazione piccola dell'energia (per esempio, con un certo ottimismo, $\Delta \approx O(10^{-5})E \approx O(10^{-5})N_0\bar{\epsilon} \approx O(10^{19})\bar{\epsilon}$), ma comunque molto maggiore delle fluttuazioni tipiche dell'energia, sugli stati che il sistema visita durante l'evoluzione dinamica, $\delta(t) \approx O(\sqrt{N_0})\bar{\epsilon} \approx O(10^{12})\bar{\epsilon}$. I valori misurati $E_A^{(\hat{k})}$ ed $E_B^{(\hat{k})}$, appariranno quindi costanti *per la maggior parte del tempo* (praticamente infinito), e possiamo identificarli come i valori di equilibrio delle energie: $E_A^{(\hat{k})} = E_A^{(eq)}$, $E_B^{(\hat{k})} = E_B^{(eq)}$.

1-Equilibrio e Ricorrenza. Supponiamo che i gas siano inizialmente isolati, con energie E_{0A} ed E_{0B} , corrispondenti a valori di non equilibrio per il sistema $A+B$ di energia totale $E_{tot} = E_{0A} + E_{0B}$. Quindi, nella fase iniziale di preparazione gli stati microscopici dei gas si trovano in una regione \mathcal{R}_k , con $k \neq \hat{k}$, il cui volume è $\mathcal{V}_k = \mathcal{V}_{(A)} \times \mathcal{V}_{(B)}$. Se i gas sono mantenuti in isolamento, i loro stati restano in \mathcal{R}_k . Si noti che per i gas isolati la regione iniziale di volume $\mathcal{V}_{(A)} \times \mathcal{V}_{(B)}$ ha probabilità 1 rispetto allo spazio delle fasi permesso, di volume \mathcal{V}_k . Se ai gas si permette di interagire, l'intero spazio delle fasi del sistema con energia E_{tot} , di volume \mathcal{V} , diventa disponibile ed essi, a questo punto, si trovano in uno stato energetico di probabilità $\mathcal{V}_k/\mathcal{V} \approx 0$ (per ipotesi $k \neq \hat{k}$ e valgono le (1.19)). È chiaro che, se non c'è alcuna proibizione, come richiede l'ipotesi ergodica, il sistema entrerà rapidamente nella regione di volume $\mathcal{V}_{(\hat{k})}$ dove le energie assumono i valori $E_A^{(\hat{k})}$ ed $E_B^{(\hat{k})}$, vedi Fig.1.2. Per quanto argomentato sopra, la probabilità di trovare i gas con queste energie è

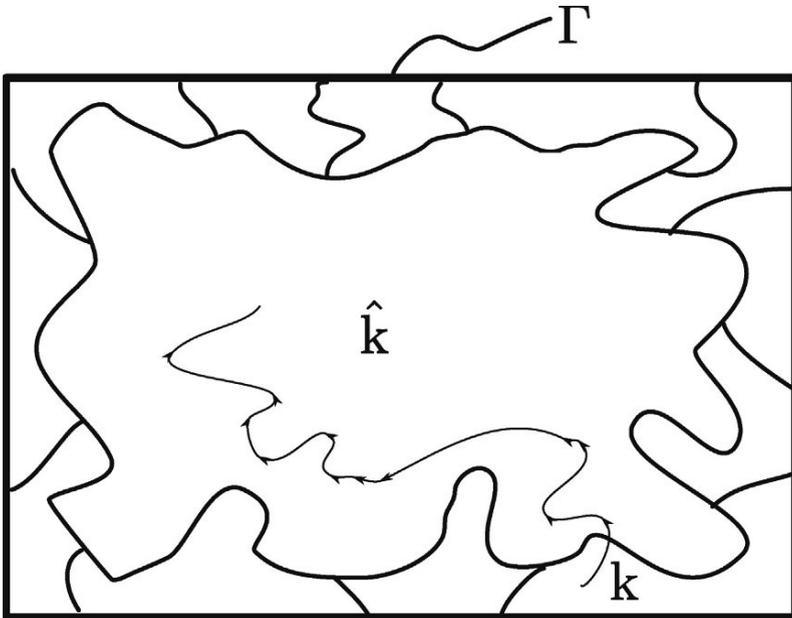


Figura 1.2 Una rappresentazione pittorica dello spazio Γ . In ognuna delle regioni l'energia totale è $E_{tot} = E_A^{(k)} + E_B^{(k)}$. La regione \hat{k} è di gran lunga la più estesa. Una coppia di sistemi parte dalla condizione di non equilibrio $k \neq \hat{k}$ e, nell'evoluzione successiva alla rimozione dei vincoli, rapidamente raggiunge la regione \hat{k} dove passa la maggior parte del tempo.

adesso $P^{(\hat{k})} = \mathcal{V}_{(\hat{k})} / \mathcal{V} \approx 1$; dal punto di vista dei tempi: le energie misurate saranno $E_A^{(\hat{k})}$ ed $E_B^{(\hat{k})}$ per un tempo praticamente infinito. Quest'ultima osservazione permette di definire $E_A^{(\hat{k})}$ ed $E_B^{(\hat{k})}$ come energie di equilibrio, e costituisce anche il modo in cui la Meccanica Statistica è in grado di liberarsi del paradosso della ricorrenza. Prima che il sistema possa ritornare nella regione iniziale \mathcal{R}_k si deve attendere che abbia visitato $\mathcal{R}_{\hat{k}}$, regione nella quale il sistema trascorre un tempo $\mathcal{T}^{(\hat{k})} \sim \mathcal{T} = O(10^{N_0})$, come si ricava dalle relazioni (1.20-1.19) e dal fatto che i tempi di ricorrenza crescono esponenzialmente col numero di molecole, secondo quanto suggerito nel punto 4) della Sezione 1.3.1. La ricorrenza deve esserci perché un teorema dimostrato in modo rigoroso lo richiede, ma scopriamo che per i sistemi macroscopici, di cui si occupa la termodinamica, il tempo tipico per osservarla, su cui il teorema non si pronuncia, è enormemente lungo anche adottando scale di tempo astronomiche.

2-Equilibrio e Reversibilità. Per quanto concerne il paradosso della reversibilità, consideriamo ancora la situazione precedente. Inizialmente i due gas si trovano in uno stato microscopico contenuto nella regione \mathcal{R}_k ; in seguito all'evoluzione, in un tempo t , il sistema raggiunge uno stato contenuto nella regione di equilibrio $\mathcal{R}_{\hat{k}}$. Se adesso si è in grado di invertire *esattamente* le velocità di tutte le molecole, mantenendole nelle loro posizioni, sappiamo di poter rimandare il sistema fuori equilibrio,

violando il secondo principio della termodinamica. L'inversione delle velocità va realizzata esattamente perché in un generico sistema dinamico il moto è caotico⁴: se le nuove condizioni iniziali (quelle con le velocità invertite e le posizioni inalterate) non sono quelle esatte, l'evoluzione del sistema si allontanerà rapidamente da quella desiderata, e il sistema non tornerà in \mathcal{R}_k ma resterà in $\mathcal{R}_{\hat{k}}$. Si noti che, anche se si riuscisse a far tornare il sistema nello stato iniziale con velocità invertite, da quel momento in poi, per gli stessi argomenti esposti sopra, avrebbe luogo un processo di ritorno all'equilibrio: si sarebbe realizzata solamente una fluttuazione macroscopica fuori dall'equilibrio. Il grosso ostacolo tuttavia è che non si è in grado di effettuare l'inversione delle velocità, neanche approssimativamente.

Non essendo praticamente in grado di compiere l'inversione, potremmo pensare di affidarci al caso per violare il secondo principio. Ci dobbiamo allora chiedere qual è la probabilità che, preparando uno stato iniziale di equilibrio, cioè contenuto in $\mathcal{R}_{\hat{k}}$, questo coincida con uno stato che è l'evoluto di uno proveniente da una qualunque regione \mathcal{R}_k , ma con le velocità invertite. Ricordiamo che per i sistemi hamiltoniani vale il teorema di Liouville: il volume occupato da un dato insieme di stati si mantiene inalterato durante l'evoluzione. Gli stati provenienti dalla regione \mathcal{R}_k occuperanno quindi un volume \mathcal{V}_k all'interno di $\mathcal{V}_{\hat{k}}$, e un ugual volume \mathcal{V}_k occuperanno anche gli stati con le velocità invertite. Dalle relazioni (1.19) segue però che la frazione del volume \mathcal{V}_k occupato dagli stati "invertiti" all'interno di $\mathcal{R}_{\hat{k}}$ tende a zero (esponenzialmente in N) quando N diventa molto grande; e con essa tende a zero la probabilità di preparare casualmente uno stato che abbandona l'equilibrio. Dato che l'insieme degli stati "invertiti" sono quelli che realizzano le fluttuazioni macroscopiche dell'energia, e sono una frazione trascurabile del volume della regione di equilibrio, dobbiamo pensare che la grande maggioranza degli stati in $\mathcal{R}_{\hat{k}}$ appartiene a traiettorie che descrivono piccole fluttuazioni di energia intorno ai valori di equilibrio, per un tempo praticamente infinito. Anche per la reversibilità vale la conclusione del punto precedente: possiamo osservarla se ripetiamo l'esperimento un numero molto grande di volte, di ordine 10^{N_0} .

⁴ Come Poincaré aveva già compreso verso il 1880, anche in sistemi regolati da leggi deterministiche, in cui lo stato futuro del sistema è univocamente determinato dal suo stato presente, si può avere un comportamento apparentemente irregolare. Inoltre una piccola incertezza iniziale δ_0 può crescere molto velocemente riducendo drasticamente la possibilità di fare previsioni:

$$\delta_t \sim \delta_0 e^{\lambda t},$$

dove λ , chiamato esponente di Lyapunov, dipende dal sistema e non dalla nostra bravura. L'amplificazione esponenziale dell'incertezza iniziale, cioè la sensibile dipendenza dalle condizioni iniziali (o *effetto farfalla*), implica che il sistema può essere predetto con una tolleranza Δ solo fino al tempo di predicibilità T_p che dipende dall'incertezza iniziale e dall'esponente di Lyapunov $\lambda > 0$ nel modo seguente:

$$T_p \sim \frac{1}{\lambda} \ln\left(\frac{\Delta}{\delta_0}\right).$$

Poiché la funzione logaritmo cresce molto lentamente, il tempo di predicibilità è determinato sostanzialmente dall'esponente di Lyapunov e poco dallo sforzo (e dal costo) per determinare la condizioni iniziali con grande precisione (cioè per diminuire δ_0).

È importante sottolineare che i comportamenti caotici non costituiscono una patologia: sono la regola e non l'eccezione essendo onnipresenti in geofisica, astronomia, ottica, biologia, chimica etc.

Concludiamo sottolineando che stiamo discutendo la difficoltà nell'effettuare una inversione dello stato che descrive il sistema con un'indeterminazione arbitrariamente piccola, ovvero con grande dettaglio. Se a questo si rinuncia e ci si limita a una descrizione parziale del sistema, in opportune condizioni si è in grado di realizzare la reversibilità meccanica. Per fare un esempio, un caso limite e banale è quello di una sfera che sia fatta cadere verticalmente da una certa altezza e urti elasticamente su una parete orizzontale, che funge da "invertitore" del moto: trascurando l'attrito, la sfera tornerà nel punto di partenza con velocità invertita (eventualmente nulla) con un tempo di risalita pari a quello di discesa. La inversione realizzata non riguarda il moto di tutte le molecole di cui è costituita la sfera, ma solamente il moto del suo centro di massa, che è un particolare "modo normale", completamente disaccoppiato dalle restanti variabili che tengono conto della dinamica "interna". In alcune situazioni possono esistere insiemi di modi normali, con cui si tiene (parzialmente) conto anche della dinamica interna, oppure insiemi di variabili legate ai momenti di spin delle particelle, che risultano caratterizzati da tempi di rilassamento all'equilibrio abbastanza lunghi; ciò permette di effettuare esperimenti che realizzano l'inversione del moto di quelle particolari variabili, con effetti decisamente molto suggestivi. Tali comportamenti sono detti *echo phenomena*. Il primo esperimento di questo tipo fu quello di "spin echo", che coinvolge momenti di dipolo magnetici nucleari, realizzato nel 1950 da E. L. Hahn.

1.5 Ruolo della dinamica

Le considerazioni precedenti sono basate essenzialmente sulla statistica degli stati nello spazio delle fasi del sistema esaminato. Dopo averli contati e suddivisi in classi per mezzo di variabili macroscopiche, abbiamo discusso degli argomenti che suggeriscono che dobbiamo attenderci un certo comportamento conforme alle leggi della termodinamica. In particolare abbiamo osservato che, se inizialmente i due gas si trovano in una regione di volume "piccolo", ci dobbiamo attendere che l'evoluzione li porti nella regione di volume dominante, realizzando così il raggiungimento dell'equilibrio. Affinché questa congettura, che sembra ragionevolmente motivata, non resti un semplice desiderio, c'è bisogno di mostrare che, almeno in qualche caso semplificato, il comportamento atteso sia una conseguenza delle leggi della dinamica.

Un primo passo in questa direzione fu compiuto da Boltzmann stesso con la sua equazione del trasporto (nota anche semplicemente come *equazione di Boltzmann*) che sarà discussa più accuratamente nel Capitolo 4.

L'idea di base è quella di studiare come si comporta, in funzione del tempo, una particolare variabile di tipo "collettivo": $f_X(\mathbf{q}, \mathbf{p})$, la distribuzione delle molecole nei vari stati di singola molecola. Per definizione, $f_X(\mathbf{q}, \mathbf{p}) d\mathbf{q} d\mathbf{p}$ dà il numero di molecole che, in uno stato \mathbf{X} di un sistema, si trovano in uno degli stati contenuti in un volumetto $d\mathbf{q} d\mathbf{p}$ intorno al punto (\mathbf{q}, \mathbf{p}) dello spazio delle fasi di singola molecola. Sottolineiamo che la funzione introdotta, come sta ad indicare il pedice X , è definita sugli stati $\mathbf{X} \equiv (\mathbf{q}_1, \dots, \mathbf{q}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$ di un sistema di N molecole. Essen-

do la f_X una densità nello spazio delle fasi di singola molecola, $f_X(\mathbf{q}, \mathbf{p})/N$ fornisce una densità di probabilità “empirica”, dipendente dallo stato microscopico \mathbf{X} . La f_X dipende dal tempo, in quanto funzione dello stato microscopico del sistema \mathbf{X} , la cui evoluzione è guidata dalle equazioni hamiltoniane (1.4).

Si noti che $f_X d\mathbf{q} d\mathbf{p}$ e $f_X d\mathbf{q} d\mathbf{p}/N$ (definite per ogni stato di N molecole) sono gli analoghi, nel continuo, degli insiemi $\{n_k\}$ e $\{v_k\}$, (introdotti nel punto 3 della Sezione 1.3.1 e definiti per ogni sequenza lunga N), per i quali vale la legge dei grandi numeri. Se quanto suggerisce l’analogia è corretto, ci aspettiamo

- 1) che lo spazio delle fasi del gas si possa suddividere in tante regioni, \mathcal{R}_f , caratterizzate dai valori che assume la f_X : tutti gli stati $\mathbf{X} \in \mathcal{R}_f$ condividono gli stessi valori di f_X , entro una tolleranza adeguatamente definita;
- 2) che esista una regione, $\mathcal{R}_{\hat{f}}$, di volume dominante (esponenzialmente in N) su tutti gli altri, tale che \hat{f} possa essere considerata la distribuzione di equilibrio.

Partendo da queste considerazioni puramente “statiche”, quello che si vorrebbe vedere è che le equazioni del moto fanno evolvere una $f_X(t)$ secondo le attese.

Per un gas diluito, cioè la cui densità è abbastanza bassa e la temperatura abbastanza alta, Boltzmann, partendo da un’equazione esatta per l’evoluzione di una generica $f_X(t)$ e introducendo l’ipotesi di *caos molecolare*, scrive la sua equazione del trasporto. Dall’equazione di Boltzmann si ottiene poi che a) esiste una soluzione stazionaria \hat{f} , data dalla distribuzione di Maxwell-Boltzmann; b) l’evoluzione porta ogni $f \neq \hat{f}$ verso \hat{f} .

L’equazione di Boltzmann descrive il comportamento “giusto” per la $f_X(t)$, ma è ottenuta al prezzo di una ipotesi extra (*caos molecolare*) non dinamica. Questo difetto è stato eliminato, negli anni ’70 del secolo passato, dal *teorema di Lanford*. Il teorema tratta le molecole come sfere dure e mostra che, in un limite opportuno per descrivere il gas diluito (limite di Boltzmann-Grad), l’evoluzione di una data f_X , almeno per tempi brevi, è guidata *esattamente* dall’equazione di Boltzmann *per la maggior parte* degli stati \mathbf{X} che essa rappresenta: la frazione di stati per cui ciò è vero può essere resa vicina a 1 quanto si vuole, prendendo il numero di molecole abbastanza grande.

Tutto questo mostra che lo scenario che abbiamo costruito per spiegare il comportamento termodinamico dei sistemi macroscopici a partire dalle equazioni microscopiche ha fondamento teorico rigoroso, almeno per un particolare sistema e in un certo limite.

1.6 Letture consigliate

In questo capitolo iniziale diamo qualche indicazione generale sui tanti buoni libri di meccanica statistica. Ne citiamo solo alcuni molto completi e particolarmente adatti per uno studio sistematico:

- M. Le Bellac, F. Mortessagne and G.G. Batrouni *Equilibrium and non-equilibrium statistical mechanics* (Cambridge University Press, 2004);

- L. Peliti *Appunti di Meccanica Statistica* (Bollati Boringhieri, Torino 2009);
- L.E. Reichl *A modern Course in Statistical Physics* (Wiley, 1998);
- J.P. Sethna *Statistical Mechanics: Entropy, Order Parameters and Complexity* (Oxford University Press, 2005).

Segnaliamo inoltre alcuni testi che non sono proprio dei manuali, ma possono essere utili per riflettere su alcuni aspetti della meccanica statistica:

- S.K. Ma *Statistical Mechanics* (World Scientific, Singapore 1985)
Discute in modo molto profondo e sottile i concetti chiave della meccanica statistica. Lo stile è discorsivo ed il formalismo matematico non particolarmente avanzato, questo non deve trarre in inganno. Non è affatto un libro facile, al lettore è richiesto uno sforzo notevole.
- G. Gallavotti *Statistical mechanics. A short treatise* (Springer-Verlag, Berlin 1995)
Raccoglie una serie di contributi molto accurati (e spesso con un approccio non convenzionale) scritti originariamente per l'Enciclopedia Treccani. Un testo non semplice che ripaga l'impegno richiesto per la sua lettura.
- G.G. Emch and C. Liu *the Logic of the Thermostatistical Physics* (Springer-Verlag, Berlin 2001)
Contiene una dettagliatissima analisi dei molti aspetti fondamentali della meccanica statistica ed anche del calcolo delle probabilità.
- P. Castiglione, M. Falcioni, A. Lesne and A. Vulpiani *Chaos and coarse graining in statistical mechanics* (Cambridge University Press, 2008)
Vengono discussi il ruolo del caos, della procedura di coarse graining e dei tanti gradi di libertà per la validità della meccanica statistica.

Dalla Meccanica alla Termodinamica

Sommario Nel capitolo precedente abbiamo esposto degli argomenti a sostegno della possibilità che la Termodinamica non sia in contraddizione con la Meccanica Classica. Il comportamento di tipo termodinamico, in cui un sistema raggiunge uno stato nel quale le osservabili che lo descrivono non dipendono dal tempo, sembra riconducibile al fatto che quelle osservabili siano variabili macroscopiche, ottenute sommando i contributi di un numero enorme di quantità microscopiche. In pratica, in un sistema macroscopico, analogamente a quanto avviene con la legge dei grandi numeri, le osservabili termodinamicamente rilevanti assumono valori approssimativamente costanti nella quasi totalità dello spazio delle fasi in cui il moto si sviluppa. Dal punto di vista della meccanica l'equilibrio termodinamico corrisponde a una situazione in cui le osservabili macroscopiche subiscono variazioni "molto piccole", tipicamente non rivelabili (quindi, per molti scopi, trascurabili), durante un intervallo di tempo "molto lungo" (e quindi praticamente infinito).

Per uscire dal piano puramente qualitativo e sviluppare una teoria sistematica, riconsideriamo un sistema di N particelle e, seguendo le idee di L. Boltzmann, in questo capitolo discutiamo l'ipotesi dell'ergodicità, e introduciamo la densità di probabilità microcanonica insieme alla relazione che definisce l'entropia in meccanica statistica.

2.1 Ipotesi ergodica e densità microcanonica

Indicando con \mathbf{q}_i e \mathbf{p}_i , rispettivamente, il vettore posizione e il vettore impulso della i -ma particella, lo stato di un sistema di N particelle è rappresentato, al tempo t , da un vettore $\mathbf{X}(t) \equiv (\mathbf{q}_1(t), \dots, \mathbf{q}_N(t), \mathbf{p}_1(t), \dots, \mathbf{p}_N(t))$ in uno spazio di dimensione $6N$ detto spazio delle fasi. Le osservabili del sistema sono rappresentate da funzioni, $A(\mathbf{X})$, definite nello spazio delle fasi. Le particelle sono soggette alle leggi deterministiche della meccanica classica e quindi $\mathbf{X}(t)$ si evolve in accordo con le equazioni di Hamilton. Se la funzione Hamiltoniana non dipende esplicitamente dal tempo, come assumeremo sempre nel seguito, allora l'energia è una quantità conservata

durante il moto, il quale quindi si sviluppa su una ipersuperficie a energia fissata. Indicando con $V(\{\mathbf{q}_j\})$ il potenziale di interazione tra le particelle, l'Hamiltoniana ha la forma

$$H = \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{2m} + V(\{\mathbf{q}_j\}), \quad (2.1)$$

e le equazioni di evoluzione sono:

$$\begin{aligned} d\mathbf{q}_i/dt &= \partial H / \partial \mathbf{p}_i = \mathbf{p}_i/m \\ d\mathbf{p}_i/dt &= -\partial H / \partial \mathbf{q}_i = -\partial V / \partial \mathbf{q}_i \end{aligned} \quad (2.2)$$

con $i = 1, \dots, N$. Supponiamo di misurare un'osservabile del sistema, che si trova in equilibrio termodinamico. È fondamentale notare che la scala dei tempi macroscopici, quelli delle osservazioni sul sistema, è molto più grande della scala dei tempi della dinamica microscopica (2.2), quelli che dettano la rapidità dei cambiamenti a livello molecolare. Ciò significa che un dato sperimentale in pratica è il risultato di un'unica osservazione durante la quale, in realtà, il sistema passa attraverso un grandissimo numero di stati microscopici diversi. Se il dato si riferisce all'osservabile $A(\mathbf{X})$, esso va quindi confrontato con una media eseguita lungo l'evoluzione del sistema e calcolata su tempi molto lunghi (dal punto di vista microscopico):

$$\bar{A}(t_0, \mathcal{T}) = \frac{1}{\mathcal{T}} \int_{t_0}^{t_0+\mathcal{T}} A(\mathbf{X}(t)) dt. \quad (2.3)$$

Ad esempio quando misuriamo la pressione con un manometro il risultato che leggiamo sul display è la media su un tempo \mathcal{T} dell'impulso trasferito nell'unità di tempo su una superficie unitaria, con \mathcal{T} dipendente dalle caratteristiche dello strumento ma molto maggiore dei tempi molecolari.

Il calcolo della media temporale $\bar{A}(t_0, \mathcal{T})$ di un'osservabile qualunque richiede, in linea di principio, sia la conoscenza dello stato microscopico del sistema ad un certo istante, sia la determinazione della corrispondente traiettoria nello spazio delle fasi. La richiesta è evidentemente inesaudibile per cui, se $\bar{A}(t_0, \mathcal{T})$ dipendesse in maniera molto forte dallo stato iniziale del sistema, non si potrebbero fare previsioni utili neppure di tipo statistico, anche trascurando la difficoltà di integrare il sistema (2.2).

L'*ipotesi ergodica* indica una via per superare questo ostacolo. Essa sostanzialmente afferma che ogni ipersuperficie di energia fissata è completamente accessibile a qualunque moto con la data energia; ovvero: una ipersuperficie di energia costante non può essere suddivisa in regioni di misura finita contenenti ognuna moti completi, cioè regioni invarianti per evoluzione temporale (se questa condizione è soddisfatta la ipersuperficie si dice 'metricamente non decomponibile' o 'metricamente transitiva'). Inoltre, per ogni traiettoria il tempo medio di permanenza in una certa regione è proporzionale al volume della regione.

Se le condizioni precedenti, che costituiscono appunto il nucleo dell'ipotesi ergodica, sono soddisfatte, segue che, per \mathcal{T} sufficientemente grande, la media in (2.3) dipende solo dall'energia del sistema e assume quindi lo stesso valore su tutte

le evoluzioni con uguale energia; inoltre, questo valore comune è calcolabile eseguendo una media di $A(\mathbf{X})$ in cui tutti (e solamente) gli stati con la fissata energia contribuiscono con uguale peso. Nelle applicazioni, tenendo conto del fatto che l'energia di qualunque sistema è determinata con un'incertezza finita, è comodo considerare nella media tutti gli stati con energia compresa in un intervallo fissato. La densità di probabilità uniforme nella regione con energia fissata a meno di un'incertezza Δ definisce la *densità microcanonica*, o *insieme microcanonico*. Indicando tale densità con $\rho_{mc}(\mathbf{X})$, e l'elemento di volume dello spazio delle fasi con $d\mathbf{X} = dq_1 \cdots dq_N dp_1 \cdots dp_N$ (che, ricordiamo, è invariante nel moto), si ha

$$\rho_{mc}(\mathbf{X}) = \left(\int_{E \leq H(\mathbf{X}) \leq E + \Delta} d\mathbf{X} \right)^{-1} \equiv (\Gamma_{\Delta}(E))^{-1}$$

e l'ipotesi ergodica permette di scrivere:

$$\bar{A} \equiv \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_{t_0}^{t_0 + \mathcal{T}} A(\mathbf{X}(t)) dt = \int A(\mathbf{X}) \rho_{mc}(\mathbf{X}) d\mathbf{X} \equiv \langle A \rangle. \quad (2.4)$$

Va sottolineato che la precedente equazione, se valida, ci libera contemporaneamente dalla necessità di determinare uno stato (iniziale) del sistema, di risolvere le equazioni del moto e di effettuare l'integrale nel tempo. La validità, o meno, della (2.4), cioè la possibilità di sostituire la media di un'osservabile qualunque lungo un'evoluzione temporale con una media dell'osservabile nello spazio delle fasi, costituisce il *problema ergodico*. Poiché se un sistema isolato in equilibrio risulta descrivibile mediante l'insieme microcanonico, non è difficile mostrare, per esempio, che un sistema in contatto con un termostato è ben descritto dall'insieme canonico, la dimostrazione della (2.4) può essere ritenuta la legittimazione dinamica dell'introduzione degli insiemi statistici.

2.2 Commenti e osservazioni sull'ergodicità

A) – Il problema ergodico nasce, insieme all'ipotesi ergodica, dalle idee di L. Boltzmann sulla meccanica statistica ed è stato in seguito studiato in termini matematici generali soprattutto da J. von Neumann e G. D. Birkhoff.

L'“ipotesi ergodica” originale di Boltzmann era la seguente: la superficie di energia costante è composta di una quantità enorme (ma finita) di celle, che possono essere numerate; durante l'evoluzione temporale una traiettoria passa attraverso tutte le celle; ciò fornisce la possibilità di sostituire una media nel tempo con una media nello spazio delle fasi.

Una interpretazione, dovuta agli Ehrenfest, nel contesto usuale della meccanica classica nello spazio delle fasi continuo, ha riformulato l'ipotesi originale nell'ipotesi che una traiettoria evolventesi su una superficie di energia costante finisce per visitare tutti i suoi punti.

La (quasi ovvia) impossibilità per una singola traiettoria di visitare ogni punto di una (iper)superficie ha poi condotto gli Ehrenfest alla formulazione della cosiddet-

ta “ipotesi quasi ergodica”, secondo la quale ogni evoluzione su una superficie di energia assegnata copre densamente la superficie stessa, con esclusione, al più, di un insieme di punti iniziali di misura nulla.

La moderna teoria ergodica può essere considerata una branca della teoria della misura e dell’integrazione, con obiettivi che vanno ben al di là del problema originale formulato da Boltzmann nel contesto della meccanica statistica. Il problema ergodico adesso può essere posto nei termini seguenti. Si considera un sistema dinamico, cioè una legge di evoluzione deterministica in uno spazio delle fasi Ω

$$\mathbf{X}(0) \rightarrow \mathbf{X}(t) = U^t \mathbf{X}(0) \quad (2.5)$$

e una misura $d\mu(\mathbf{X})$ invariante sotto l’evoluzione data da U^t , i.e. $d\mu(\mathbf{X}) = d\mu(U^{-t}\mathbf{X})$. Per esempio, nel caso di un sistema meccanico Hamiltoniano, Ω è lo spazio delle fasi descritto dalle variabili canoniche, U^t è un operatore che dipende da H e la legge (2.5) è un modo compatto di indicare le soluzioni delle equazioni di Hamilton (2.2), infine si avrebbe $d\mu(\mathbf{X}) = \rho_{mc}(\mathbf{X}) d\mathbf{X}$. Il sistema dinamico $(\Omega, U^t, d\mu(\mathbf{X}))$ si dice ergodico, rispetto alla misura $d\mu(\mathbf{X})$, se, per ogni funzione integrabile $A(\mathbf{X})$ e per quasi tutte (rispetto a μ) le condizioni iniziali $\mathbf{X}(t_0)$, si ha:

$$\bar{A} \equiv \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_{t_0}^{t_0 + \mathcal{T}} A(\mathbf{X}(t)) dt = \int A(\mathbf{X}) d\mu(\mathbf{X}) \equiv \langle A \rangle, \quad (2.6)$$

dove $\mathbf{X}(t) = U^{t-t_0} \mathbf{X}(t_0)$. La domanda che ci si pone è quindi sotto quali condizioni un sistema dinamico risulti ergodico. Il problema è stato affrontato da G. D. Birkhoff che ha dimostrato i seguenti teoremi. Posto $t_0 = 0$ e $\mathbf{X}(0) = \mathbf{X}_0$,

Theorem 2.1. *Per quasi tutte le condizioni iniziali \mathbf{X}_0 la media su tempo infinito*

$$\bar{A}(\mathbf{X}_0) \equiv \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} A(U^t \mathbf{X}_0) dt \quad (2.7)$$

esiste e il limite non dipende dalla scelta del punto iniziale su una data traiettoria.

Theorem 2.2. *Una condizione necessaria e sufficiente perché un sistema sia ergodico, cioè la media temporale $\bar{A}(\mathbf{X}_0)$ non dipenda dalla condizione iniziale (per quasi ogni \mathbf{X}_0), è che lo spazio delle fasi Ω sia metricamente indecomponibile (o metricamente transitivo). Quest’ultima proprietà significa che Ω non può essere suddiviso in due parti, ognuna di misura positiva, che siano invarianti rispetto alla dinamica U^t .*

Il teorema 2.1 è abbastanza generale e, almeno dal punto di vista della fisica, non molto stringente, infatti assicura che la media nel tempo $\bar{A}(\mathbf{X}_0)$ esiste, ma non garantisce la sua indipendenza dalle condizioni iniziali, al variare della traiettoria. Il risultato 2.2 è più interessante ma praticamente poco conclusivo per la meccanica statistica, dato che, in generale, non è possibile decidere se un dato sistema è metricamente indecomponibile. In pratica il teorema 2.2 si limita a spostare il problema.

B) – In realtà, per un sistema macroscopico il problema dell'ergodicità, oltre che di difficile soluzione, potrebbe essere sostanzialmente irrilevante nel contesto della meccanica statistica. A causa del grande numero di particelle, e quindi dell'enormità delle regioni di spazio delle fasi coinvolte, i tempi \mathcal{T} necessari perché le due medie nell' eq. (2.4) risultino confrontabili (ammesso che siano uguali), possono diventare molto più grandi dell'età dell' Universo, *nel caso di osservabili qualunque*. In questa situazione \mathcal{T} non avrebbe nessun interesse fisico, e ugualmente la relazione (2.4). Una questione importante è quindi quanto deve essere grande \mathcal{T} affinché $\overline{A}(t_0, \mathcal{T})$ risulti vicino a $\langle A \rangle$. Ci si aspetta che la risposta a questa domanda, in generale, possa dipendere sia dall'osservabile A che dal numero di particelle N . Come esempio si consideri la funzione

$$A_G(\mathbf{X}) = \begin{cases} 1 & \text{se } \mathbf{X} \in G \\ 0 & \text{altrimenti} \end{cases} \quad (2.8)$$

dove G è una regione contenuta nello spazio delle fasi accessibile a un sistema con una certa energia. $A_G(\mathbf{X})$ segnala la presenza del vettore di stato del sistema in G ; la sua media microcanonica è:

$$\langle A_G \rangle = \int_{E \leq H(\mathbf{X}) \leq E+\Delta} d\mathbf{X} \frac{A_G(\mathbf{X})}{\Gamma_\Delta(E)} = \frac{\Gamma(G)}{\Gamma_\Delta(E)}, \quad (2.9)$$

dove $\Gamma(G)$ indica il volume della regione G . Si noti che, per definizione di microcanonico, il volume relativo $\Gamma(G)/\Gamma_\Delta(E)$ è anche $P(G)$, la probabilità di trovare il sistema in G . La quantità $\overline{A_G}(\mathcal{T})$ è, in questo caso, la frazione di tempo che la traiettoria $\mathbf{X}(t)$ passa in G durante l'intervallo $[0, \mathcal{T}]$. Perché $\overline{A_G}(\mathcal{T})$ dia una stima credibile di $\Gamma(G)/\Gamma_\Delta(E)$ la traiettoria deve aver esplorato una buona parte del volume $\Gamma_\Delta(E)$. Se (in accordo con l'ipotesi ergodica) escludiamo che ci siano regioni preferite, il segnale che una "buona esplorazione" è stata compiuta è dato dal ritorno nella regione G , da cui possiamo assumere che il sistema sia partito. Una stima di \mathcal{T}_{eq} , tempo necessario per avere un buon accordo tra $\overline{A_G}(\mathcal{T})$ e $\langle A_G \rangle$, può quindi fornirla $\langle \mathcal{T}_G \rangle$, il tempo medio di ricorrenza in G . Possiamo farci un'idea su $\langle \mathcal{T}_G \rangle$ nel modo seguente.

Utilizzando il lemma di Kac¹ (vedi Appendice H), possiamo scrivere che il tempo medio di ricorrenza in G , $\langle \mathcal{T}_G \rangle$, è dato da $\tau/P(G)$:

$$\langle \mathcal{T}_G \rangle = \tau \frac{\Gamma_\Delta(E)}{\Gamma(G)}, \quad (2.10)$$

dove τ indica l'intervallo di tempo tra le osservazioni fatte sul sistema. Se si indica con μ_n , la proiezione di G sullo spazio a 6 dimensioni generato dalle variabili

¹ In questo caso il lemma va usato con cautela: per ipotesi, lo spazio delle fasi accessibile contiene energie diverse ed è quindi costituito da molte regioni invarianti, su ognuna delle quali si può applicare il lemma. Si può però pensare a una media dei tempi di ricorrenza medi.

$(\mathbf{q}_n, \mathbf{p}_n)$ descrittivi della n -ma particella, si ha

$$\Gamma(G) = \prod_{n=1}^N \mu_n . \quad (2.11)$$

Consideriamo un gas contenuto in un recipiente di volume V e poniamo $\mu_n = \varepsilon^3$ (ε ha le dimensioni di una azione e a ogni coppia di variabili canoniche è associato un fattore ε). Si ha quindi

$$\Gamma(G) = (\varepsilon)^{3N} . \quad (2.12)$$

Nell'approssimazione di gas ideale, in cui l'energia è tutta cinetica, il calcolo di $\Gamma_{\Delta}(E)$ si riduce al calcolo del volume di una sfera in $3N$ dimensioni², (dimensione dello spazio degli impulsi) e all'applicazione dell'approssimazione di Stirling³ per i fattoriali. Il risultato è

$$\Gamma_{\Delta}(E) \cong \frac{\Delta}{E} \sqrt{\frac{3N}{4\pi}} \left(V \left(\frac{4}{3} \pi m \varepsilon \frac{E}{N} \right)^{3/2} \right)^N . \quad (2.13)$$

Una prima osservazione da fare è che il fattore davanti all'esponenziale in N è irrilevante: il suo contributo, che si può riscrivere in forma esponenziale come $\exp[\ln(\Delta/E) \sqrt{3N/4\pi}]$, porta una correzione di $O(\ln N)$ al termine dominante con esponente $N = O(10^{24})$.

La seconda osservazione è che il volume accessibile al gas (la parte importante, escludendo la correzione che abbiamo visto essere trascurabile) può essere pensato come il prodotto, su tutte le particelle, dello spazio delle fasi disponibile a ognuna di esse. Se indichiamo con $e_1 = E/N$ l'energia per particella e calcoliamo il volume di spazio degli impulsi V_p accessibile alla singola particella (quello con energia fino a e_1), otteniamo $V_p = (2me_1)^{3/2} (4\pi/3)$. La (2.13) diventa

$$\Gamma_{\Delta}(E) \cong \left(\sqrt{\frac{e^3 \pi}{6}} V V_p \right)^N . \quad (2.14)$$

Tornando alla (2.10) si ha:

$$\langle \mathcal{T}_G \rangle \cong \tau \left(\sqrt{\frac{e^3 \pi}{6}} \left(\frac{V V_p}{\varepsilon^3} \right) \right)^N ; \quad (2.15)$$

vediamo quindi che il tempo di ricorrenza dipende dal rapporto tra il volume dello spazio delle fasi disponibile per ogni particella (dettato dalle condizioni esterne, V ed E/N) e il volume dello spazio delle fasi su cui si fissa l'attenzione (determinato da G). Ma, in realtà, il fattore N a esponente fa sì che la dipendenza di $\langle \mathcal{T}_G \rangle$ dal rapporto

² Il volume di una sfera di raggio R in d dimensioni è $Vol_d(R) = \pi^{d/2} R^d / \Gamma(\frac{d}{2} + 1)$, dove $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ indica la funzione Gamma di Eulero.

³ $\Gamma(x) \cong \sqrt{2\pi} x^{x-(1/2)} e^{-x}$.

sia estremamente debole: $\langle \mathcal{T}_G \rangle$ è enormemente grande, praticamente per ogni ε ; solamente nel caso $\varepsilon^3 \cong V V_p$ si può avere un tempo di ricorrenza non astronomico, ma questo significa che la regione che stiamo guardando è praticamente tutta quella accessibile: $\Gamma(G) \cong \Gamma_\Delta(E)$.

Nel caso di una generica osservabile $A(\mathbf{X})$ di un sistema macroscopico, per la quale le medie nel tempo e nello spazio delle fasi sono intercambiabili solo dopo un buon campionamento della regione accessibile, i tempi di ricorrenza enormi rendono praticamente inutilizzabile l'eq. (2.4). D'altra parte le considerazioni sui tempi di ritorno sono anche alla base della risposta data originariamente da Boltzmann alle critiche sulla validità del suo teorema H . Infatti, a causa dell'enorme numero di particelle presenti, in un sistema macroscopico anche il tempo di ricorrenza di uno stato di non equilibrio è inimmaginabilmente grande. Per esempio, Boltzmann stimò che, in una sfera di aria di raggio 1 cm alla temperatura di 300 K e a pressione standard, per ritornare a uno stato in cui la concentrazione delle molecole differisca dal valore medio per l'1% il tempo di attesa è $10^{10^{14}}$ secondi !!

I tempi esponenzialmente grandi (in N), che appaiono in questi due esempi, hanno la loro comune origine nella piccolezza (esponenziale) dei volumi relativi delle regioni coinvolte: quello dove $A(\mathbf{X})$ è differente da zero e quello dove lo stato del sistema è leggermente fuori equilibrio. Tuttavia questi tempi lunghi ci danno due informazioni molto differenti. Nel primo caso un tempo così grande non è positivo, perché qualifica come praticamente inutile l'eq. (2.4). Nel secondo caso un tempo così lungo è benvenuto, perché permette di introdurre la nozione di equilibrio di un sistema macroscopico in meccanica statistica.

C) – A questo punto è necessario ritornare sulle considerazioni fatte nel primo capitolo e sottolineare nuovamente che le osservabili rilevanti per la termodinamica, quelle con le quali sono caratterizzati gli stati di equilibrio, non sono funzioni generiche. Sono poche e soprattutto di un tipo particolare, cosicché la questione fisicamente interessante è se i tempi per raggiungere l'equilibrio (cioè l'uguaglianza delle medie nel tempo e nello spazio delle fasi) possano essere corti abbastanza per queste funzioni termodinamicamente interessanti. In effetti il recupero della (2.4) è possibile basandosi sulle seguenti considerazioni:

- a) nei sistemi termodinamici il numero di costituenti microscopici è molto grande;
- b) la questione interessante per la meccanica statistica è la validità della (2.4) non per osservabili generiche, bensì per le poche grandezze rilevanti nella termodinamica (per esempio, l'energia cinetica, la pressione, la densità) che hanno una struttura particolare, cioè sono esprimibili, esattamente o con buona approssimazione, come somma di contributi separati dovuti ai costituenti microscopici;
- c) è accettabile che l'equazione (2.4) possa non valere per condizioni iniziali contenute in regioni di misura complessivamente piccola (tendente a zero se $N \rightarrow \infty$);
- d) è accettabile che l'equazione (2.4) non sia vera esattamente ma solo approssimativamente, cioè

$$|\bar{A} - \langle A \rangle| < \varepsilon$$

con ε tendente a zero per $N \rightarrow \infty$.

A proposito del punto *b*), notiamo che la funzione definita in (2.8) può essere scritta anche:

$$A(\mathbf{X}) = \prod_n \chi_{\mu_n}(\mathbf{q}_n, \mathbf{p}_n),$$

dove $\chi_{\mu_n}(\mathbf{q}_n, \mathbf{p}_n)$ è la funzione caratteristica del sottoinsieme μ_n , ovvero $\chi_{\mu_n}(\mathbf{q}_n, \mathbf{p}_n) = 1$ se $(\mathbf{q}_n, \mathbf{p}_n) \in \mu_n$ e $\chi_{\mu_n}(\mathbf{q}_n, \mathbf{p}_n) = 0$ altrimenti. Questo rende evidente che una tale funzione non è di tipo additivo, ma moltiplicativo, cioè non rientra nella classe delle funzioni interessanti.

Tenendo conto di queste richieste fisiche, matematicamente meno vincolanti, si può pensare di ottenere risultati interessanti, anche se non così generali come i teoremi di Birkhoff, i quali valgono per sistemi dinamici generici, anche di bassa dimensionalità, per osservabili non specifiche, e per quasi tutte le condizioni iniziali.

Come esempio di conclusioni che si possono trarre da questa impostazione del problema, riportiamo i seguenti risultati, dovuti a A. J. Khinchin, che riguardano un gas ideale.

D) – Si consideri un sistema con Hamiltoniana separabile:

$$H = \sum_{n=1}^N H_n(\mathbf{q}_n, \mathbf{p}_n) \quad (2.16)$$

e una classe speciale di osservabili (dette *funzioni somma*) che sono esprimibili come somma di N componenti, dipendenti ognuna dalle variabili di una sola particella, cioè della forma

$$f(\mathbf{X}) = \sum_{n=1}^N f_n(\mathbf{q}_n, \mathbf{p}_n) \quad (2.17)$$

dove $f_n = O(1)$, e nel caso più semplice e frequente le componenti f_n sono tutte uguali. Esempi importanti di funzioni somma sono dati dalla pressione, l'energia cinetica, l'energia totale e la funzione di distribuzione di singola particella. Si noti che il cambiamento in una singola f_n , di $O(1)$, produce una variazione relativa $O(1/N)$ in $f(\mathbf{X})$: le funzioni somma sono “buone” funzioni macroscopiche, poiché non sono tanto sensibili ai dettagli microscopici.

Dato che l'Hamiltoniana è separabile si ha:

$$\langle f \rangle = O(N) \quad \text{e} \quad \sigma^2 = \langle (f - \langle f \rangle)^2 \rangle = O(N). \quad (2.18)$$

Si ricordi che $\langle \rangle$ indica la media sull'insieme microcanonico (stiamo considerando un sistema con energia assegnata). Considerando la media nel tempo $\bar{f}(\mathbf{X})$ di un'osservabile f , lungo una traiettoria che parte da \mathbf{X} , sotto ipotesi abbastanza generali (e senza invocare la transitività metrica, che entra invece nell'ipotesi ergodica) si ha, usando la disuguaglianza di Markov:

$$\text{Prob}(|\bar{f} - \langle f \rangle| > a) < \frac{4}{a} \langle |f - \langle f \rangle| \rangle < \frac{4}{a} \sigma, \quad (2.19)$$

dove l'ultima disuguaglianza è dovuta alla relazione (disuguaglianza di Lyapunov o di Schwartz)

$$\langle |f - \langle f \rangle| \rangle \leq \langle (f - \langle f \rangle)^2 \rangle^{1/2}.$$

La disuguaglianza (2.19) implica

$$\text{Prob} \left(\left| \frac{\bar{f}}{\langle f \rangle} - 1 \right| > \frac{a}{|\langle f \rangle|} \right) < \frac{4}{a} \sigma. \quad (2.20)$$

Se si sceglie $a = \sigma^{3/2}$ e si tiene conto della (2.18) si può scrivere

$$\text{Prob} \left(\left| \frac{\bar{f}}{\langle f \rangle} - 1 \right| > \frac{K_1}{N^{1/4}} \right) < \frac{K_2}{N^{1/4}} \quad (2.21)$$

dove K_1 e K_2 sono costanti $O(1)$. Si ha quindi che, per la classe delle funzioni somma l'insieme dei punti iniziali per i quali la media nel tempo differisce da quella nello spazio delle fasi più di una certa quantità, *che va a zero per $N \rightarrow \infty$* , costituisce una frazione dei punti con energia data che va a zero per $N \rightarrow \infty$. Sostanzialmente, in questo limite e per questa classe di funzioni la (2.4) è vera (tranne che in una zona dello spazio delle fasi, che è sempre più piccola all'aumentare di N) e questo indipendentemente dai dettagli della dinamica.

Spostando l'attenzione dalla ergodicità come proprietà del sistema, alla ergodicità come proprietà delle osservabili, si trova che, nei casi che interessano, è possibile sostituire una media temporale con la media microcanonica. Ciò però lascia ancora aperto il problema fondamentale dei tempi necessari per raggiungere l'equilibrio. Anche su questo punto si possono usare le assunzioni (2.16 - 2.17) e la (2.18) per mostrare che:

$$\text{Prob} \left(\left| \frac{f}{\langle f \rangle} - 1 \right| > \frac{K_3}{N^{1/4}} \right) < \frac{K_4}{N^{1/4}} \quad (2.22)$$

ovvero: le osservabili fisicamente rilevanti sono auto-medianti, cioè per grandi valori di N , le osservabili di questo tipo sono soggette alla legge dei grandi numeri e, su una ipersuperficie di data energia, assumono un valore pressoché costante (macroscopicamente), tranne che in regioni di misura complessiva molto piccola.

La relazione (2.22) è importante per due motivi. Da una parte assegna alla quantità $\langle f \rangle$ il ruolo di rappresentare il valore quasi costante (macroscopicamente) dell'osservabile $f(\mathbf{X})$, dandole un significato fisico indipendentemente da ogni media nel tempo. Dall'altra, poichè in queste condizioni è ragionevole supporre che il risultato di una media temporale non dipenda molto da \mathcal{T} né, generalmente, dallo stato di partenza, suggerisce che la relazione (2.4) possa valere anche per 'piccoli' \mathcal{T} , e cioè abbia interesse fisico, consentendo di collegare $\langle f \rangle$ a una misura di f su tempi finiti. Questi risultati confermano con adeguato rigore le considerazioni, per lo più qualitative, fatte nel primo capitolo. Assumeremo quindi che la media sulla distribuzione microcanonica fornisca i valori corretti delle osservabili termodinamiche che si misurano sul sistema in equilibrio.

Si può ritenere che l'essenza di questi risultati sia che la meccanica statistica basata sugli insiemi funziona indipendentemente dalla validità dell'ergodicità (in senso strettamente matematico). In effetti questo era anche il punto di vista di Boltzmann stesso.

E) – L'aspetto debole, dal punto di vista fisico, dell'approccio alla Khinchin riguarda l'assunzione di non interazione, eq. (2.16), in quanto un requisito essenziale per un comportamento termodinamico è la possibilità di scambiare energia tra particelle. Naturalmente a Khinchin il problema era presente essendo chiaro anche che una vera Hamiltoniana può essere al più solo approssimativamente separabile. L'idea di fondo è che l'interazione tra le particelle contribuisce molto poco nel calcolo delle grandezze macroscopiche all'equilibrio e che, nella maggior parte dei calcoli in meccanica statistica, questi termini possono essere trascurati.

La poco desiderabile restrizione sulla struttura separabile dell'Hamiltoniana, può essere superata. Mazur e van der Linden studiando sistemi di particelle interagenti tramite potenziali a corto raggio d'azione, hanno mostrato che il punto di vista di Khinchin sull'importanza dell'interazione tra particelle è essenzialmente corretto. Lasciando da parte gli aspetti tecnici del lavoro, l'interpretazione fisica del risultato è che, a causa della corta portata dell'interazione, un sistema di molte particelle si comporta come se fosse costituito da un grande numero di componenti non interagenti. Come scrivono Mazur e van der Linden: *One might think of subsystems consisting of large numbers of particles; the interaction between these subsystems is then a surface effect and very small compared to the energy content of the subsystems themselves.* I loro calcoli mostrano che *the energies of these subsystems behave as almost independent random variables, so that a central limit theorem still applies.* È anche interessante notare che il risultato ottenuto dai due autori è valido a esclusione, al più, di un numero finito di temperature, alle quali il sistema può subire una transizione di fase.

F) – È utile sottolineare due aspetti:

- nei risultati sopra esposti, oltre alla condizione che la portata dell'interazione sia limitata, la dinamica non ha un ruolo molto importante: per le osservabili interessanti, l'esistenza di buone proprietà statistiche su ogni superficie di energia assegnata, non dipende dai dettagli della dinamica, ma è legata al fatto che $N \gg 1$. Questo aiuta anche a capire perché normalmente si trascurino gli altri integrali del moto, che esistono in ogni sistema Hamiltoniano;
- il fatto che le osservabili interessanti abbiano valore (macroscopicamente) quasi costante *sulla maggior parte* della ipersuperficie di energia assegnata, ma *non su tutta*, conferma la descrizione qualitativa esposta nella Sezione 1.4 (vedi Fig. 1.2) e permette l'esistenza degli stati di non equilibrio in questo quadro concettuale. Il passo successivo, per dare sostanza all'idea di equilibrio e perché il quadro risulti coerente, sarà mostrare che, partendo da uno di questi stati di non equilibrio, la dinamica Hamiltoniana fa evolvere il sistema nella "direzione giusta" (equazione di Boltzmann, vedi capitolo 4).

2.3 L'entropia

Abbiamo introdotto l'insieme microcanonico per descrivere l'equilibrio di un sistema isolato; immaginiamo, ad esempio, che il sistema sia un gas il cui stato termodinamico è determinato dal valore dell'energia (E) e, poniamo, dal volume (V) e dal numero di particelle (N). Per stabilire un legame tra descrizione statistica e termodinamica, è allora utile identificare, tra tutti i potenziali termodinamici, l'entropia (S), perché avendo la sua espressione in funzione di E, V, N , che sono le variabili naturali di S , possiamo poi ricavare in modo abbastanza diretto la termodinamica del sistema.

Il legame cercato è quello stabilito da Boltzmann

$$S(E, V, N) = k_B \ln W(E, V, N), \quad (2.23)$$

dove $W(E, V, N)$ indica la molteplicità di stati microscopici corrispondenti allo stato di equilibrio dato. Poiché in meccanica classica gli stati dinamici costituiscono un insieme continuo, la loro molteplicità deve necessariamente esprimersi attraverso i volumi nello spazio delle fasi; quindi la definizione di entropia in meccanica statistica si trova anche scritta

$$S(E, V, N) = k_B \ln \left[\frac{\Gamma_{\Delta}(E, V, N)}{z^{3N}} \right], \quad (2.24)$$

dove z ha le dimensioni fisiche di un'azione e z^{3N} dà il volume di una cella elementare di riferimento nello spazio delle fasi; questo permette di "contare" le configurazioni microscopiche. Si possono dare due argomenti per motivare tale corrispondenza.

1 – In termodinamica l'entropia di un gas ideale monoatomico si scrive:

$$S(E, V, N) = N k_B \ln \left[\left(\frac{V}{N} \right) \left(\frac{E}{N} \right)^{3/2} + s_0 \right], \quad (2.25)$$

come è richiesto dall'equazione di stato dei gas perfetti e dall'additività di S .

La costante s_0 nella (2.25) può essere determinata invocando la terza legge (*teorema di Nernst*), ponendo quindi pari a zero l'entropia del sistema a $T = 0$. Nel 1912 O. Sackur e H. Tetrode, utilizzando i dati sperimentali sul mercurio nelle fasi solida, liquida e gassosa, arrivarono indipendentemente al risultato

$$S(E, V, N) = N k_B \left[\ln \left(\frac{V}{N} \right) \left(\frac{E}{N} \right)^{3/2} + \frac{5}{2} + \frac{3}{2} \ln \frac{4\pi m}{3h^2} \right]; \quad (2.26)$$

la (2.26), nota come equazione di Sackur-Tetrode, verifica i dati sperimentali se la costante h che vi compare ha il valore della costante di Planck. Essa si può riscrivere

come:

$$S(E, V, N) = N k_B \ln \left[(V) \left(\frac{4}{3} \pi e m \frac{E}{N} \right)^{3/2} \frac{1}{h^3} \right] - N k_B \ln \frac{N}{e}. \quad (2.27)$$

Usando l'approssimazione di Stirling, $\ln N! = N \ln(N/e) + O(\ln N) \cong N \ln(N/e)$, si ha quindi

$$S(E, V, N) = k_B \ln \frac{\left[V \left(\frac{4}{3} \pi e m \frac{E}{N} \right)^{3/2} \right]^N}{h^{3N} N!}, \quad (2.28)$$

ovvero, dalle (2.13) e (2.14),

$$S(E, V, N) = k_B \ln \frac{\Gamma_{\Delta}(E, V, N)}{h^{3N} N!} = k_B \ln \left(e \sqrt{\frac{e^3 \pi}{6}} \frac{(V/N) V_p}{h^3} \right)^N, \quad (2.29)$$

dove, ricordiamo, $V_p = (2mE/N)^{3/2} (4\pi/3)$ è il volume di spazio degli impulsi accessibile a una singola particella con energia fino a E/N . Per un gas ideale, l'entropia risulta quindi proporzionale al logaritmo della molteplicità dello stato di equilibrio, se questa è calcolata dividendo il volume occupato nello spazio delle fasi, Γ_{Δ} , per un volume di riferimento pari a h^{3N} (cioè nella (2.24) z è la costante di Planck), e ancora per $N!$. Sia la presenza di h^{3N} che quella di $N!$ si possono giustificare con considerazioni di meccanica quantistica. Essendo il prodotto delle indeterminazioni sulla misura simultanea di ogni coppia di variabili coniugate (posizione e impulso) limitato inferiormente da una quantità $O(h)$, h^{3N} rappresenta la massima precisione con cui può essere determinato uno stato dinamico di N particelle: lo spazio delle fasi all'interno del volumetto h^{3N} rappresenta un unico stato. Ma il principio di indeterminazione rende anche fisicamente indistinguibili particelle identiche, per cui all'interno del volume Γ_{Δ} gli stati ottenuti permutando le N particelle non sono stati diversi: la divisione per $N!$ determina il conteggio corretto degli stati fisici. La seconda uguaglianza della (2.29) mostra che la molteplicità dello stato di equilibrio

$$W(E, V, N) = \left(e \sqrt{\frac{e^3 \pi}{6}} \frac{(V/N) V_p}{h^3} \right)^N, \quad (2.30)$$

analogamente a quanto vale per Γ_{Δ} , è il prodotto su tutte le particelle del numero di stati disponibili per la singola particella, dove uno stato di singola particella occupa il volume h^3 e lo spazio reale disponibile non è V ma V/N , questo a causa della indistinguibilità.

2 – La prima equazione della (2.29) mostra che, per un sistema con numero fissato di particelle, le variazioni di entropia si possono ottenere anche dalla definizione (2.24) nella forma $S = k_B \ln \Gamma_{\Delta} + \text{costante}$.

L'idea che l'entropia sia collegata ai volumi nello spazio delle fasi ha la sua origine nel *Teorema di Helmholtz* e risale ai tentativi iniziali di far discendere “esattamente” la termodinamica dalla meccanica. Abbiamo visto che questo non è possibile in linea di principio, e che invece la termodinamica consegue “praticamente” dalla meccanica, in quanto le sue leggi descrivono comportamenti veri nella quasi totalità dei casi. Tuttavia il collegamento suggerito dal teorema si è rivelato fondamentale nello sviluppo della teoria, grazie alla generalizzazione fattane da Boltzmann.

L'argomento, poco noto anche ai cultori di storia della fisica, è stato riconsiderato recentemente da Gallavotti e discusso in termini molto chiari (accessibile a studenti) da Campisi e Kobe.

Ricordiamo che per un sistema con numero fissato di particelle, il cui stato termodinamico è individuato da 2 variabili, E e V (per esempio, un gas contenuto in un volume V), per il secondo principio della termodinamica si ha

$$dS = \frac{dE + PdV}{T} = \text{differenziale esatto.} \quad (2.31)$$

Consideriamo un sistema meccanico unidimensionale con Hamiltoniana

$$H(q, p, \tilde{V}) = \frac{p^2}{2m} + \phi(q, \tilde{V}) \quad (2.32)$$

dove \tilde{V} è un parametro di controllo che può essere fatto variare, si pensi, per esempio, a un pendolo di lunghezza \tilde{V} . Assumiamo che il potenziale $\phi(q, \tilde{V})$ diverga per $|q| \rightarrow \infty$, di modo che per ogni valore di E e \tilde{V} il moto risulta periodico; indichiamo con $\tau(E, \tilde{V})$ il periodo, con $q_-(E, \tilde{V})$ e $q_+(E, \tilde{V})$ il valore minimo e massimo della q .

Possiamo calcolare le medie su un ciclo delle seguenti grandezze meccaniche

$$\frac{1}{\tau(E, \tilde{V})} \int_0^\tau \frac{p^2(t)}{2m} dt \equiv \left\langle \frac{p^2}{2m} \right\rangle_t, \quad (2.33)$$

$$\frac{1}{\tau(E, \tilde{V})} \int_0^\tau \frac{\partial \phi(q, \tilde{V})}{\partial \tilde{V}} dt \equiv \left\langle \frac{\partial \phi(q, \tilde{V})}{\partial \tilde{V}} \right\rangle_t, \quad (2.34)$$

e definire le 2 quantità $\tilde{T}(E, \tilde{V})$ e $\tilde{P}(E, \tilde{V})$, tramite le relazioni

$$\frac{1}{2} k_B \tilde{T} = \left\langle \frac{p^2}{2m} \right\rangle_t, \quad \tilde{P} = - \left\langle \frac{\partial \phi(q, \tilde{V})}{\partial \tilde{V}} \right\rangle_t, \quad (2.35)$$

dalle quali risulta evidente che, in questo contesto puramente meccanico, \tilde{T} , \tilde{P} e \tilde{V} vogliono rappresentare temperatura, pressione e volume termodinamici. Si ha il

Teorema di Helmholtz: per la funzione

$$\tilde{S}(E, \tilde{V}) = k_B \ln 2 \int_{q_-(E, \tilde{V})}^{q_+(E, \tilde{V})} \sqrt{2m[E - \phi(q, \tilde{V})]} dq = k_B \ln \oint p(q) dq \quad (2.36)$$

valgono le seguenti uguaglianze

$$\frac{\partial \tilde{S}}{\partial E} = \frac{1}{\tilde{T}}, \quad \frac{\partial \tilde{S}}{\partial \tilde{V}} = \frac{\tilde{P}}{\tilde{T}}. \quad (2.37)$$

La dimostrazione non è difficile; il lettore interessato può trovare i dettagli nell'articolo di Campisi e Kobe.

Il teorema di Helmholtz implica, risultato decisamente non banale, che esiste un analogo meccanico del secondo principio della termodinamica. L'espressione

$$\frac{dE + \tilde{P}d\tilde{V}}{\tilde{T}} \quad (2.38)$$

dove \tilde{T} e \tilde{P} sono espresse in termini di medie temporali di quantità meccaniche, è un differenziale esatto, per cui, nell'analogia meccanica, la funzione $\tilde{S}(E, \tilde{V})$ ha il ruolo di entropia.

Come è chiaro dalla (2.36), $\tilde{S}(E, \tilde{V})$ è proporzionale al logaritmo dell'area dello spazio delle fasi (bidimensionale, in questo caso) racchiusa dall'orbita di energia E e parametro \tilde{V} ; essa quindi può essere scritta nella forma usuale

$$\tilde{S}(E, \tilde{V}) = k_B \ln \int_{H(q, p, \tilde{V}) < E} dpdq. \quad (2.39)$$

La peculiarità del caso unidimensionale è la sua ergodicità: dato \tilde{V} , ad ogni energia E è associato un unico ciclo, si possono definire univocamente $\tilde{T}, \tilde{P}, \tilde{S}$ come funzioni di E e \tilde{V} e le medie nel tempo sono equivalenti a medie nello spazio delle fasi (sul ciclo). La dimostrazione è abbastanza semplice: data un'osservabile $f(p, q)$ il suo valor medio (microcanonico) sulla curva $H(p, q) = E$ si può scrivere

$$\langle f \rangle_E = \frac{\int dpdq f(p, q) \delta(p^2/2m + \phi(q) - E)}{\int dpdq \delta(p^2/2m + \phi(q) - E)}. \quad (2.40)$$

Considerando la $\delta(p^2/2m + \phi(q) - E)$ come funzione di p e ricordando la proprietà

$$\delta(g(x)) = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|}, \quad (2.41)$$

dove la somma è fatta sulle radici dell'equazione $g(x) = 0$, si ha

$$\delta(p^2/2m + \phi(q) - E) = \frac{\delta(p - p_+)}{p_+/m} + \frac{\delta(p - p_-)}{|p_-/m|}, \quad (2.42)$$

con $p_{\pm} = \pm\sqrt{2m(E - \phi(q))} = p_{\pm}(q)$. Si ha quindi

$$\begin{aligned}
 \langle f \rangle_E &= \frac{\int_{q_-}^{q_+} dq \left[(m/p_+) f(q, p_+) + (m/p_-) f(q, p_-) \right]}{\int_{q_-}^{q_+} dq \left[(m/p_+) + (m/p_-) \right]} = \\
 &= \frac{\int_{q_-}^{q_+} dq (m/p_+) f(q, p_+) + \int_{q_+}^{q_-} dq (m/p_-) f(q, p_-)}{\int_{q_-}^{q_+} dq (m/p_+) + \int_{q_+}^{q_-} dq (m/p_-)} = \quad (2.43) \\
 &= \frac{\oint dq (m/p) f(p, q)}{\oint dq (m/p)} ;
 \end{aligned}$$

poiché $dq = (p/m)dt$ si ha infine

$$\langle f \rangle_E = \frac{\int_0^{\tau} dt f(p(t), q(t))}{\int_0^{\tau} dt} = \langle f \rangle_t . \quad (2.44)$$

In generale in un sistema con N particelle l'ergodicità non è più garantita: possono esistere cicli diversi con uguale energia, ma diverse medie temporali per le altre grandezze meccaniche. Per generalizzare il risultato a sistemi con molte particelle, cioè trovare una funzione $\tilde{S}(E, \tilde{V})$ tale che le relazioni (2.35) e (2.37) siano valide, Boltzmann fa l'ipotesi ergodica. Assume che la traiettoria visiti tutta la superficie $H = E$, e quindi che la media temporale sia equivalente alla media microcanonica. Il moto periodico del caso unidimensionale è rimpiazzato da moto ergodico sulla superficie a energia costante. Fatta l'ipotesi ergodica, è possibile mostrare un **Teorema di Helmholtz generalizzato**⁴ definendo

$$\tilde{S}(E, \tilde{V}) = k \ln \int_{H(\mathbf{q}, \mathbf{p}, \tilde{V}) < E} d\mathbf{q} d\mathbf{p} . \quad (2.45)$$

L'osservazione finale per recuperare la definizione (2.24) è che in uno spazio di dimensione molto grande, come notato nel primo capitolo, il volume di un solido convesso (come quello definito dall'integrale qua sopra) è praticamente tutto concentrato in uno strato sottile al di sotto della superficie esterna: per sistemi macroscopici (2.24) e (2.45) sono equivalenti.

⁴ La dimostrazione non si discosta molto da quella del caso unidimensionale, vedi Campisi e Kobe.

2.4 Letture consigliate

- G. Gallavotti *Statistical mechanics. A short treatise* (Springer-Verlag, Berlin 1995);
- N. Zanghì *I fondamenti concettuali dell'approccio statistico in fisica*, in *La Natura delle Cose* Ed. V. Allori, M. Dorato, F. Laudisa e N. Zanghì, pag. 139 (Carocci, Roma 2005);
- A.I. Khinchin *Mathematical Foundations of Statistical Mechanics* (Dover Publications, 1960);
- M. Campisi and D. Kobe *Derivation of the Boltzmann principle* *Am. J. Phys.* **78**, 608 (2010);
- P. Mazur and J. van der Linden *Asymptotic form of the structure function for real systems* *J. Math. Phys.* **4**, 271 (1963).

Il Microcanonico e oltre

Sommario Nei capitoli precedenti abbiamo fatto delle considerazioni che, attraverso l'ipotesi ergodica, portano a concludere che un accettabile punto di partenza, per sviluppare una teoria microscopica della Termodinamica, sia la descrizione di un sistema macroscopico in equilibrio mediante la densità di probabilità microcanonica. In linea di principio, questa densità è adatta a descrivere un sistema isolato e il legame tra Meccanica Statistica e Termodinamica è la relazione di Boltzmann, che, a meno di una costante, identifica l'entropia termodinamica con il logaritmo del numero di stati microscopici disponibili per il sistema

$$S(E, V, N) = k_B \ln W(E, V, N), \quad (3.1)$$

o del volume di spazio delle fasi che essi occupano

$$S(E, V, N) = k_B \ln \left(\Gamma_{\Delta}(E, V, N) / h^{3N} \right), \quad (3.2)$$

tenendo presente che nel caso di gas di particelle identiche Γ_{Δ} va diviso anche per $N!$. Partendo da queste basi, in questo capitolo deriviamo qualche risultato interessante, come la distribuzione di Maxwell-Boltzmann, che dà la probabilità di occupazione degli stati di singola particella, e la “densità canonica”, che fornisce la distribuzione di probabilità degli stati microscopici di una porzione del sistema isolato.

Svilupperemo inoltre una impostazione alternativa, basata sull'idea di “ensemble”, per introdurre le densità adatte a descrivere sistemi in equilibrio e discuteremo brevemente l'importante strumento delle simulazioni numeriche per calcolare le medie all'equilibrio.

3.1 Osservazioni preliminari sui volumi e sull'entropia

Anche in questo capitolo, se ciò non crea ambiguità, useremo la notazione compatta \mathbf{X} per indicare l'insieme delle coordinate canoniche di un sistema Hamiltoniano; porremo inoltre, per comodità, $h = 1$. Notiamo che si può scrivere

$$\begin{aligned}\Gamma_{\Delta}(E, V, N) &= \int_{E < H < E + \Delta} d\mathbf{X} = \int_{H < E + \Delta} d\mathbf{X} - \int_{H < E} d\mathbf{X} \\ &\equiv \Sigma(E + \Delta, V, N) - \Sigma(E, V, N),\end{aligned}\quad (3.3)$$

dove abbiamo introdotto la nuova funzione

$$\Sigma(E, V, N) = \int_{H < E} d\mathbf{X}, \quad (3.4)$$

che è il volume occupato dagli stati del sistema con energie fino a E . Possiamo usare il teorema di Lagrange per scrivere la variazione di Σ che compare nella (3.3):

$$\Sigma(E + \Delta, V, N) - \Sigma(E, V, N) = \left. \frac{\partial \Sigma(E, V, N)}{\partial E} \right|_{\hat{E}} \Delta; \quad (3.5)$$

dove \hat{E} è una opportuna energia nell'intervallo di incertezza Δ . Definendo quindi $\omega(E, V, N) = \partial \Sigma(E, V, N) / \partial E$, otteniamo

$$\Gamma_{\Delta}(E, V, N) = \int_{E < H < E + \Delta} d\mathbf{X} = \omega(\hat{E}, N, V) \Delta. \quad (3.6)$$

Trascurando la differenza tra E e \hat{E} (che sono due valori che differiscono meno dell'incertezza sull'energia), si ha quindi

$$\ln \Gamma_{\Delta}(E, V, N) = \ln \omega(E, V, N) + \ln \Delta. \quad (3.7)$$

Nel capitolo precedente abbiamo calcolato $\Gamma_{\Delta}(E, V, N)$ per un gas ideale e abbiamo trovato (vedi (2.13) e (2.14)) che si può scrivere nella forma di un volume di spazio delle fasi di singola particella (*vol*) alla potenza N :

$$\Gamma_{\Delta}(E, V, N) = \exp \left[N \ln(\text{vol}) + O(\ln N) \right] = (\text{vol})^N, \quad (3.8)$$

dove abbiamo trascurato correzioni di ordine $\ln N$ rispetto a N , cosa certamente lecita nei casi in cui N è dell'ordine del numero di Avogadro. Una simile dipendenza deve essere vera in generale, cioè non solo per un gas ideale: un volume in uno spazio di dimensione d si deve esprimere come (qualcosa) ^{d} . Nel caso di N particelle, con $d = 6N$, i volumi si potranno mettere nella forma $\left((\text{qualcosa})^6 \right)^N = (\text{vol})^N$, dove *vol* rappresenta il volume caratteristico dello spazio delle fasi di singola particella nello stato macroscopico individuato da (E, V, N) . Infatti, per un gas ideale abbiamo

visto che $vol \propto (V) (E/N)^{3/2}$. Riconsiderando l'equazione (3.7), e tenendo conto del fatto che l'incertezza sull'energia è proporzionale a E , possiamo concludere che, trascurando correzioni di ordine $\ln N$, si ha

$$S(E, V, N) = k_B \ln \Gamma_{\Delta}(E, V, N) = k_B \ln \omega(E, V, N). \quad (3.9)$$

Vediamo quindi che l'entropia di un sistema non dipende dall'incertezza sulla misura dell'energia e può essere determinata anche per mezzo della funzione $\omega(E, V, N) = \partial \Sigma / \partial E$.

In realtà si può vedere che, nella stessa approssimazione di trascurare termini di ordine $\ln N$ rispetto a N , è anche permesso scrivere

$$S(E, V, N) = k_B \ln \Sigma(E, V, N). \quad (3.10)$$

Sia quest'ultima equazione che la (3.9) sono conseguenze del peculiare comportamento dei volumi negli spazi di alta dimensionalità.

Prima di concludere questa parentesi è utile far presente che, partendo dalla definizione (3.4), $\Sigma(E, V, N)$ si può scrivere anche

$$\Sigma(E, V, N) = \int \theta(E - H(\mathbf{X})) d\mathbf{X}, \quad (3.11)$$

dove $\theta(y)$ è la funzione “gradino” (o di Heaviside)

$$\theta(y) = 0 \quad \text{se } y < 0; \quad \theta(y) = 1 \quad \text{se } y > 0. \quad (3.12)$$

Dalla (3.11) segue che, essendo ω la derivata di Σ rispetto al parametro E , si ha anche

$$\omega(E, V, N) = \int \delta(E - H(\mathbf{X})) d\mathbf{X}, \quad (3.13)$$

dove $\delta(y)$ indica la delta di Dirac, avendo usato la relazione $\theta'(y) = \delta(y)$. Questa rappresentazione della funzione ω è utile negli sviluppi teorici. La presenza della $\delta(E - H(\mathbf{X}))$, che seleziona tutti e soli i punti con energia pari a E , fa vedere che ω è una misura dei punti sulla ipersuperficie di energia assegnata. In più, essa è una misura invariante nell'evoluzione hamiltoniana che regge il sistema; infatti dato che per definizione

$$\omega(E, V, N) = \lim_{dE \rightarrow 0} \frac{\int_{H(\mathbf{X}) < E + dE} d\mathbf{X} - \int_{H(\mathbf{X}) < E} d\mathbf{X}}{dE}, \quad (3.14)$$

e dato che le quantità a numeratore e denominatore sono invarianti del moto, anche $\omega(E, V, N)$ lo è. La forma esplicita della misura invariante sulla superficie isoenergetica si trova nell'Appendice C. Si noti che, come è evidente dall'equazione qua sopra, $\omega(E, V, N) dE$ dà il volume occupato dagli stati con energia compresa tra E ed $E + dE$, per cui $\omega(E, V, N)$ è detta *densità degli stati*.

3.2 Dall'insieme microcanonico a quello canonico

Iniziamo ricordando che, se si ha una densità di probabilità per n variabili $p(x_1, x_2, \dots, x_n)$ ma si è interessati solamente ad alcune di esse, si può ricavare l'informazione richiesta mediante la cosiddetta "marginalizzazione", cioè l'integrazione sulle variabili che non interessano. Per esempio, nel caso di 3 variabili (x_1, x_2, x_3) si può essere interessati a una sola (ad esempio x_1) oppure a due (ad esempio x_1 e x_2) per cui si introducono le densità di probabilità marginali

$$\begin{aligned} p(x_1) &= \int p(x_1, x_2, x_3) dx_2 dx_3, \\ p(x_1, x_2) &= \int p(x_1, x_2, x_3) dx_3. \end{aligned} \quad (3.15)$$

Consideriamo un sistema isolato, costituito da N particelle in un volume V , con energia compresa nell'intervallo $[E, E + \Delta]$ e descritto da $6N$ variabili, complessivamente indicate con il vettore \mathbf{X} . Supponiamo che il sistema sia diviso in 2: il sistema A_1 , con N_1 particelle nel volume V_1 , descritto dalle $6N_1$ variabili \mathbf{X}_1 , e il sistema A_2 , con $N_2 = N - N_1$ particelle nel volume $V_2 = V - V_1$, descritto dalle $6N_2$ variabili \mathbf{X}_2 . Si ha quindi $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$. Per fissare le idee, possiamo pensare a due gas in contatto termico tra loro, mediante una parete divisoria, ma globalmente isolati.

Per il sistema completo, in equilibrio, usiamo la densità microcanonica

$$\rho_{mc}(\mathbf{X}) = \frac{1}{\Gamma_{\Delta}(E, V, N)} = \rho_{mc}(\mathbf{X}_1, \mathbf{X}_2), \quad (3.16)$$

dove

$$\begin{aligned} \rho_{mc}(\mathbf{X}_1, \mathbf{X}_2) &= \frac{1}{\Gamma_{\Delta}(E, N, V)} \quad \text{se } H((\mathbf{X}_1, \mathbf{X}_2)) \in [E, E + \Delta] \\ \rho_{mc}(\mathbf{X}_1, \mathbf{X}_2) &= 0 \quad \text{se } H((\mathbf{X}_1, \mathbf{X}_2)) \notin [E, E + \Delta]. \end{aligned} \quad (3.17)$$

L'Hamiltoniana totale si può scrivere nella forma $H((\mathbf{X}_1, \mathbf{X}_2)) = H_1(\mathbf{X}_1) + H_2(\mathbf{X}_2) + H_{12}(\mathbf{X}_1, \mathbf{X}_2)$ e si può normalmente trascurare il contributo di interazione¹ H_{12} .

La densità di probabilità marginale per il solo sistema A_1 , come segue dalle (3.15), si ottiene integrando sulle variabili del sistema A_2 :

$$\rho_{A_1}(\mathbf{X}_1) = \int \rho_{mc}(\mathbf{X}_1, \mathbf{X}_2) d\mathbf{X}_2; \quad (3.18)$$

l'integrazione, che va estesa agli stati del sistema A_2 con energia pari all'energia totale meno quella che compete al sistema A_1 mentre si trova nello stato \mathbf{X}_1 , dà

¹ Abbiamo visto che questo è fisicamente sensato se $N_1 \gg 1$, $N_2 \gg 1$ e il raggio di interazione tra le coppie di particelle è piccolo rispetto alla grandezza lineare dei sistemi descritti dalle variabili \mathbf{X}_1 e \mathbf{X}_2 .

come risultato

$$\rho_{A_1}(\mathbf{X}_1) = \int \frac{d\mathbf{X}_2}{\Gamma_{\Delta}(E, V, N)} = \frac{\Gamma_{\Delta}^{(2)}(E - H_1(\mathbf{X}_1), V_2, N_2)}{\Gamma_{\Delta}(E, V, N)}. \quad (3.19)$$

Usando la definizione di entropia (3.2) si può scrivere (ricordarsi che abbiamo posto $h = 1$)

$$\Gamma_{\Delta}^{(2)}(E - H_1(\mathbf{X}_1), V_2, N_2) = \exp \left[S_2(E - H_1(\mathbf{X}_1), V_2, N_2) / k_B \right]. \quad (3.20)$$

Supponiamo ora che il sistema A_1 sia molto più piccolo di A_2 : $N_1 \ll N_2$ e $V_1 \ll V_2$; dato che siamo in equilibrio, anche per le energie di equilibrio, indichiamole con U_1 e U_2 , è $U_1 \ll U_2$. Conseguentemente per i valori di energia che ha senso considerare, microscopicamente fluttuanti intorno a U_1 , si ha $H_1(\mathbf{X}_1) \ll E$, ed è permesso approssimare l'entropia $S_2(E - H_1)$ nella (3.20) con uno sviluppo di Taylor al primo ordine. Ricordando che $\partial S / \partial E = T^{-1}$ è l'inverso della temperatura, scriviamo

$$\begin{aligned} S_2(E - H_1) &= S_2(U_2 + U_1 - H_1) \\ &\cong S_2(U_2) + \frac{1}{T} (U_1 - H_1) \end{aligned} \quad (3.21)$$

per cui la (3.19) diventa

$$\rho_{A_1}(\mathbf{X}_1) = \frac{\exp \left[S_2(U_2) / k_B + U_1 / k_B T - H_1(\mathbf{X}_1) / k_B T \right]}{\exp \left[S_2(U_2) / k_B + S_1(U_1) / k_B \right]} \quad (3.22)$$

o anche

$$\begin{aligned} \rho_{A_1}(\mathbf{X}_1) &= \exp \left[\frac{1}{k_B T} (U_1 - T S_1(U_1)) - \frac{1}{k_B T} H_1(\mathbf{X}_1) \right] \\ &= \exp \left[\frac{1}{k_B T} (F_1(T) - H_1(\mathbf{X}_1)) \right], \end{aligned} \quad (3.23)$$

dove è stata introdotta l'energia libera di Helmholtz $F = U - TS$. Possiamo riscrivere la densità nella forma standard con cui si definisce la *distribuzione canonica* (che descrive l'*insieme canonico*) per il sottosistema A_1

$$\rho_C(\mathbf{X}_1) = \frac{e^{-\beta H_1(\mathbf{X}_1)}}{Z(\beta, V_1, N_1)}, \quad (3.24)$$

dove $Z(\beta, V_1, N_1) = \int e^{-\beta H_1(\mathbf{X}_1)} d\mathbf{X}_1$ è la *funzione di partizione canonica* e si è posto $\beta = 1/k_B T$. Il confronto tra (3.24) e (3.23) fornisce il legame tra la termodinamica e la meccanica statistica basata sulla densità canonica:

$$F_1(\beta, V_1, N_1) = -\frac{1}{\beta} \ln Z(\beta, V_1, N_1), \quad (3.25)$$

che, nel contesto dell'*insieme canonico*, costituisce l' analogo della relazione di Boltzmann (3.2) nel microcanonico.

È importante sottolineare che l'equazione (3.24) (come la (3.19)) fornisce la densità di probabilità di trovare il sistema A_1 in un particolare stato \mathbf{X}_1 e, anche se questa densità dipende solo dall'energia dello stato $H_1(\mathbf{X}_1)$, essa non è la densità di probabilità $P(E)$ di trovare A_1 con energia $E = H_1(\mathbf{X}_1)$, che si ottiene invece, come è facile mostrare, moltiplicando la $\rho_C(\mathbf{X}_1)$ per la molteplicità degli stati con energia $E = H_1(\mathbf{X}_1)$, ovvero $\omega_1(E, V_1, N_1)$. Abbiamo quindi:

$$P(E) = \frac{\omega_1(E, V_1, N_1)e^{-\beta E}}{Z(\beta, V_1, N_1)}. \quad (3.26)$$

3.3 La distribuzione di Maxwell-Boltzmann

Nella sezione precedente abbiamo ricavato la distribuzione canonica, che concerne gli stati di un sistema macroscopico costituito da una parte piccola di un sistema più grande. Il sistema "piccolo" deve essere comunque macroscopico per poter trascurare il termine di interazione $H_{12}(\mathbf{X}_1, \mathbf{X}_2)$ rispetto a $H_1(\mathbf{X}_1)$. La distribuzione di Maxwell-Boltzmann (MB) si riferisce agli stati di singola particella. Nell'ipotesi di gas perfetto essa è molto semplice da ottenere descrivendo il sistema in equilibrio con la distribuzione canonica, ma è istruttivo ricavarla, utilizzando ancora la (3.15), restando nel microcanonico.

Consideriamo un sistema isolato, costituito da N particelle non interagenti, contenute in un recipiente di volume V , con Hamiltoniana

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m}; \quad (3.27)$$

il sistema è in equilibrio con energia E , e lo descriviamo con l'insieme microcanonico.

Inizialmente notiamo che la densità di probabilità di un sottosistema, data dalla (3.19), per la (3.6) si può riscrivere

$$\rho_{A_1}(\mathbf{X}_1) = \frac{\omega_2(E - H_1(\mathbf{X}_1), N_2, V_2)}{\omega(E, V, N)}. \quad (3.28)$$

Nella formula qua sopra $\omega(E, V, N)$ dà la misura, nello spazio delle fasi delle N particelle, degli stati con energia (esattamente) E , mentre $\omega_2(E_2, V_2, N_2)$, è la misura degli stati di N_2 particelle con energia pari (esattamente) a $E_2 = E - H_1(\mathbf{X}_1)$. Nel caso di sistemi reali la (3.28), come la (3.19), è una buona approssimazione, ottenuta trascurando l'interazione tra i sottosistemi, che sono comunque macroscopici. Per un gas ideale la (3.28) è esatta e vale anche se A_1 è costituito da una singola particella. In questo caso poniamo $\mathbf{X}_1 = (\mathbf{x}, \mathbf{p})$, $N_2 = N - 1$, $V_2 = V$ e quindi, per

la densità di probabilità che una particella si trovi nella posizione \mathbf{x} con impulso \mathbf{p} si ha:

$$\rho(\mathbf{x}, \mathbf{p}) = \frac{\omega\left(E - \frac{\mathbf{p}^2}{2m}, V, N - 1\right)}{\omega(E, V, N)}. \quad (3.29)$$

La densità degli stati si calcola facilmente come derivata della funzione $\Sigma(E, V, N)$, per la quale si ha

$$\Sigma(E, V, N) = \int_{\sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} < E} d^N \mathbf{x} d^N \mathbf{p} = V^N \int_{\sum_{i=1}^N \mathbf{p}_i^2 < 2mE} d^N \mathbf{p}; \quad (3.30)$$

usando la formula per il volume di una ipersfera di raggio R in d dimensioni: $Vol_d(R) = \pi^{d/2} R^d / \Gamma(\frac{d}{2} + 1)$, con $d = 3N$ e $R = \sqrt{2mE}$, si arriva a

$$\omega(E, V, N) = V^N 3N m C_N (2mE)^{(3N/2)-1}, \quad (3.31)$$

dove $C_N = \pi^{3N/2} / \Gamma((3N/2) + 1)$. L'equazione (3.29) diventa quindi

$$\rho(\mathbf{x}, \mathbf{p}) = \frac{1}{V} \frac{\Gamma(3/2(N))}{\Gamma(3/2(N-1))} \frac{1}{(2\pi m E)^{3/2}} \left(1 - \frac{\mathbf{p}^2}{2mE}\right)^{3N/2-5/2}, \quad (3.32)$$

che è valida per $|\mathbf{p}| \leq \sqrt{2mE}$. L'espressione trovata per la distribuzione di singola particella è esatta, e vale per ogni N , ma assume una forma particolarmente semplice se $N \gg 1$. In questo caso, il fattore $5/2$ nell'esponente si può trascurare rispetto a $(3/2)N$, l'approssimazione di Stirling $\Gamma(x) \simeq \sqrt{2\pi} x^{x-(1/2)} e^{-x}$ (usata con la dovuta attenzione) dà

$$\frac{\Gamma(3/2(N))}{\Gamma(3/2(N-1))} \simeq \left(\frac{3}{2}N\right)^{3/2}, \quad (3.33)$$

per cui si può scrivere

$$\rho(\mathbf{x}, \mathbf{p}) \simeq \frac{1}{V} \left(\frac{3}{2} \frac{N}{E}\right)^{3/2} \frac{1}{(2\pi m)^{3/2}} \left(1 - \frac{\mathbf{p}^2}{2mE}\right)^{3N/2}, \quad (3.34)$$

o anche

$$\rho(\mathbf{x}, \mathbf{p}) \simeq \frac{1}{V} \frac{1}{[2\pi m(2E/3N)]^{3/2}} \left(1 - \frac{\mathbf{p}^2/2m}{(2E/3N)(3N/2)}\right)^{3N/2}. \quad (3.35)$$

A questo punto si può usare la relazione che dà l'esponenziale come limite: $\lim_{z \rightarrow \infty} (1 + b/z)^z = e^b$, con la condizione che, mentre $N \rightarrow \infty$ il rapporto E/N resti costante, ottenendo

$$\rho(\mathbf{x}, \mathbf{p}) \simeq \frac{1}{V} \frac{1}{[2\pi m(2E/3N)]^{3/2}} \exp - \frac{\mathbf{p}^2}{2m(2E/3N)}. \quad (3.36)$$

Il risultato trovato dà, per la singola particella, una distribuzione uniforme nello spazio e una distribuzione gaussiana nelle componenti dell'impulso, la cui varianza dipende dal rapporto E/N . Possiamo però usare la relazione

$$\frac{1}{T} = \frac{\partial S}{\partial E} = k_B \frac{\partial \ln \omega(E, V, N)}{\partial E} \quad (3.37)$$

per il gas delle N particelle, con ω dato dalla (3.31), ottenendo il legame tra energia per particella e temperatura

$$\frac{1}{T} = k_B \frac{3N}{2E}. \quad (3.38)$$

Il risultato finale per la distribuzione di Maxwell-Boltzmann è quindi:

$$\rho(\mathbf{x}, \mathbf{p}) \simeq \frac{1}{V} \frac{1}{\sqrt{(2\pi mk_B T)^3}} \exp - \frac{\mathbf{p}^2}{2mk_B T}. \quad (3.39)$$

La distribuzione di MB, per come è stata qui ricavata, rappresenta la densità di probabilità che una particella del gas si trovi in un certo stato. Però, come abbiamo brevemente accennato alla fine del Capitolo 1, la distribuzione di MB è anche la densità di probabilità "empirica" delle particelle che, all'equilibrio, è realizzata sulla quasi totalità degli stati del gas (e questo, come abbiamo discusso brevemente nel Capitolo 1 e come suggerisce il teorema di Lanford, anche in presenza di un'interazione). Questo doppio significato della distribuzione di MB (distribuzione di singola particella e distribuzione empirica di massima molteplicità) è ancora una manifestazione della legge dei grandi numeri, per la quale la frequenza di un evento tende alla sua probabilità.

La caratterizzazione della distribuzione di MB come distribuzione empirica di massima molteplicità è alla base del metodo, a volte usato, di ricavare le distribuzioni di Bose-Einstein e Fermi-Dirac impostando un problema di massimizzazione per le distribuzioni empiriche, tenendo conto dei vincoli quantistici sugli stati di bosoni e fermioni identici. In questo modo è possibile ricavare le formule che danno i numeri medi di occupazione di uno stato di singola particella, s , sia per i bosoni

$$\langle n_s \rangle_B = \frac{1}{e^{\beta(\varepsilon_s - \mu)} - 1}, \quad (3.40)$$

che per i fermioni

$$\langle n_s \rangle_F = \frac{1}{e^{\beta(\varepsilon_s - \mu)} + 1}; \quad (3.41)$$

dove ε_s indica l'energia dello stato considerato e μ è un parametro che si introduce per fissare il numero di particelle nel sistema: $\sum_s \langle n_s \rangle = N$, il cui significato fisico è quello di potenziale chimico.

3.4 Gli ensembles

Abbiamo affrontato il problema della Meccanica Statistica cercando di capire se quello che avviene durante l'evoluzione dello stato microscopico di un dato sistema macroscopico sia compatibile con un comportamento di tipo termodinamico, cosa che, a prima vista, potrebbe sembrare impossibile. La conclusione a cui siamo giunti è che, prestando attenzione alle sole osservabili che definiscono lo stato termodinamico del sistema (osservabili *macroscopiche*) e non al dettaglio microscopico, il comportamento del sistema è *praticamente* quello termodinamico. In particolare, se il sistema è inizialmente preparato in uno stato di non equilibrio termodinamico, è perfettamente concepibile che la sua evoluzione microscopica lo porti in uno stato che, macroscopicamente, possiamo descrivere come di equilibrio. Naturalmente il sistema considerato subirebbe la stessa sorte partendo da uno stato microscopico diverso, ma che desse luogo alla stessa configurazione macroscopica iniziale.

Discutendo di stati microscopici iniziali diversi di un dato sistema, si può pensare alla ripetizione (in un certo laboratorio, in tempi successivi) di un esperimento in cui il controllo delle variabili macroscopiche permette di fissare i loro valori iniziali. È anche possibile pensare a un insieme di repliche del sistema (in diversi laboratori, nello stesso tempo), inizialmente tutte nello stato macroscopico assegnato, ma ognuna in uno stato microscopico diverso. Questa osservazione dà la possibilità di introdurre un punto di vista alternativo, molto diffuso, per affrontare il problema della Meccanica Statistica: il punto di vista degli *ensembles*.

Consideriamo un sistema isolato che inizialmente, per come è stato preparato, non è in equilibrio; per esempio, nel caso del sistema modello considerato nel primo capitolo, quando i due gas sono messi a contatto le loro energie non corrispondono a quelle dell'equilibrio finale. L'insieme degli stati microscopici che realizzano la condizione iniziale può essere caratterizzata da una densità di probabilità diversa da zero solo su questi stati, indichiamola con $\rho_0(\mathbf{X})$. La densità ρ_0 descrive la statistica dello stato termodinamico iniziale: $\rho_0(\mathbf{X}) d\mathbf{X}$, dal punto di vista delle repliche, fornisce la frazione di sistemi che si trovano in un volumetto $d\mathbf{X}$ intorno a \mathbf{X} ; alternativamente, dal punto di vista di un singolo sistema, dà la probabilità di trovarlo in un intorno di \mathbf{X} all'inizio del processo. Ci aspettiamo che, essendo la situazione iniziale di non equilibrio, le conseguenti variazioni macroscopiche nel sistema saranno tali che $\rho_0(\mathbf{X}) d\mathbf{X}$ non darà più una buona descrizione di esso. Tuttavia sappiamo che, una volta raggiunto l'equilibrio finale, la densità adatta alla descrizione statistica del sistema è quella microcanonica: la probabilità di trovare il sistema in un intorno di \mathbf{X} è $\rho_{mc}(\mathbf{X}) d\mathbf{X}$. La differenza sostanziale tra $\rho_0(\mathbf{X})$ e $\rho_{mc}(\mathbf{X})$ è che la prima dipende da come è stato preparato il sistema e cambia con il tempo, la seconda no, ciò che ne fa una distribuzione di equilibrio.

Il problema della Meccanica Statistica si può quindi porre anche nel modo seguente:

- data una arbitraria densità iniziale $\rho_0(\mathbf{X})$ si ricavi la legge che guida la sua evoluzione;

- si determini la forma delle $\rho_{st}(\mathbf{X})$, le densità indipendenti dal tempo o stazionarie, quelle adatte a descrivere gli stati di equilibrio;
- si mostri, se possibile, che ogni $\rho_0(\mathbf{X})$ ha come limite qualche $\rho_{st}(\mathbf{X})$.

Il nostro interesse è per i sistemi Hamiltoniani, ma il problema degli ensembles qui sopra esposto è riferibile a un qualunque sistema dinamico, anche con evoluzione a tempi discreti (vedi oltre). Nel caso di un sistema Hamiltoniano mentre i primi due punti hanno soluzione abbastanza semplice, il terzo è decisamente complicato da affrontare; per cui quasi sempre si adotta un atteggiamento pragmatico: essendo note le densità stazionarie si assume che un sistema in equilibrio sia descritto da una opportuna di esse.

Restiamo in un contesto generale, considerando un sistema i cui stati sono univocamente identificati da un vettore \mathbf{X} di dimensione d , che si evolve secondo la legge

$$\dot{\mathbf{X}} \equiv \frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X}), \quad (3.42)$$

dove abbiamo assunto che il campo delle velocità \mathbf{F} sia indipendente dal tempo. Il caso Hamiltoniano si ottiene come caso particolare in cui la dimensione è pari, le variabili sono metà di tipo posizione e metà di tipo impulso e \mathbf{F} si ricava derivando una Hamiltoniana, vedi eq. (4) del Capitolo 1.

A un certo istante è data una distribuzione $\rho(\mathbf{X}, t)$ e vogliamo trovare la legge con cui essa cambia nel tempo. Partiamo dal fatto che l'insieme dei sistemi che stiamo studiando per mezzo della $\rho(\mathbf{X}, t)$ si deve conservare nel tempo, per cui, ad ogni istante, deve essere $\int d\mathbf{X} \rho(\mathbf{X}, t) = 1$. Possiamo ragionare, per esempio, come nel caso della carica elettrica: se ρ è pensata come una densità di carica, la conservazione della carica totale si può esprimere in forma differenziale come $\partial\rho/\partial t + \nabla \cdot (\rho \mathbf{v}) = 0$, mediante l'introduzione della densità di corrente ($\rho \mathbf{v}$). Questa equazione trasforma la legge di conservazione globale in una locale, affermando che la variazione della carica in ogni dato volume, in assenza di sorgenti o pozzi, è determinata dal bilanciamento tra il flusso di carica entrante e quello uscente. Per la densità di probabilità il vettore velocità è dato dalla (3.42): $\dot{\mathbf{X}} = \mathbf{F}$, per cui, ponendo $\partial/\partial X_i = \partial_i$, la conservazione della probabilità (o delle repliche) si scrive:

$$\frac{\partial \rho}{\partial t} = -\nabla_{\mathbf{X}} \cdot (\rho \dot{\mathbf{X}}) = -\sum_{i=1}^d [(\partial_i \rho) \dot{X}_i + \rho (\partial_i \dot{X}_i)] \quad (3.43)$$

o anche

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^d (\partial_i \rho) \dot{X}_i = -\rho \sum_{i=1}^d (\partial_i \dot{X}_i). \quad (3.44)$$

L'equazione trovata regola l'evoluzione delle densità per un sistema dinamico generico, e assume una forma particolarmente interessante nel caso di sistemi in cui la divergenza del campo di velocità è nulla (sistemi *conservativi*):

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^d (\partial_i \rho) \dot{X}_i = \frac{d\rho}{dt} = 0 \quad \text{se} \quad \sum_{i=1}^d (\partial_i \dot{X}_i) = \sum_{i=1}^d (\partial_i F_i) = 0; \quad (3.45)$$

per questi sistemi la densità di probabilità è costante lungo le traiettorie. I sistemi Hamiltoniani sono conservativi infatti, usando le equazioni di Hamilton, si ha:

$$\sum_{i=1}^{3N} \left(\frac{\partial}{\partial p_i} \dot{p}_i + \frac{\partial}{\partial q_i} \dot{q}_i \right) = \sum_{i=1}^{3N} \left(\frac{\partial}{\partial p_i} \left(- \frac{\partial H}{\partial q_i} \right) + \frac{\partial}{\partial q_i} \left(\frac{\partial H}{\partial p_i} \right) \right) = 0. \quad (3.46)$$

L'evoluzione delle densità per i sistemi Hamiltoniani è quindi:

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0, \quad (3.47)$$

che, introducendo la parentesi di Poisson tra ρ e H

$$\{\rho, H\} = \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right), \quad (3.48)$$

diventa

$$\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0. \quad (3.49)$$

L'equazione (3.49), che è detta *equazione di Liouville* si trova anche scritta

$$\frac{\partial \rho}{\partial t} + \mathcal{L}\rho = 0 \quad (3.50)$$

se si definisce l'*operatore di Liouville* nel modo seguente:

$$\mathcal{L} = \sum_{i=1}^{3N} \left(\frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} \right). \quad (3.51)$$

Una volta trovata la legge di evoluzione per le densità di un sistema hamiltoniano, possiamo caratterizzare le densità stazionarie. Dalla (3.49) vediamo che una ρ che non dipenda esplicitamente dal tempo deve soddisfare l'equazione

$$\{\rho, H\} = 0 = \mathcal{L}\rho, \quad (3.52)$$

cioè una densità stazionaria, ρ_{st} , deve essere un integrale del moto. Arrivati a questo punto per andare avanti, analogamente a quanto fatto nei capitoli precedenti, si assume che l'unico integrale significativo sia l'energia, per cui ρ_{st} si scrive come funzione di H e che, nel caso di sistemi isolati, la densità adatta sia quella microcanonica.

Concludiamo il paragrafo accennando all'evoluzione di una densità di probabilità per un sistema dinamico con evoluzione a tempi discreti, detto anche *mappa*. In questo caso la dinamica nello spazio delle fasi in cui vive il vettore \mathbf{X} di dimensione d è regolata da una legge del tipo

$$\mathbf{X}(t+1) = \mathbf{G}[\mathbf{X}(t)], \quad (3.53)$$

con t che assume valori interi. Data una densità $\rho_t(\mathbf{X})$ essa cambierà secondo la seguente regola

$$\rho_{t+1}(\mathbf{X}) = \int \rho_t(\mathbf{Y}) \delta(\mathbf{X} - \mathbf{G}[\mathbf{Y}]) d\mathbf{Y} \equiv \mathcal{L}_{PF} \rho_t(\mathbf{X}), \quad (3.54)$$

dove \mathcal{L}_{PF} è l'operatore di Perron-Frobenius; la (3.54) è l'analogia per i sistemi a tempi discreti della (3.43), valida per i sistemi che si evolvono seguendo equazioni differenziali. Il significato della (3.54) è chiaro: al tempo $t + 1$ la densità di probabilità $\rho_{t+1}(\mathbf{X})$ è determinata dalla densità di probabilità di tutti i punti \mathbf{Y} tali che $\mathbf{X} = \mathbf{G}(\mathbf{Y})$. Le densità stazionarie per una mappa soddisfano quindi l'equazione

$$\rho_{st}(\mathbf{X}) = \mathcal{L}_{PF} \rho_{st}(\mathbf{X}), \quad (3.55)$$

cioè sono autofunzioni dell'operatore \mathcal{L}_{PF} con autovalore 1. Le densità stazionarie dei sistemi Hamiltoniani, come si vede dalla (3.52), sono invece autofunzioni con autovalore zero dell'operatore di Liouville.

3.5 Ergodicità e simulazioni numeriche

L'ipotesi ergodica, in senso stretto, è sicuramente falsa; questo è ben chiaro sia da risultati matematici, come il teorema KAM (Kolmogorov, Arnold e Moser), sia da simulazioni numeriche, a iniziare dal famoso lavoro di Fermi, Pasta ed Ulam. Tuttavia, come abbiamo visto, Khinchin è stato in grado di dimostrare che l'ergodicità, in una forma più debole, è sostanzialmente valida in sistemi con molti gradi di libertà.

Notiamo che l'ergodicità, pur con tutti i caveat a cui abbiamo precedentemente accennato, ha meriti indiscussi: primo tra tutti, permette di introdurre in modo naturale la probabilità nel contesto dei sistemi deterministici. In particolare, poiché l'ergodicità è equivalente ad assumere che diverse traiettorie abbiano le stesse proprietà asintotiche, si ha un modo non ambiguo per interpretare la probabilità in termini frequentistici. Assumendo l'ergodicità, la probabilità di un evento (definita attraverso la sua frequenza asintotica) è una proprietà oggettiva e misurabile: basta seguire una traiettoria per un tempo sufficientemente lungo. Sottolineiamo che stiamo parlando di un unico sistema (anche se ha tanti gradi di libertà) e non di un insieme di sistemi identici.

Adottando il punto di vista di Boltzmann, riteniamo corretto assumere come approccio statistico fisicamente ben fondato (a livello concettuale) quello in termini di medie ottenute seguendo l'evoluzione temporale del sistema. Notare che questo è proprio ciò che si effettua esplicitamente sia negli esperimenti di laboratorio sia nelle simulazioni numeriche (con il metodo Monte Carlo o con la dinamica molecolare). In questo modo di affrontare il problema l'introduzione degli insiemi statistici non ha niente di innaturale e misterioso²: un insieme statistico non è altro che una

² In alcuni libri di meccanica statistica si possono trovare frasi un po' oscure del tipo *Un insieme statistico è una collezione di sistemi identici*. Non è difficile convincersi, ad esempio dal libro di Gibbs *Elementary principles in statistical mechanics*, che una collezione di sistemi identici non è altro che un modo per indicare una densità di probabilità.

densità di probabilità nello spazio delle fasi, ed è introdotto solo per una questione pratica. Le medie temporali, che sono le sole fisicamente ben fondate, sono equivalenti alle medie rispetto ad una certa densità di probabilità nello spazio delle fasi, e questo permette di effettuare esplicitamente il calcolo, o almeno facilita il compito.

3.5.1 La dinamica molecolare

Discutiamo brevemente come l'ergodicità sia utilizzata in pratica nelle simulazioni numeriche. Consideriamo un sistema costituito da N particelle di massa m , soggette a un potenziale esterno U_e (prodotto, per esempio, dalle pareti della scatola in cui è contenuto il sistema), che interagiscono con un potenziale a coppie dipendente dalla distanza $U(r)$; il sistema è descritto dall'Hamiltoniana

$$H = \sum_{i=1}^N \left[\frac{\mathbf{p}_i^2}{2m} + U_e(\mathbf{q}_i) \right] + \sum_{i,j} U(|\mathbf{q}_i - \mathbf{q}_j|).$$

Il vettore $\mathbf{X} = (\mathbf{q}_1, \dots, \mathbf{q}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$ evolve seguendo la dinamica Newtoniana

$$m \frac{d^2 \mathbf{q}_j}{dt^2} = \mathbf{f}_j(\mathbf{Q}) \quad (3.56)$$

dove $\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_N)$ ed \mathbf{f}_j è la forza totale che agisce sulla j -ma particella.

Le equazioni precedenti in genere non hanno una soluzione analitica. L'idea della *Dinamica Molecolare* è concettualmente molto semplice: data una condizione iniziale $\mathbf{X}(0)$, determinare numericamente la traiettoria

$$\mathbf{X}(\Delta t), \mathbf{X}(2\Delta t), \dots, \mathbf{X}(k\Delta t), \dots, \mathbf{X}(\mathcal{M}\Delta t),$$

dove Δt è il passo di integrazione, e quindi calcolare delle medie statistiche su un intervallo di tempo $\mathcal{T} = \mathcal{M}\Delta t$:

$$\bar{A}^{\mathcal{T}} = \frac{1}{\mathcal{M}} \sum_{k=1}^{\mathcal{M}} A(\mathbf{X}(k\Delta t)).$$

Assumendo l'ergodicità, se \mathcal{T} è sufficientemente grande allora la media temporale sarà vicina a quella calcolata con la densità microcanonica; in questo modo si è in grado di ricostruire la termodinamica dalla dinamica. Per esempio, una volta fissata l'energia E , calcolando la media dell'energia cinetica, possiamo determinare la temperatura $T(E)$ e, a questo punto, ripetendo il calcolo per diversi valori dell'energia, possiamo ricavare il calore specifico C_V

$$\frac{1}{C_V} = \frac{\partial T(E)}{\partial E}$$

e quindi l'entropia

$$S(T_2) - S(T_1) = \int_{T_1}^{T_2} \frac{C_V dT}{T}.$$

3.5.2 L'algoritmo di Verlet

Per generare la traiettoria $\mathbf{X}(\Delta t), \mathbf{X}(2\Delta t), \dots, \mathbf{X}(k\Delta t), \dots, \mathbf{X}(\mathcal{M}\Delta t)$ dobbiamo usare un metodo per risolvere numericamente le Eq. (3.56). L'algoritmo più usato, che come vedremo ha molti pregi, è molto semplice ed è dovuto a Verlet uno dei pionieri delle simulazioni numeriche in meccanica statistica. Solo per semplicità di notazione consideriamo il caso unidimensionale con massa unitaria (il lettore può facilmente ripetere i calcoli per il caso generico)

$$\frac{d^2x(t)}{dt^2} = f(x(t)).$$

Sviluppiamo $x(t + \Delta t)$ in serie di Taylor:

$$x(t + \Delta t) = x(t) + \Delta t \frac{d}{dt}x(t) + \frac{1}{2}\Delta t^2 \frac{d^2}{dt^2}x(t) + \frac{1}{6}\Delta t^3 \frac{d^3}{dt^3}x(t) + O(\Delta t^4)$$

analogamente

$$x(t - \Delta t) = x(t) - \Delta t \frac{d}{dt}x(t) + \frac{1}{2}\Delta t^2 \frac{d^2}{dt^2}x(t) - \frac{1}{6}\Delta t^3 \frac{d^3}{dt^3}x(t) + O(\Delta t^4)$$

quindi con precisione $O(\Delta t^4)$ abbiamo:

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + \Delta t^2 f(x(t)).$$

Questo metodo di integrazione è di facile utilizzo ed ha la notevole proprietà di rispettare la natura Hamiltoniana del sistema. Indicando con x e p la posizione e l'impulso al tempo t e con x' e p' la analoghe quantità al tempo $t + \Delta t$, l'algoritmo di Verlet può essere scritto nella forma

$$x' = x + \Delta t p, \quad p' = p + \Delta t f(x'). \quad (3.57)$$

È facile mostrare che l'evoluzione dal tempo t a $t + \Delta t$: $(x, p) \rightarrow (x', p')$ è una trasformazione canonica, infatti introducendo la funzione generatrice

$$S(x', p) = x'p - \frac{1}{2}\Delta t p^2 + \Delta t F(x') \quad \text{con} \quad F(x') = \int^{x'} f(z) dz$$

si ha

$$x = \frac{\partial S}{\partial p} = x' - \Delta t p, \quad p' = \frac{\partial S}{\partial x'} = p + \Delta t f(x')$$

equivalenti alle Eq. (3.57).

Ovviamente nella dinamica molecolare ci sono molti problemi pratici da affrontare, ad esempio:

- a) la scelta della condizione iniziale³;
- b) metodi efficienti per il calcolo delle forze⁴;
- c) la stima dell'errore nel calcolo delle medie⁵;

non è questa la sede per entrare in dettagli, diciamo solo che è necessario un certo tirocinio.

3.5.3 Il metodo Monte Carlo

Un'altra tecnica molto usata nelle simulazioni numeriche è il metodo Monte Carlo che, per semplicità, discutiamo solo nel caso di stati discreti. Il problema è il seguente: si vuole calcolare la quantità

$$\langle Q \rangle = \frac{1}{M} \sum_{j=1}^M Q_j p_j,$$

dove p_j è la probabilità dello stato j . Se M è molto grande il calcolo può essere proibitivo; per esempio in meccanica statistica il valore di M è enorme anche in sistemi piccoli: se consideriamo un modello sul reticolo (come il modello di Ising) in cui in ogni sito ci sono due possibilità, con solo 100 siti si ha $M = 2^{100} \simeq 10^{30}$.

L'idea del Monte Carlo è di rimpiazzare la media precedente con una media temporale ottenuta seguendo un viaggiatore che salta tra gli stati di una catena di Markov che hanno probabilità invarianti $\{p_j\}$:

$$\frac{1}{\mathcal{M}} \sum_{i=1}^{\mathcal{M}} Q_{j_i}$$

dove $j_1, j_2, \dots, j_{\mathcal{M}}$ indica la successione delle posizioni del viaggiatore. Se la catena di Markov è ergodica (condizione che si può accertare dalle probabilità di transizione tra gli stati $\{P_{i \rightarrow j}\}$) allora per \mathcal{M} grande si ha che la media temporale tende a $\langle Q \rangle$. Notare che nel metodo Monte Carlo c'è un'ampia arbitrarietà, infatti non ci sono vincoli particolari nella scelta delle probabilità di transizione, purché abbiano

³ Chiaramente si deve evitare di prendere una condizione iniziale atipica, cioè lontana da quelle di equilibrio. Per esempio nel caso di un liquido è opportuno, per evitare transienti non significativi, prendere le posizioni delle particelle uniformemente distribuite e le velocità con una distribuzione di Maxwell-Boltzmann.

⁴ In un sistema di N particelle con interazioni a coppie per calcolare le forze che agiscono su ogni particella, se si procede in modo ingenuo, si devono effettuare $O(N^3)$ operazioni.

⁵ Nella stima dell'errore la difficoltà principale è dovuta al fatto che le variabili $\mathbf{X}(\Delta t), \mathbf{X}(2\Delta t), \dots, \mathbf{X}(k\Delta t), \dots, \mathbf{X}(\mathcal{M}\Delta t)$ non sono indipendenti, quindi l'errore sulla media dipende dal tempo caratteristico di correlazione τ_c (che può essere stimato dall'andamento delle funzioni di correlazione), e l'incertezza sulla media è $O(\sqrt{\tau_c/\mathcal{T}})$.

come probabilità invarianti le $\{p_j\}$ e la catena di Markov sia ergodica. A causa di questa libertà di scelta⁶, al contrario di quanto avviene nella dinamica molecolare, le proprietà dinamiche (come le funzioni di correlazione) generate con il Monte Carlo non hanno alcun significato fisico.

Naturalmente, sia nella dinamica molecolare che nel metodo Monte Carlo, in sistemi con tante particelle la traiettoria non può essere tanto lunga da visitare tutto lo spazio delle fasi (abbiamo visto che sarebbe necessario un tempo enorme che cresce esponenzialmente con il numero di particelle). Ci si potrebbe quindi domandare quale sia il segreto della loro efficacia. In poche parole possiamo dire che i loro punti di forza sono nel fatto che tipicamente vengono calcolate medie di quantità non troppo strane, diciamo del tipo di quelle individuate da Khinchin; nel metodo Monte Carlo si ha inoltre che il viaggiatore non perde tempo ad avventurarsi in stati con probabilità troppo bassa.

Concludiamo con un rapido cenno al fatto che le simulazioni numeriche non sono un mero strumento per produrre numeri, spesso siamo di fronte a veri e propri esperimenti numerici che permettono di scoprire fenomeni non scontati e magari in contrasto con idee radicate. Come esempi possiamo citare il celebre esperimento di Fermi, Pasta ed Ulam⁷, che ha mostrato come una piccola perturbazione di un sistema integrabile in genere non è sufficiente a rendere ergodico un sistema Hamiltoniano, e la presenza delle così dette code idrodinamiche: nei liquidi alcune funzioni di correlazione non decadono esponenzialmente (come per molto tempo si è creduto) bensì con una legge a potenza.

3.6 Letture consigliate

- L. Peliti *Appunti di Meccanica Statistica* (Bollati Boringhieri, Torino 2009);
- L.E. Reichl *A modern Course in Statistical Physics* (Wiley, 1998);
- O. Häggström *Finite Markov Chains and Algorithmic Applications* (Cambridge University Press, 2002);
- M. Falcioni e A. Vulpiani *Il contributo di Enrico Fermi ai sistemi non lineari: L'influenza di un articolo mai pubblicato in Conoscere Fermi* Ed. C. Bernardini e L. Bonolis, pag. 274 (Editrice Compositori, 2001).

⁶ L'algoritmo che tipicamente viene utilizzato è quello chiamato di Metropolis originariamente pubblicato in N. Metropolis, A. Rosenbluth, M. Rosenbluth, M. Teller and E. Teller *Equation of state calculations by fast computing machines* J. Chem. Phys. **21**, 1087 (1953); ma non è certo l'unico, un altro particolarmente potente è stato introdotto da Swendsen e Wang.

⁷ Sarebbe forse più corretto aggiungere tra gli autori la programmatrice Mary Tsingou.

Il problema dell'irreversibilità macroscopica

Sommario In questo capitolo ritorniamo sul problema di come riconciliare l'irreversibilità in termodinamica con la reversibilità delle equazioni del moto microscopico. Discuteremo l'equazione di Boltzmann, le obiezioni su di essa sollevate da Zermelo e Loschmidt e le risposte fornite da Boltzmann. Illustreremo un semplice modello probabilistico, dovuto agli Ehrenfest, in cui le intuizioni di Boltzmann sono state dimostrate in modo chiaro e comprensibile.

4.1 Il problema dell'irreversibilità

Nella vita di tutti i giorni incontriamo continuamente fenomeni che suggeriscono, in modo evidente, l'esistenza di processi naturali intrinsecamente irreversibili. Basti pensare allo sciogliersi di una zolletta di zucchero nel caffè, o il frantumarsi di un piatto sul pavimento.

Se si effettua la ripresa cinematografica di un evento, come la rottura di un bicchiere, l'accensione di un fiammifero, o il tuffo di un nuotatore da un trampolino e si proietta la pellicola in senso inverso si ha la sensazione di qualcosa di impossibile (ed anche un po' ridicolo): i pezzi del bicchiere si riuniscono, la cenere scompare e spunta la fiammella riformando il fiammifero, tutti gli schizzi d'acqua si ricompongono e dalla superficie della piscina tornata piatta riemerge il tuffatore. Al contrario nel filmato di fenomeni elementari, come le oscillazioni di un pendolo o gli urti tra sfere elastiche, non si nota niente di innaturale nella proiezione in senso inverso. Sembra quindi esistere una netta distinzione tra i fenomeni reversibili e quelli irreversibili e, di conseguenza, una separazione netta tra le leggi del moto che reggono le due diverse classi di fenomeni. Da una parte le leggi di evoluzione che descrivono fenomeni come la conduzione del calore e la diffusione, con una struttura intrinsecamente irreversibile, e il secondo principio della termodinamica, che stabilisce l'esistenza di una ben definita direzionalità temporale: il calore non passa mai spontaneamente dal corpo più freddo a quello più caldo, le sostanze in soluzione diffondono sempre dalle zone a concentrazione più elevata verso quel-

le con concentrazione più bassa e così via. Dall'altra le equazioni del moto, come quelle di Newton, che regolano fenomeni "semplici", come i pendoli, ma anche il moto dei pianeti, per i quali l'evoluzione inversa (cioè il film proiettato al contrario) è indistinguibile da quella diretta.

Ovviamente tutto questo pone un evidente problema se si accetta che i costituenti ultimi della materia siano atomi (o molecole) che interagiscono obbedendo alla meccanica di Newton (una trattazione del problema in termini quantistici non modifica in modo rilevante gli aspetti concettuali del problema). Poiché le leggi fondamentali della fisica (sia classica che quantistica) hanno una struttura reversibile, la dicotomia tra fenomeni irreversibili e leggi fondamentali reversibili non può essere accettata a cuor leggero.

Il secondo principio della dinamica specifica le equazioni che regolano l'evoluzione di un sistema di N particelle, di massa m_1, m_2, \dots, m_N , interagenti tra loro:

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} = \mathbf{F}_i(\mathbf{x}_1, \dots, \mathbf{x}_N) = - \frac{\partial U(\mathbf{x}_1, \dots, \mathbf{x}_N)}{\partial \mathbf{x}_i}. \quad (4.1)$$

Facciamo evolvere il sistema descritto dalle equazioni (4.1), a partire da una certa condizione iniziale $(\mathbf{x}_1(0), \dots, \mathbf{x}_N(0); \mathbf{v}_1(0), \dots, \mathbf{v}_N(0))$ fino ad un certo tempo $t > 0$. Questo equivale a proiettare il film in avanti. Al tempo t si "inverte il tempo": lasciando invariate le posizioni $\mathbf{x}_1(t), \dots, \mathbf{x}_N(t)$ si invertono le velocità sostituendo le $\mathbf{v}_1(t), \dots, \mathbf{v}_N(t)$ con $-\mathbf{v}_1(t), \dots, -\mathbf{v}_N(t)$, e si faccia evolvere nuovamente il sistema; questa operazione è l'analogo matematico del proiettare il film all'indietro. Poiché le equazioni di Newton sono invarianti rispetto alla trasformazione di inversione temporale

$$t \rightarrow -t, \quad \mathbf{x}_i \rightarrow \mathbf{x}_i, \quad \mathbf{v}_i \rightarrow -\mathbf{v}_i, \quad (4.2)$$

l'evoluzione diretta e quella inversa sono ugualmente possibili: il sistema ripercorrerà all'indietro la sua storia e, al tempo t dopo l'"inversione temporale", ritornerà nella stessa posizione iniziale ma con le velocità invertite.

Anche la meccanica quantistica è simmetrica rispetto ad inversione temporale, infatti l'equazione di Schrödinger

$$i \frac{\hbar}{2\pi} \partial_t \psi = \hat{H} \psi \quad (4.3)$$

è invariante rispetto alla trasformazione

$$t \rightarrow -t, \quad \psi \rightarrow \psi^*. \quad (4.4)$$

Il dibattito sulla freccia del tempo è stato spesso acceso e molte sono state le proposte (anche diametralmente opposte), non è questa la sede per una discussione dettagliata di questo problema. Da una parte non è facile rinunciare alla reversibilità delle leggi fondamentali, ma pochi (anche se tra questi pochi c'è A. Einstein) riescono ad accettare che *la distinzione fra passato, presente e futuro è solo un'illusione, anche se tenace*. Dall'altra parte è difficile abbandonare la percezione intuitiva di una realtà che si sviluppa ed evolve suggerendo che l'irreversibilità non sia qualcosa di

meramente soggettivo. Chi avrà mai il coraggio di dire che Hiroshima e Auschwitz sono state illusioni?

4.2 L'equazione di Boltzmann

Il primo (e nell'opinione di molti il maggiore) contributo al problema di come conciliare l'irreversibilità a livello macroscopico con la reversibilità a livello microscopico è dovuto a L. Boltzmann, con il suo teorema H e i successivi sviluppi originati dall'acceso dibattito nato in risposta alle varie critiche che gli vennero mosse.

Boltzmann cercò di affrontare i seguenti problemi:

- a) ottenere, nell'ambito della meccanica newtoniana la distribuzione di Maxwell-Boltzmann (MB) per le velocità delle molecole di un gas in equilibrio termodinamico (sarà chiaro in seguito che oltre alle leggi della dinamica è necessario introdurre ipotesi aggiuntive);
- b) mostrare, sempre nell'ambito della meccanica classica, che partendo da una generica distribuzione di probabilità, asintoticamente si ottiene sempre la distribuzione di MB.

Consideriamo un gas di N particelle identiche di massa m contenute in una scatola (di volume V) con pareti lisce ed elastiche (oppure con condizioni periodiche ai bordi). Le particelle sono soggette a:

- I) una forza esterna $\mathbf{F}(\mathbf{x})$ che agisce sulla singola particella nella posizione \mathbf{x} ;
- II) un'interazione a coppie a corto raggio, i.e. diversa da zero solo se $|\mathbf{x}_i - \mathbf{x}_j| \leq \sigma$.

Si può, ad esempio, pensare ad un gas di sfere dure di raggio r , allora il potenziale è ∞ se $|\mathbf{x}_i - \mathbf{x}_j| \leq 2r$ e zero se $|\mathbf{x}_i - \mathbf{x}_j| > 2r$.

Assumiamo inoltre che il gas sia diluito, cioè σ sia molto minore della tipica distanza tra particelle $\ell = (V/N)^{1/3}$.

Introduciamo la funzione di distribuzione $f(\mathbf{x}, \mathbf{v}, t)$, definita su ogni stato microscopico \mathbf{X} del gas e tale che $f(\mathbf{x}, \mathbf{v}, t)d\mathbf{x}d\mathbf{v}$ dà la frazione di molecole che si trovano nel volumetto $d\mathbf{x}d\mathbf{v}$, dello spazio delle fasi di singola molecola, con velocità \mathbf{v} nel punto \mathbf{x} al tempo t ; per ragioni storiche si usano le variabili (\mathbf{x}, \mathbf{v}) , sarebbe più coerente con il formalismo della meccanica analitica usare (\mathbf{x}, \mathbf{p}) .

Cerchiamo l'equazione di evoluzione per la f . Nel caso non siano presenti interazioni tra particelle e si abbia solo la forza esterna $\mathbf{F}(\mathbf{x})$ si ha

$$\frac{\partial f}{\partial t} + (\mathbf{v} \cdot \nabla)f + \left(\frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}}\right)f = 0 \quad (4.5)$$

dove $\nabla_{\mathbf{v}}$ è il gradiente rispetto alle componenti della velocità.

La dimostrazione della formula precedente è semplice. Consideriamo un volume $\Delta V(t)$ centrato intorno al punto (\mathbf{x}, \mathbf{v}) . Dopo un tempo Δt lo stato (\mathbf{x}, \mathbf{v}) si sarà

evoluto nello stato

$$(\mathbf{x} + \mathbf{v}\Delta t, \mathbf{v} + \frac{\mathbf{F}}{m}\Delta t). \quad (4.6)$$

Per la conservazione del numero delle molecole abbiamo

$$f(\mathbf{x}, \mathbf{v}, t)\Delta V(t) = f(\mathbf{x} + \mathbf{v}\Delta t, \mathbf{v} + \frac{\mathbf{F}}{m}\Delta t, t + \Delta t)\Delta V(t + \Delta t), \quad (4.7)$$

mentre per la conservazione del volume dello spazio delle fasi $\Delta V(t) = \Delta V(t + \Delta t)$, quindi

$$f(\mathbf{x}, \mathbf{v}, t) = f(\mathbf{x} + \mathbf{v}\Delta t, \mathbf{v} + \frac{\mathbf{F}}{m}\Delta t, t + \Delta t) = \quad (4.8)$$

$$f(\mathbf{x}, \mathbf{v}, t) + \Delta t \left(\frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial t} + \sum_j v_j \frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial x_j} + \sum_j \frac{F_j}{m} \frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial v_j} \right) + O(\Delta t^2), \quad (4.9)$$

da cui la (4.5).

Che succede in presenza di interazione? L'equazione (4.5) sarà modificata in

$$\frac{Df}{Dt} = C \quad (4.10)$$

dove C indica i contributi dovuti alle interazioni e D/Dt indica la derivata sostanziale lungo la traiettoria

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla) + \left(\frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} \right). \quad (4.11)$$

Supponendo, per semplicità, che siano assenti forze esterne e trascurando la dipendenza spaziale in f la (4.10) si riduce a

$$\frac{\partial f(\mathbf{v}, t)}{\partial t} = C. \quad (4.12)$$

Nel caso di gas diluito è sufficiente considerare la variazione di $f(\mathbf{v}, t)$ dovuta alle collisioni binarie in quanto il contributo delle collisioni che coinvolgono 3 o più particelle è trascurabile. Si hanno due tipi di collisione:

a) quelle in cui 2 particelle (diciamo quelle con indice 1 e 2) collidono in modo tale che dopo l'urto una particella ha velocità \mathbf{v}

$$(\mathbf{v}'_1, \mathbf{v}'_2) \rightarrow (\mathbf{v}_1, \mathbf{v}_2) \quad \text{con} \quad \mathbf{v}_1 = \mathbf{v}, \quad (4.13)$$

ovviamente bisogna sommare su tutte le possibili collisioni compatibili con la dinamica;

b) quelle in cui la particella 1 con velocità $\mathbf{v}_1 = \mathbf{v}$ collide con un'altra

$$(\mathbf{v}_1, \mathbf{v}_2) \rightarrow (\mathbf{v}'_1, \mathbf{v}'_2). \quad (4.14)$$

Naturalmente collisioni di tipo a) danno un contributo positivo a C e fanno aumentare $f(\mathbf{v}, t)$ mentre quelle di tipo b) contribuiscono negativamente a C e fanno diminuire $f(\mathbf{v}, t)$. Il termine collisione è leggermente improprio, ma è giustificato dal fatto che nel caso di gas diluito ogni particella interagisce con le altre solo una frazione di tempo molto breve.

Per semplicità di notazione consideriamo il caso in cui la velocità può assumere solo valori discreti, ed indichiamo con $f_i(t)$ il corrispondente discreto di $f(\mathbf{v}, t)$. Poiché ci sono solo collisioni binarie l'equazione per $f_i(t)$ deve avere la seguente struttura:

$$\frac{df_i}{dt} = \sum_{k,l,j} T_{(k,l) \rightarrow (i,j)} - \sum_{k,l,j} T_{(i,j) \rightarrow (k,l)}, \quad (4.15)$$

dove il primo e secondo termine a destra dell'equazione rappresentano, rispettivamente, il contributo delle collisioni di tipo a) e quelle di tipo b), e $T_{(k,l) \rightarrow (i,j)}$ è l'apporto della "collisione" $(k, l) \rightarrow (i, j)$.

Il punto delicato è come determinare $T_{(k,l) \rightarrow (i,j)}$; sicuramente si può scrivere

$$T_{(k,l) \rightarrow (i,j)} = W_{(k,l) \rightarrow (i,j)} F_{k,l}^{(2)}(t) \quad (4.16)$$

dove $F_{k,l}^{(2)}(t)$ è proporzionale al numero di coppie con una particella nello stato k ed un'altra nello stato l , e $W_{(k,l) \rightarrow (i,j)}$ tiene conto della dinamica della collisione ed è indipendente da $F_{k,l}^{(2)}(t)$. In questo modo il sistema di equazioni non è chiuso e dobbiamo scrivere l'equazione di evoluzione per $F_{k,l}^{(2)}(t)$; ciò comporta l'introduzione di nuove variabili $F_{k,l,m}^{(3)}(t)$ la cui evoluzione richiede $F_{k,l,m,n}^{(4)}(t)$ etc (questo è il problema della gerarchia di BBKGY, che è presente in tutti i problemi non lineari). L'ipotesi di Boltzmann (in genere indicata col termine *caos molecolare*, in tedesco *Stoßzahlansatz*) è che sia lecito fattorizzare $F_{k,l}^{(2)}(t)$ come prodotto di $f_k(t)$ ed $f_l(t)$:

$$T_{(k,l) \rightarrow (i,j)} = W_{(k,l) \rightarrow (i,j)} f_k(t) f_l(t), \quad (4.17)$$

con $W_{(k,l) \rightarrow (i,j)}$ indipendente da $f_k(t)$ ed $f_l(t)$. Vedremo che questa assunzione è piuttosto delicata. Non è difficile mostrare che, se il potenziale di interazione tra due particelle dipende solo dalla mutua distanza, si ha:

$$W_{(k,l) \rightarrow (i,j)} = W_{(i,j) \rightarrow (k,l)}; \quad (4.18)$$

questo segue dal fatto che, nella dinamica newtoniana, la "collisione inversa" $(k', l') \rightarrow (i', j')$ (ove l'apice indica l'inversione delle velocità) è equivalente a quella diretta $(k, l) \rightarrow (i, j)$ ed anche alla $(i', j') \rightarrow (k', l')$.

Introduciamo la funzione $H(t)$:

$$H(t) = \sum_i f_i(t) \ln f_i(t). \quad (4.19)$$

Con calcoli elementari (ricordarsi che $\sum_i f_i(t) = 1$) si ottiene

$$\frac{dH(t)}{dt} = \sum_i \frac{df_i(t)}{dt} \ln f_i(t), \quad (4.20)$$

$$\frac{dH(t)}{dt} = \sum_{i,j,k,l} W_{(k,l) \rightarrow (i,j)} \{f_k(t)f_l(t) - f_i(t)f_j(t)\} \ln f_i(t). \quad (4.21)$$

Utilizzando la simmetria tra le collisioni $(k,l) \rightarrow (i,j)$ e $(k,l) \rightarrow (j,i)$ si ha

$$\frac{dH(t)}{dt} = \frac{1}{2} \sum_{i,j,k,l} W_{(k,l) \rightarrow (i,j)} \{f_k(t)f_l(t) - f_i(t)f_j(t)\} \ln[f_i(t)f_j(t)], \quad (4.22)$$

ed usando la simmetria tra $(k,l) \rightarrow (i,j)$ e $(i,j) \rightarrow (k,l)$ si ottiene

$$\begin{aligned} \frac{dH(t)}{dt} = & \\ & -\frac{1}{4} \sum_{i,j,k,l} W_{(k,l) \rightarrow (i,j)} \{f_i(t)f_j(t) - f_k(t)f_l(t)\} \{\ln[f_i(t)f_j(t)] - \ln[f_k(t)f_l(t)]\}. \end{aligned} \quad (4.23)$$

È facile dimostrare che $dH/dt \leq 0$ (teorema H), infatti $W_{(k,l) \rightarrow (i,j)} \geq 0$ inoltre $(\ln x - \ln y)(x - y) \geq 0$ per ogni $x > 0$ e $y > 0$ ¹. Abbiamo quindi che $dH/dt \leq 0$, ed è zero solo quando le $\{f_i\}$ soddisfano l'equazione

$$f_i^{(s)} f_j^{(s)} = f_k^{(s)} f_l^{(s)}. \quad (4.25)$$

A questo punto torniamo all'equazione per variabili continue nella forma semplificata (4.12) che scriviamo

$$\begin{aligned} \partial_t f(\mathbf{v}_1, t) = & \\ & \int W[(\mathbf{v}'_1, \mathbf{v}'_2) \rightarrow (\mathbf{v}_1, \mathbf{v}_2)] \{f(\mathbf{v}'_1, t)f(\mathbf{v}'_2, t) - f(\mathbf{v}_1, t)f(\mathbf{v}_2, t)\} d\mathbf{v}'_1 d\mathbf{v}'_2 d\mathbf{v}_2. \end{aligned} \quad (4.26)$$

È importante sottolineare il fatto che l'espressione esplicita di $W[(\mathbf{v}'_1, \mathbf{v}'_2) \rightarrow (\mathbf{v}_1, \mathbf{v}_2)]$ per molti aspetti è irrilevante; ad esempio per la forma della distribuzione stazionaria e per il teorema H. La (4.26) è l'equazione di Boltzmann, chiave fondamentale per la teoria cinetica. Da essa segue che l'equazione per determinare la distribuzione

¹ Consideriamo la funzione

$$G(x, y) = (\ln x - \ln y)(x - y) = y \ln \frac{x}{y} (x - y), \quad (4.24)$$

se $x/y > 1$ allora $\ln(x/y) > 0$ e $((x/y) - 1) > 0$, analogamente se $x/y \leq 1$ allora $\ln(x/y) \leq 0$ e $((x/y) - 1) \leq 0$, quindi $G(x, y) \geq 0$.

stazionaria è

$$f^{(s)}(\mathbf{v}_1)f^{(s)}(\mathbf{v}_2) = f^{(s)}(\mathbf{v}'_1)f^{(s)}(\mathbf{v}'_2) \quad (4.27)$$

o equivalentemente

$$\ln f^{(s)}(\mathbf{v}_1) + \ln f^{(s)}(\mathbf{v}_2) = \ln f^{(s)}(\mathbf{v}'_1) + \ln f^{(s)}(\mathbf{v}'_2). \quad (4.28)$$

Ricordiamo che nelle collisioni binarie si devono conservare impulso ed energia e quindi, per particelle identiche (o di uguale massa), deve essere:

$$\mathbf{v}'_1 + \mathbf{v}'_2 = \mathbf{v}_1 + \mathbf{v}_2 \quad e \quad |\mathbf{v}'_1|^2 + |\mathbf{v}'_2|^2 = |\mathbf{v}_1|^2 + |\mathbf{v}_2|^2, \quad (4.29)$$

per cui:

$$\ln f^{(s)}(\mathbf{v}) = a + bv_x + cv_y + dv_z + e|\mathbf{v}|^2 \quad (4.30)$$

con a, b, c, d, e costanti da determinare. Sfruttando le proprietà di simmetria (invarianza per rotazione) e in assenza di una corrente media, si ottiene

$$f^{(s)}(\mathbf{v}) = A e^{-B|\mathbf{v}|^2} \quad (4.31)$$

cioè proprio la distribuzione di MB; ovviamente $B = m/(2k_B T)$ ove k_B è la costante di Boltzmann, T la temperatura ed A è determinata dalla condizione di normalizzazione $\int f^{(s)}(\mathbf{v})d\mathbf{v} = 1$, $A = (B/\pi)^{3/2}$.

Dall'equazione di Boltzmann segue quindi che una generica distribuzione di probabilità iniziale $f(\mathbf{v}, 0)$ tende alla distribuzione di MB. Se poniamo l'entropia, S , proporzionale a $-H$, con $H = \int f(\mathbf{v}) \ln f(\mathbf{v})d\mathbf{v}$, si ha che essa non diminuisce mai ed è massima in corrispondenza della distribuzione di Maxwell-Boltzmann. A questo punto sembrerebbe che il teorema H fornisca una dimostrazione del secondo principio della termodinamica, una delle leggi fondamentali della fisica macroscopica, in termini della teoria cinetica.

Ma le cose non sono affatto così semplici...

4.2.1 Due paradossi apparentemente insormontabili

La soluzione proposta da Boltzmann per il problema dell'irreversibilità si scontra con i paradossi della ricorrenza (dovuto a E. Zermelo), e della reversibilità (dovuto a J. Loschmidt). Zermelo notò che il teorema H è in disaccordo con il teorema della ricorrenza di Poincaré: in ogni sistema meccanico, il cui moto avviene in una regione limitata dello spazio delle fasi, a un certo tempo finito T_R il sistema ritorna vicino, quanto si vuole, alle condizioni iniziali. Poiché la funzione di distribuzione f , e quindi H , dipende dalle posizioni e dalle velocità delle molecole, quando il sistema torna, dopo un tempo T_R , vicino al suo stato iniziale anche H deve necessariamente assumere un valore prossimo al suo valore iniziale. Come conseguenza $H(t)$ non può essere una funzione monotona del tempo.

Il paradosso della ricorrenza è usualmente attribuito a Zermelo che lo formulò nel 1896, ma sicuramente non fu il primo: venne preceduto da Poincaré che, nel 1893, in uno scritto divulgativo notò:

Un teorema facile da stabilire ci dice che un mondo limitato, sottoposto alle sole leggi della meccanica, ripasserà sempre attraverso uno stato molto vicino al suo stato iniziale. Al contrario, secondo le leggi sperimentali conosciute (se si volesse attribuire ad esse valore assoluto e si volesse spingere le conseguenze fino in fondo), l'universo tende verso un certo stato finale dal quale non potrà più venir fuori. In questo stato finale, che sarà una specie di morte, tutti i corpi saranno in riposo e alla stessa temperatura.

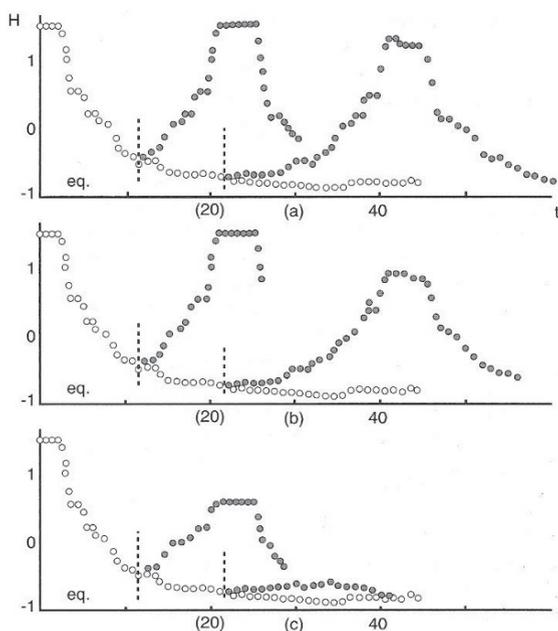
Anticipando, almeno a livello qualitativo, anche la risposta di Boltzmann all'obiezione di Zermelo:

il mondo resterà in quello stato soltanto per un periodo di tempo enorme, tanto più lungo quanto più le molecole saranno numerose.

Veniamo ora al paradosso della reversibilità: se H diminuisce dal tempo 0 al tempo t , e al tempo t si invertono le velocità, allora, a causa della simmetria delle equazioni del moto sotto inversione del tempo, il sistema ripercorrerà al contrario la sua storia e per tempi maggiori di t si deve avere un aumento di H .

La Fig. 4.1 mostra i risultati di una simulazione numerica in un sistema di dischi rigidi che collidono elasticamente.

Figura 4.1 $H(t)$ vs t in un sistema di 100 dischi elastici (con condizioni periodiche ai bordi). I pallini indicano l'evoluzione diretta, i cerchietti neri l'evoluzione della configurazione con inversione temporale dopo 50 e 100 collisioni. L'inversione viene eseguita aggiungendo un errore di 10^{-8} in a), 10^{-5} in b) e 10^{-2} in c). La figura è tratta da: J. Orban and A. Bellemans *Velocity-inversion and irreversibility in a dilute gas of hard disks* Physics Letter A **24**, 620 (1967).



4.2.2 La risposta di Boltzmann

Il problema del teorema H non è nella derivazione del risultato $dH(t)/dt \leq 0$ a partire dall'equazione di Boltzmann, derivazione matematicamente corretta, bensì nell'equazione di Boltzmann stessa, nel modo in cui l'eq. (4.26) è stata ottenuta. Lo stesso Boltzmann riuscì a dare una prima risposta (che in seguito si è rivelata sostanzialmente corretta) ai paradossi sollevati da Zermelo e Loschmidt. Per quanto riguarda il paradosso della ricorrenza, egli notò che il tempo di ricorrenza di Poincaré è, nei sistemi macroscopici, enorme lungo e di fatto non osservabile. Ad esempio in un centimetro cubo di gas, a pressione e temperatura normali, per avere un ritorno alla condizione iniziale, con una precisione di 10^{-9} m nelle posizioni e 1 m/s nelle velocità, bisognerebbe aspettare un tempo dell'ordine di $10^{10^{19}}$ anni, cioè un tempo enorme anche rispetto all'età dell'Universo. In termini moderni l'argomento di Boltzmann è una conseguenza del lemma di Kac (vedi la sezione Appendice H del capitolo con le Appendici):

in un sistema ergodico, il tempo medio di ritorno in una regione A è dato da

$$\langle \tau_A \rangle = \frac{\tau_c}{\mu(A)}, \quad (4.32)$$

dove τ_c è un tempo caratteristico e $\mu(A)$ è la percentuale del volume dello spazio delle fasi occupato dalla regione A.

Consideriamo uno spazio a d dimensioni (in un gas $d = 6N$), allora se A è un ipercubo di lato ε , si avrà $\mu(A) \sim (\varepsilon/L)^d$, ove L è la variazione tipica della singola componente del vettore di stato. Quindi si ottiene

$$\langle \tau_A \rangle \sim \tau_c C^d \quad \text{con } C > 1, \quad (4.33)$$

e i valori precisi di τ_c e C non sono importanti quando d è dell'ordine del numero di Avogadro.

Discutiamo ora il paradosso delle reversibilità, che è basato sulla possibilità di invertire² a un dato istante le velocità delle particelle con precisione infinita.

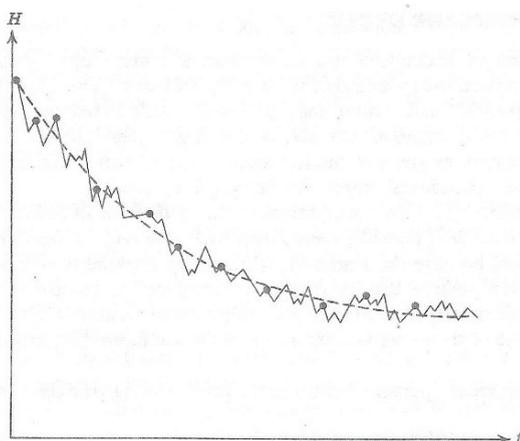
La curva $H(t)$ non è esattamente monotona, però lo è "in media" e con alta probabilità, se si è abbastanza lontani dall'equilibrio termodinamico, cioè la distribuzione di probabilità è lontana dalla MB, vedi Fig. 4.2

Nelle parole di Boltzmann

la seconda legge [della termodinamica] non può essere provata matematicamente solamente dalle equazioni della dinamica... Quello che ho dimostrato è il seguente fatto: è estremamente probabile che $H(t)$ sia molto vicino al suo valore minimo; se è più grande, allora può crescere o decrescere, ma la probabilità che decresca è [molto] maggiore

² Non è necessario determinare le velocità, si deve essere in grado di effettuare l'inversione delle variabili, come avviene, per esempio, negli esperimenti di inversione dei momenti di dipolo nucleari, cui si è accennato nella Sezione 4 del Capitolo 1.

Figura 4.2 Andamento tipico di $H(t)$ vs t in un sistema con condizioni iniziali lontane dall'equilibrio termodinamico, la curva tratteggiata mostra la previsione del teorema H . I pallini neri sono i punti in cui c 'è caos molecolare.

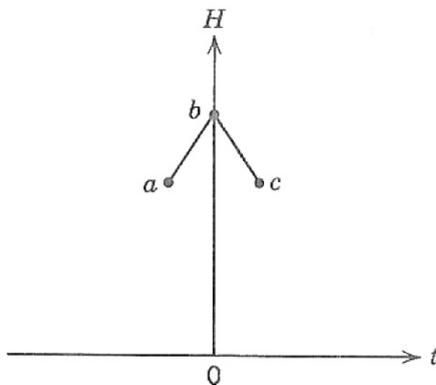


È evidente che considerare solo le collisioni binarie è legittimo nel limite di gas diluito, quindi il punto debole nella derivazione dell'equazione di Boltzmann può essere solo l'ipotesi di caos molecolare con la fattorizzazione della probabilità congiunta delle velocità delle due particelle che collidono, i.e. $F_{i,j}^{(2)} = f_i f_j$.

Se si introduce il concetto di preparazione del sistema in uno stato macroscopico, cioè si considera l'insieme delle configurazioni microscopiche che corrispondono a un dato stato macroscopico, allora la validità del teorema H è riconducibile al fatto, in qualche modo più semplice, che se lo stato macroscopico è molto lontano dall'equilibrio allora sono molto più numerose le configurazioni microscopiche che danno una diminuzione di H rispetto a quelle che ne danno un aumento.

Boltzmann fece notare che il paradosso della reversibilità non si applica alle condizioni iniziali in cui $H(t)$ è un massimo (queste sono quelle in cui l'ipotesi di caos molecolare è valida), vedi Fig. 4.3. Questo tipo di argomento può

Figura 4.3 $H(t)$ ha un picco locale quando c 'è caos molecolare.



sembrare un po' vago, ed addirittura paradossale in quanto si vuole che $H(t)$ abbia (in prevalenza) le "punte", ed allo stesso tempo sia una funzione decrescente, tuttavia è possibile mostrare in alcuni modelli probabilistici la sostanziale correttezza (che, in un opportuno limite, diventa rigore matematico) dell'intuizione di Boltzmann.

4.3 Un modello probabilistico in cui tutto è chiaro

Discutiamo brevemente il modello delle pulci su due cani, introdotto da P. e T. Ehrenfest, che pur nella sua semplicità è illuminante. Su due cani (A e B) sono distribuite N pulci (per semplicità prendiamo N pari): ad ogni secondo una pulce, scelta a caso, cambia cane. Si può pensare che ogni pulce sia numerata e che la pulce che deve effettuare il salto venga scelta con una regola probabilistica tipo tombola (cioè con l'estrazione di palline numerate che poi una volta estratte vengono nuovamente imbussolate). Si può formalizzare il modello in termini di una catena di Markov (vedi la sezione Appendice G delle Appendici) in cui le uniche transizioni permesse sono quelle $m \rightarrow m \pm 1$ con probabilità

$$P_{m \rightarrow m-1} = \frac{m}{N}, \quad P_{m \rightarrow m+1} = 1 - \frac{m}{N}, \quad (4.34)$$

ovvero, indicando con n_t il numero di pulci sul cane A al tempo t

$$P(n_{t+1} = m \pm 1 | n_t = m) = P_{m \rightarrow m \pm 1}. \quad (4.35)$$

Nelle parole di Kac questo sistema è *probabilmente uno dei più istruttivi modelli di tutta la fisica e, benché non sia altro che un esempio di catena di Markov finita, è di notevole interesse* e permette di dare un'interpretazione non ambigua delle intuizioni di Boltzmann. Ci si può domandare quale sia la probabilità stazionaria Π_n di avere n pulci sul cane A. È facile convincersi che vale l'equazione

$$\Pi_n = \Pi_{n-1} P_{n-1 \rightarrow n} + \Pi_{n+1} P_{n+1 \rightarrow n}. \quad (4.36)$$

Una ispezione diretta mostra che

$$\Pi_n = \frac{N!}{n!(N-n)!} 2^{-N} \quad (4.37)$$

e inoltre si ha la seguente relazione (detta di bilancio dettagliato):

$$\Pi_{n \pm 1} P_{n \pm 1 \rightarrow n} = \Pi_n P_{n \rightarrow n \pm 1}. \quad (4.38)$$

Π_n è massima in $N/2$ e nel limite $N \gg 1$, per valori di n tali che $|n - N/2| < \sqrt{N}$, essa è ben approssimata da una Gaussiana centrata in $N/2$.

Notiamo che in questo modello stocastico si ha l'analogo della reversibilità meccanica dato che

$$P(n_t = m_0, n_{t+1} = m_1, \dots, n_{t+k-1} = m_{k-1}, n_{t+k} = m_k) = \quad (4.39)$$

$$P(n_t = m_k, n_{t+1} = m_{k-1}, \dots, n_{t+k-1} = m_1, n_{t+k} = m_0) \quad (4.40)$$

cioè la probabilità della “traiettoria diretta” $\{n_t = m_0, n_{t+1} = m_1, \dots, n_{t+k-1} = m_{k-1}, n_{t+k} = m_k\}$ è identica a quella della “traiettoria inversa” $\{n_t = m_k, n_{t+1} = m_{k-1}, \dots, n_{t+k-1} = m_1, n_{t+k} = m_0\}$. È sufficiente scrivere

$$P(n_t = m_0, n_{t+1} = m_1, \dots, n_{t+k-1} = m_{k-1}, n_{t+k} = m_k) = \quad (4.41)$$

$$\Pi_{m_0} P_{m_0 \rightarrow m_1} P_{m_1 \rightarrow m_2} \dots P_{m_{k-1} \rightarrow m_k}$$

ed usare il bilancio dettagliato (4.38).

Si ha anche la ricorrenza in quanto, per ogni condizione iniziale $n_0 = m$, c'è probabilità non nulla che dopo un certo tempo T_m si abbia $n_{T_m} = m$.

4.3.1 Caos molecolare nel modello di Ehrenfest

Assumiamo come condizione iniziale n_0 vicino a N , cioè quasi tutte le pulci sono sul cane A . Ci si aspetta che all'aumentare del tempo n_t diminuisca, pur con qualche fluttuazione, fino a stabilizzarsi intorno ad $N/2$ con escursioni dell'ordine di \sqrt{N} . È facile mostrare che³

$$\left[\langle n_{t+1} \rangle - \frac{N}{2} \right] = \left[1 - \frac{2}{N} \right] \left[\langle n_t \rangle - \frac{N}{2} \right] \quad (4.43)$$

e quindi

$$\left[\langle n_t \rangle - \frac{N}{2} \right] = \left[1 - \frac{2}{N} \right]^t \left[\langle n_0 \rangle - \frac{N}{2} \right] \quad (4.44)$$

il numero medio al tempo t tende (in modo esponenziale) al valore medio stazionario $N/2$. Un analogo calcolo permette di calcolare la varianza σ_t^2 . Consideriamo la variabile $\eta_{t+1} = n_{t+1} - N/2 = \eta_t + \Delta_t$:

$$\eta_{t+1}^2 = \eta_t^2 + \Delta_t^2 + 2\Delta_t \eta_t \quad (4.45)$$

$$\langle \eta_{t+1}^2 \rangle = \langle \eta_t^2 \rangle + 1 + 2E_{\eta_t}(\langle \Delta_t \eta_t | \eta_t \rangle), \quad (4.46)$$

³ Basta scrivere $n_{t+1} = n_t + \Delta(n_t)$ ove $\Delta(n_t)$ vale 1 con probabilità $P_{n_t \rightarrow n_t+1} = 1 - n_t/N$ oppure -1 con probabilità $P_{n_t \rightarrow n_t-1} = n_t/N$. Mediando si ha

$$\langle n_{t+1} \rangle = \langle n_t \rangle + \langle P_{n_t \rightarrow n_t+1} \rangle - \langle P_{n_t \rightarrow n_t-1} \rangle \quad (4.42)$$

utilizzando la (4.34) si ottiene la (4.43).

dove $\langle |n_t| \rangle$ indica il valor medio condizionato al valore di n_t , e $E_{\eta_t}(\cdot)$ il valor medio rispetto alla variabile η_t . È facile calcolare $E_{\eta_t} \langle \Delta_t \eta_t | \eta_t \rangle = -(2/N) \langle \eta_t^2 \rangle$ si ha quindi

$$\langle \eta_t^2 \rangle = \frac{N}{4} + \left(1 - \frac{4}{N}\right)^t \left[\langle \eta_0^2 \rangle - \frac{N}{4} \right]. \quad (4.47)$$

A questo punto abbiamo l'espressione esplicita della varianza che può essere scritta nella forma

$$\sigma_t^2 = \langle \eta_t^2 \rangle - \frac{N^2}{4} + N \langle n_t \rangle - \langle n_t \rangle^2 \quad (4.48)$$

Se al tempo $t = 0$ il valore di n è determinato con certezza ($n_0 = m$) allora $\sigma_0^2 = 0$ e σ_t^2 cresce monotonicamente, con un tempo caratteristico $O(N/2)$, fino a $\sigma_\infty^2 = N/4$.

Introducendo la funzione $H(t) = |\langle n_t \rangle - N/2|$, la (4.44) può essere interpretata come un teorema H. Ovviamente $H(t)$ è monotona solo perché è stata effettuata l'operazione di media.

Comunque iniziando con $n_0 = m$ lontano dal valore di equilibrio, e.g. m vicino ad N , n_t fluttua tra $\langle n_t \rangle - \delta_t$ e $\langle n_t \rangle + \delta_t$ ove $\delta_t \propto \sigma_t$; quindi finché $\sigma_t \ll \sqrt{N}$, cosa che accade per $t < T_N = O(N)$, si ha $\delta_t \ll \langle n_t \rangle$ e quindi n_t , a parte qualche fluttuazione, segue l'andamento decrescente di $\langle n_t \rangle$. Naturalmente anche per $t < T_N$ ci sono dei tempi per i quali si ha un comportamento "atipico" (diciamo "antitermodinamico"), i.e. $n_{t+1} > n_t$, ma questo accade molto raramente se $N \gg 1$ e $|n_t - N/2| \gg \sqrt{N}$.

Tutto ciò può essere mostrato esplicitamente osservando che le configurazioni con

$$C_{\mathcal{A}} = \{n_{t-1} = m - 1, \quad n_t = m, \quad n_{t+1} = m - 1\} \quad (4.49)$$

sono preponderanti rispetto a tutte le altre, i.e.

$$C_{\mathcal{B}} = \{n_{t-1} = m - 1, \quad n_t = m, \quad n_{t+1} = m + 1\} \quad (4.50)$$

$$C_{\mathcal{C}} = \{n_{t-1} = m + 1, \quad n_t = m, \quad n_{t+1} = m - 1\} \quad (4.51)$$

$$C_{\mathcal{D}} = \{n_{t-1} = m + 1, \quad n_t = m, \quad n_{t+1} = m + 1\} \quad (4.52)$$

quando $N \gg 1$ e si è molto lontani dall'equilibrio, in pratica se $|m - N/2| \gg \sqrt{N}$. Notare che in questo modello le configurazioni \mathcal{A} sono quelle in cui si ha il caos molecolare (vedi Fig.4.3). È sufficiente un calcolo esplicito della probabilità della configurazione $C_{\mathcal{A}}$ condizionata all'evento $n_t = m$:

$$P_m(C_{\mathcal{A}}) = P(n_{t-1} = m - 1, n_t = m, n_{t+1} = m - 1 | n_t = m) = \quad (4.53)$$

$$\frac{\Pi_{m-1} P_{m-1 \rightarrow m} P_{m \rightarrow m-1}}{\Pi_m} = \frac{m^2}{N^2}, \quad (4.54)$$

se m è vicina ad N cioè $m = N - k$ con $k \ll N$ allora

$$P_m(C_{\mathcal{A}}) = \left(1 - \frac{k}{N}\right)^2 \simeq 1 - 2\frac{k}{N}, \quad (4.55)$$

e quindi molto vicino ad uno se $N \gg 1$. Un calcolo diretto mostra che se m è vicino a N ed $N \gg 1$ allora $P_m(C_{\mathcal{A}})$ è molto maggiore di $P_m(C_{\mathcal{B}})$, $P_m(C_{\mathcal{C}})$ e $P_m(C_{\mathcal{D}})$.

Questo fornisce un'interpretazione ed una spiegazione non ambigua dell'ipotesi (apparentemente paradossale) di Boltzmann che, nei sistemi macroscopici, i punti della curva $H(t)$ vs t sono, in stragrande maggioranza, dei massimi.

Il lettore è invitato a simulare il modello di Ehrenfest, con $N \gg 1$ e n_0 tale che risulti $|n_0 - N/2| \gg \sigma_\infty$, confrontando le singole realizzazioni di n_t con il valor medio $\langle n_t \rangle$.

4.3.2 Tempi di ricorrenza

Dato $n_0 = m$ quanto tempo si deve aspettare (in media) per aver un ritorno? La risposta a questa domanda è data dal lemma di Kac

$$\langle T_m \rangle = \frac{1}{\Pi_m}. \quad (4.56)$$

Il calcolo del tempo medio di ricorrenza è molto istruttivo per capire il significato fisico della ricorrenza in sistemi macroscopici. È facile vedere che se m è molto distante da $N/2$ allora il valore di $\langle T_m \rangle$ è enorme anche per valori di N non grandissimi. Ad esempio se $m = N$ allora $\langle T_{m=N} \rangle = 2^N$ (per $N = 10000$ si ha $\langle T \rangle = 2^{10000} \simeq 10^{3000}$); invece per m intorno ad $N/2$ il tempo di ricorrenza è molto più piccolo. Usando la formula di Stirling per $m = N/2$ si ottiene $\langle T_{m=N/2} \rangle \simeq \sqrt{\pi N/2} \ll 2^N$.

L'analisi del modello di Ehrenfest mostra in modo non vago che le intuizioni di Boltzmann sono perfettamente verificate se $N \gg 1$ e la condizione iniziale è molto lontana dall'equilibrio (ad esempio $n_0 \simeq N$):

- a) n_t è vicino a $\langle n_t \rangle$ che decresce monotonamente verso $N/2$;
- b) gli stati con caos molecolare (cioè le configurazioni $C_{\mathcal{A}}$) sono predominanti;
- c) i tempi di ricorrenza sono enormi.

Come commento finale possiamo ricordare le parole di Smoluchowski:

Un processo appare irreversibile (reversibile) quando lo stato iniziale ha un tempo di ricorrenza lungo (breve) rispetto al tempo di osservazione.

4.4 Ancora sull'irreversibilità

È importante ribadire che l'irreversibilità fisica si riferisce a un singolo sistema e non è in alcun modo una proprietà di un ensemble. Quando parliamo del secondo principio della termodinamica intendiamo che una pentola di acqua calda lasciata a se stessa si raffredda; si tratta di una singola pentola (quella nella nostra cucina) e non

di un ensemble di tante pentole. Su questo punto c'è una certa confusione, in particolare sul ruolo del caos deterministico, e spesso vengono associate all'irreversibilità proprietà che, pur interessanti, ne hanno solo una superficiale somiglianza.

Consideriamo un sistema i cui stati iniziali siano descritti da una densità di probabilità $\rho(\mathbf{x}, 0)$. Possiamo calcolare $\rho(\mathbf{x}, t)$ per $t > 0$ e domandarci cosa succede nel limite $t \rightarrow \infty$. Sotto opportune condizioni, cioè se il sistema è mixing⁴ (una proprietà più forte dell'ergodicità), per $t \rightarrow \infty$ la densità di probabilità tende ad una densità invariante

$$\rho(\mathbf{x}, t) \rightarrow \rho_{inv}(\mathbf{x}). \quad (4.57)$$

Questa proprietà è importante e interessante ma non ha niente a che vedere, nonostante quanto sostenuto da alcuni, con l'irreversibilità della fisica, infatti la (4.57) descrive l'evoluzione di un ensemble di condizioni iniziali e non il comportamento del singolo sistema.

Si può mostrare che esistono sistemi che, anche se vale la (4.57), non presentano nessun comportamento irreversibile. Per convincersene consideriamo una mappa caotica e mixing, per esempio il famoso *gatto di Arnold*:

$$x_{t+1} = x_t + y_t \text{ mod } 1, \quad y_{t+1} = y_t + x_{t+1} \text{ mod } 1,$$

e osserviamo una singola realizzazione seguendo l'evoluzione diretta

$$\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{t-1}, \mathbf{x}_t \quad (4.58)$$

e quella inversa (cioè proiettata al contrario)

$$\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_0. \quad (4.59)$$

Nonostante il sistema sia caotico e le successioni (4.58) e (4.59) appaiano irregolari, non si noterà niente di strano nell'evoluzione inversa, a differenza di quanto accade nei sistemi macroscopici.

A questo punto si potrebbe obiettare che nella discussione del modello di Ehrenfest abbiamo fatto ricorso a considerazioni di tipo probabilistico e che anche Boltzmann nel suo approccio usa delle densità di probabilità; il problema è delicato e merita un po' di attenzione.

Nel modello di Ehrenfest l'importanza del calcolo delle probabilità è nella possibilità di mostrare che per $N \gg 1$ e condizioni iniziali lontano dall'equilibrio la singola traiettoria è vicina alla previsione statistica. Nell'approccio di Boltzmann all'irreversibilità si utilizza la funzione $f(\mathbf{v}, t)$, interpretabile come una densità di probabilità empirica, tuttavia non c'è alcun ricorso agli ensemble statistici. Infatti la $f(\mathbf{v}, t)$ dà la distribuzione delle velocità in un singolo sistema con $N \gg 1$ particelle; questo è ben chiaro quando si fanno le simulazioni in cui si studia un unico sistema

⁴ Un sistema dinamico con densità invariante $\rho_{inv}(\mathbf{x})$ è mixing se per ogni coppia di insiemi A e B dello spazio delle fasi, per $t \rightarrow \infty$ si ha

$$\mu(S^{-t}A \cap B) \rightarrow \mu(A)\mu(B)$$

dove $\mu(A) = \int_A \rho_{inv}(\mathbf{x}) d\mathbf{x}$ e $S^{-t}A$ indica l'insieme dei punti che dopo un tempo t finiscono in A .

e la $f(\mathbf{v}, t)$ è calcolata con un istogramma dalle velocità delle particelle al tempo t : $\mathbf{v}_1(t), \mathbf{v}_2(t), \dots, \mathbf{v}_{N-1}(t), \mathbf{v}_N(t)$.

4.5 Considerazioni finali

Quanto visto precedentemente per il modello cani-pulci lascia sperare che sia possibile, in opportuni limiti, dimostrare la validità del teorema H . L'apparato tecnico per portare a termine questo programma di ricerca è imponente e non può certo essere presentato in questa sede. Partendo da un fondamentale lavoro di H. Grad del 1948 si è arrivati a formulare e dimostrare in modo rigoroso quanto da Boltzmann intuito con profondo senso fisico. Tra i tanti che hanno partecipato a questo significativo progresso citiamo R. Illner, O.E. Lanford, N. Shinbrot, R. DiPerna, P.-L. Lions, M. Pulvirenti e C. Cercignani. La sostanza del loro lavoro può essere riassunta come segue:

Nel limite di Grad:

$$N \rightarrow \infty \quad , \quad \sigma \rightarrow 0 \quad , \quad N\sigma^2 \rightarrow \text{cost.} \quad (4.60)$$

dove N è il numero di molecole nell'unità di volume e σ è il diametro delle molecole, la probabilità di avere il caos molecolare tende ad uno (certezza) se la condizione iniziale del sistema è fuori dall'equilibrio termodinamico, e $f(\mathbf{x}, \mathbf{v}, t)$ evolve in accordo con l'equazione di Boltzmann e quindi vale il teorema H .

Si può capire il significato fisico di $N\sigma^2 \rightarrow \text{cost.}$ osservando che l'equazione di Boltzmann nel caso di sfere dure di diametro σ è:

$$\frac{\partial f}{\partial t} = N\sigma^2 \int_+ d\mathbf{n} \int d\mathbf{v}'_2 |\mathbf{n} \cdot (\mathbf{v}_1 - \mathbf{v}_2)| \{f(\mathbf{v}'_1, t)f(\mathbf{v}'_2, t) - f(\mathbf{v}_1, t)f(\mathbf{v}_2, t)\} \quad (4.61)$$

dove N è il numero di particelle nell'unità di volume, \mathbf{n} il versore nella direzione congiungente i centri di massa delle due molecole collidenti e le velocità \mathbf{v}'_1 e \mathbf{v}'_2 devono soddisfare la regola d'urto

$$\mathbf{v}'_1 = \mathbf{v}_1 - \left((\mathbf{v}_1 - \mathbf{v}_2) \cdot \mathbf{n} \right) \mathbf{n}, \quad \mathbf{v}'_2 = \mathbf{v}_2 + \left((\mathbf{v}_1 - \mathbf{v}_2) \cdot \mathbf{n} \right) \mathbf{n}. \quad (4.62)$$

Il pedice $+$ nel primo integrale rappresenta la restrizione $\mathbf{n} \cdot (\mathbf{v}_1 - \mathbf{v}_2) < 0$ necessaria perché avvenga l'urto. La sezione urto della singola collisione è proporzionale a σ^2 , e $N\sigma^2$ è proporzionale alla rapidità di variazione di $f(\mathbf{v}, t)$, vedi eq. (4.61). Nei gas il limite di Grad è fisicamente realistico, ad esempio a temperatura ambiente e pressione atmosferica in 1 cm^3 si hanno $N \sim 10^{20}$ molecole e $\sigma \sim 10^{-8} \text{ cm}$, quindi $N\sigma^2 \sim 10^4 \text{ cm}^2 = 1 \text{ m}^2$, mentre il volume occupato dalle particelle è piccolo rispetto al volume totale, infatti $N\sigma^3 \sim 10^{-4} \text{ cm}^3$.

È impressionante l'accordo di questo risultato con quanto intuito da Boltzmann:

Nelle equazioni della meccanica non c'è niente di analogo a quanto si ha con la Seconda Legge della termodinamica che può essere ricondotta a termini meccanici solo con assunzioni sulle condizioni iniziali.

La grandezza del lavoro di Boltzmann non è tanto (come a volte viene detto) nell'aver "ridotto la termodinamica alla meccanica", quanto nell'aver compreso l'impossibilità di ricondurre l'irreversibilità alle sole leggi della meccanica, comprendendo la sottile natura dell'emergenza dell'irreversibilità per la quale sono necessari due ingredienti fondamentali:

- a) il grande numero di particelle (atomi o molecole) coinvolte e quindi la grande disparità tra scala microscopica e macroscopica;
- b) opportune condizioni iniziali (quelle per le quali si ha il caos molecolare).

Possiamo aggiungere un terzo elemento (in qualche modo connesso con b):

- c) l'uso della probabilità: non tutti gli stati microscopici evolvono in modo irreversibile ma solo la "maggior parte". Nei sistemi macroscopici, che coinvolgono un elevatissimo numero di particelle, si ha la pratica certezza di avere un comportamento irreversibile.

Nella trattazione dell'irreversibilità è essenziale considerare il problema dei livelli di realtà, cioè di come si guarda un fenomeno. Come osservato da P. Duhem, se ci fosse dato di scorgere le molecole di un fluido noteremmo un'agitazione tumultuosa sia nel caso di un fluido in regime laminare, che a livello macroscopico appare regolare, sia in un fluido turbolento che è caratterizzato da un comportamento irregolare anche a livello macroscopico. Consideriamo il seguente esperimento concettuale: si versi del profumo in un angolo di una stanza; le molecole del profumo, inizialmente concentrate in una piccola regione, velocemente occuperanno tutta la stanza. Si immagini ora di poter filmare le molecole. Proiettando la pellicola all'indietro, si vedrà un fenomeno innaturale: tutte le molecole sparse nella stanza si riuniranno in un angolo. Guardando invece una sola molecola non si noterà niente di anormale nel film proiettato al contrario. Analogamente non si nota niente di strano nel film proiettato al contrario se si limita l'osservazione alle molecole contenute in una piccola zona della stanza su un intervallo di tempo non troppo lungo. L'impressione di un comportamento innaturale si ha solo guardando un numero elevato di molecole oppure una zona abbastanza grande della stanza.

L'apparente aspetto "soggettivo" dell'irreversibilità non dovrebbe essere sottolineato in modo eccessivo. L'irreversibilità è un fenomeno reale e non può certo dipendere dalle conoscenze, o dall'abilità matematica, dell'osservatore. Anche l'uso di concetti probabilistici, pur di grande importanza, non deve far dimenticare il fatto che l'irreversibilità si riferisce ad un singolo sistema con un elevato numero di gradi di libertà.

La rilevanza delle condizioni iniziali e dei tanti gradi di libertà non è accettato da tutti; ad esempio Prigogine e la sua scuola hanno un'opinione ben diversa, come è evidente dalla netta affermazione:

l'irreversibilità o è vera a tutti i livelli o non lo è affatto: non può emergere dal nulla passando da un livello all'altro.

È questo un punto di vista (che possiamo indicare come riduzionismo di tipo estremo), a nostro avviso errato, che nega la possibilità di nuovi fenomeni nel passaggio da un livello all'altro. Al contrario l'irreversibilità non è certo un "emergere dal nulla", bensì dal limite singolare $N \rightarrow \infty$, $\sigma \rightarrow 0$, come intuito da Boltzmann e successivamente dimostrato in modo rigoroso. L'irreversibilità può essere vista come una proprietà emergente nel passaggio dal microscopico al macroscopico.

4.6 Letture consigliate

Per una presentazione non troppo difficile all'equazione di Boltzmann sono particolarmente indicati i primi capitoli di

- G. M. Kremer *An Introduction to the Boltzmann Equation and Transport* (Springer-Verlag, 2010).

Per i problemi dell'irreversibilità, equazione di Boltzmann, discussione dei paradossi e di modelli stocastici è ottimo:

- M. Kac *Probability and Related Topics in Physical Sciences* (Am. Math. Soc. 1957).

Per una presentazione dei diversi punti di vista sull'irreversibilità (incluso quello di Prigogine) è interessante

- F. Guerra *Reversibilità/Irreversibilità*, Enciclopedia Einaudi Volume XI, pagina 1067 (Torino 1980).

Per una discussione del ruolo dei limiti singolari per le proprietà emergenti in fisica si veda

- M. Berry *Asymptotics, singularities and the reduction of theories*, in *Logic, Methodology and Philosophy of Science IX* pag. 597, D. Prawitz, B. Skyrms and D. Westerstahl (Editori), (Elsevier Science B.V., 1994);
- H. Primas *Emergence in exact natural science*, Acta Polytechnica Scandinavica **Ma 91**, 83 (1998).

Studiare il mondo atomico con la meccanica statistica

Sommario L'esistenza degli atomi è adesso un fatto indiscusso e la cosa è così scontata che non tutti sono completamente consapevoli dell'enorme importanza concettuale della realtà fisica degli atomi. Ecco come si esprimeva Feynman:

Se in un cataclisma andasse distrutta tutta la conoscenza scientifica, e soltanto una frase potesse essere trasmessa alle generazioni successive, quale affermazione conterrebbe la massima quantità di informazioni nel numero minimo di parole? Io credo che sarebbe l'ipotesi atomistica (o dato di fatto atomico, o comunque vogliamo chiamarlo) secondo cui tutte le cose sono fatte di atomi, piccole particelle che si agitano con un moto perpetuo, attraendosi quando sono un po' distanti una dall'altra, ma respingendosi quando sono schiacciate una contro l'altra. In questa singola frase c'è un'enorme quantità di informazione sul mondo che ci circonda, se soltanto ci si riflette sopra con un po' di immaginazione.

L'atomismo è una grande visione del mondo, che inizia nel V secolo avanti Cristo con Leucippo e Democrito. L'idea principale è che la realtà sensibile sia descrivibile in termini di elementi (gli atomi) che costituiscono la vera essenza dei fenomeni, nascosta alla nostra percezione. Nell'antichità classica l'atomismo non ebbe grande successo e venne avversato dalle due grandi scuole filosofiche (la platonica e l'aristotelica), comunque ebbe seguaci come Epicuro e Lucrezio con il suo grande poema *De Rerum Natura*.

Una ovvia obiezione a questa idea è che gli atomi non si vedono, mentre noi sperimentiamo solo colori, forme, odori e così via. Oggi con le moderne tecnologie siamo praticamente in grado di "vedere" gli atomi e di osservarne le proprietà, ad esempio la loro grandezza e come interagiscono. È interessante il fatto che la convinzione, fondata su fortissime evidenze sperimentali, della reale esistenza degli atomi risale a tempi molto precedenti al microscopio a effetto tunnel. Vedremo come utilizzando i risultati della meccanica statistica, anche se gli atomi non si possono vedere (almeno fino a poco tempo fa) è possibile contarli, cioè misurare il numero di Avogadro.

Mostriamo inoltre come, usando un po' di meccanica statistica e concetti elementari di teoria cinetica, sia possibile determinare le dimensioni tipiche degli atomi e l'energia tipica del potenziale di interazione.

5.1 Meccanica statistica ed esperimenti

La meccanica statistica è in grado di prevedere comportamenti non banali che possono essere verificati o falsificati in esperimenti di laboratorio. Forse l'esempio più spettacolare è quello dei fenomeni critici (che presenta aspetti tecnicamente non semplici e non sarà affrontato in questa sede).

Discuteremo brevemente alcune evidenze sperimentali dirette e indirette della validità delle più rilevanti previsioni della meccanica statistica. Mostriamo inoltre come con un confronto tra dati sperimentali e previsioni teoriche sia possibile determinare le proprietà più importanti dei potenziali di interazione.

5.1.1 La distribuzione di Maxwell

La distribuzione di Maxwell per la densità di probabilità delle velocità delle molecole, benché ottenuta da Maxwell (e poi da Boltzmann, a partire dall'equazione di Boltzmann, vedi Capitolo 4) già nella seconda metà del XIX secolo, ha dovuto aspettare vari decenni per una conferma sperimentale.

Una verifica diretta

La distribuzione di Maxwell può essere misurata direttamente con esperimenti che utilizzano fasci molecolari. A grandi linee l'esperimento è il seguente. Un fascio molecolare viene generato scaldando atomi di potassio in un forno a temperatura di qualche centinaio di gradi Celsius a cui sia praticata una piccola apertura. Per selezionare le particelle con velocità v si utilizzano due diaframmi posti a distanza ℓ , che ruotano con la stessa velocità angolare ω , sui quali sono praticate due fenditure sfasate di un angolo θ , vedi Fig. 5.1. Affinchè una particella riesca a passare è necessario che il tempo di volo tra i due diaframmi ℓ/v sia uguale al tempo θ/ω che i diaframmi impiegano a ruotare di un angolo θ , quindi saranno selezionate le particelle con velocità $v = \ell\omega/\theta$.

Ricordiamo che il numero di particelle di massa m con velocità tra v e $v + dv$ è prevista essere

$$\text{cost.} \times v^2 e^{-mv^2/2k_B T} dv ; \quad (5.1)$$

poichè solo una frazione proporzionale a v riesce a passare attraverso le due fenditure abbiamo che il numero di particelle che arrivano sul rivelatore nell'unità di tempo è

$$\text{cost.} \times v^3 e^{-mv^2/2k_B T} dv . \quad (5.2)$$

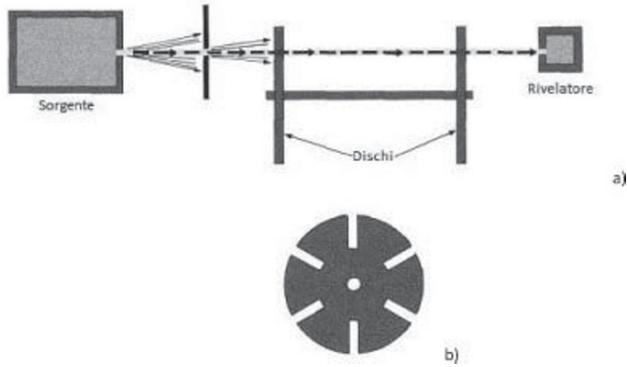


Figura 5.1 a) Schema dell'apparato per lo studio della distribuzione delle velocità molecolari. b) Vista frontale di un diaframma.

Il confronto mostra perfetto accordo tra le misure sperimentali e la distribuzione di Maxwell. Questo tipo di esperimento è piuttosto delicato, è necessario infatti usare una meccanica di precisione ed efficaci pompe per avere che all'esterno del forno ci sia praticamente il vuoto.

Allargamento delle linee spettrali

Consideriamo un gas in un cubo di lato ℓ contenente un gas di atomi di massa m a temperatura T . Ciascun atomo da fermo emette una radiazione monocromatica di lunghezza d'onda λ_0 . La radiazione viene rivelata ad una distanza $L \gg \ell$; a causa dell'effetto Doppler e dell'agitazione termica si ha un allargamento della riga spettrale.

Indicando con v la velocità nella direzione della congiungente la scatola contenente il gas e il rivelatore, se un atomo si muove con velocità v , per effetto Doppler, la lunghezza d'onda λ rivelata sarà:

$$\lambda = \lambda_0 \left(1 + \frac{v}{c}\right). \quad (5.3)$$

Poiché la densità di probabilità di v è l'usuale gaussiana (distribuzione di Maxwell)

$$p(v) = \sqrt{\frac{m}{2\pi k_B T}} e^{-mv^2/2k_B T} \quad (5.4)$$

usando la relazione $v = c(\lambda - \lambda_0)/\lambda_0$ è immediato trovare $p(\lambda)$ (la densità di probabilità di λ) e quindi l'intensità $I(\lambda)d\lambda$ della luce osservata nell'intervallo tra λ e $\lambda + d\lambda$:

$$I(\lambda)d\lambda \propto p(\lambda)d\lambda \propto \exp\left[-\frac{(\lambda - \lambda_0)^2}{\delta^2}\right]d\lambda \quad (5.5)$$

dove $\delta = \sqrt{(2k_B T \lambda_0^2)/(mc^2)}$ è detta semi larghezza Doppler della riga. La formula precedente per l'allargamento delle righe spettrali è molto importante nello studio della fisica stellare.

5.1.2 La distribuzione di Fermi-Dirac

Nello studio dell'emissione termoionica si deve calcolare la corrente di elettroni che sono emessi da una superficie metallica in funzione della temperatura. Schematizziamo gli elettroni di conduzione di un metallo come un gas di particelle libere e assumiamo che tra l'interno e l'esterno del metallo sia presente una barriera di potenziale di altezza $e\phi$; per uscire dal metallo un elettrone deve quindi avere energia cinetica sufficientemente alta che gli permetta di superare la barriera. Indicando con x la direzione perpendicolare alla superficie del metallo, gli elettroni che contribuiscono alla corrente sono solo quelli con

$$\frac{p_x^2}{2m} > e\phi, \quad (5.6)$$

cioè $p_x > p_0 = \sqrt{2me\phi}$, o $v_x > v_0 = \sqrt{2e\phi/m}$.

Sappiamo che il gas degli elettroni di conduzione in un metallo è degenere, dobbiamo quindi usare una trattazione quantistica. Usando l'approssimazione del continuo per gli stati del gas, il numero medio per unità di volume di fermioni di spin $1/2$ che si trovano in un volumetto $dp_x dp_y dp_z$ dello spazio degli impulsi si scrive:

$$\frac{2}{h^3} \frac{dp_x dp_y dp_z}{e^{(\varepsilon - \mu)/k_B T} + 1} \quad (5.7)$$

con $\varepsilon = (p_x^2 + p_y^2 + p_z^2)/2m$. La carica elettrica emessa dal metallo, per unità di superficie e unità di tempo è quindi¹

$$j_x = \frac{2}{h^3} \int_{p_0}^{\infty} dp_x \frac{p_x}{m} \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z \frac{1}{e^{(\varepsilon - \mu)/k_B T} + 1}. \quad (5.8)$$

Introducendo le variabili $\varepsilon_x = p_x^2/2m$ e $\varepsilon_t = (p_y^2 + p_z^2)/2m$ abbiamo

$$\begin{aligned} j_x &= \frac{4\pi m e}{h^3} \int_{e\phi}^{\infty} d\varepsilon_x \int_0^{\infty} d\varepsilon_t \frac{1}{e^{(\varepsilon_x + \varepsilon_t - \mu)/k_B T} + 1} = \\ &= \frac{4\pi m e k_B T}{h^3} \int_{e\phi}^{\infty} d\varepsilon_x \ln \left[1 + e^{-(\varepsilon_x - \mu)/k_B T} \right]. \end{aligned} \quad (5.9)$$

¹ Si noti che stiamo assumendo che un elettrone con energia sufficiente oltrepassi sicuramente la barriera di potenziale, cioè stiamo completamente trascurando l'effetto quantistico della riflessione. Questo effetto, che causa una correzione da considerare nel confronto coi dati sperimentali, non è importante per il nostro scopo.

Nell'esponenziale qua sopra si ha $(\epsilon_x - \mu) > (e\phi - \mu) > (e\phi - \epsilon_F)$, dove ϵ_F è l'energia di fermi del metallo. Tipicamente la differenza $e\phi - \epsilon_F$ è qualche elettronvolt; dato che $1 \text{ eV} \approx 12.000 \text{ K}$ e inoltre la maggior parte dei comuni metalli a 2.000 K sono già fusi (il tungsteno fonde a $\approx 3.700 \text{ K}$), per le temperature che ha senso considerare risulta $(e\phi - \mu) \gg k_B T$, si può quindi sviluppare il logaritmo in termini di $e^{-(\epsilon_x - \mu)/k_B T} \ll 1$, ottenendo

$$j_x = \frac{4\pi m e k_B T}{h^3} \int_{e\phi}^{\infty} d\epsilon_x e^{-(\epsilon_x - \mu)/k_B T}, \quad (5.10)$$

e quindi la legge di Richardson

$$j_x = \frac{4\pi e m k_B^2}{h^3} T^2 e^{-W/k_B T} \quad (5.11)$$

dove $W = e\phi - \mu$ è il lavoro di estrazione, cioè la minima quantità di energia necessaria affinché un elettrone lasci il metallo.

La trattazione del problema con la meccanica statistica classica è molto semplice; usando la distribuzione di Maxwell per le velocità, la corrente per unità di superficie si scrive:

$$j_x = \rho e \left(\frac{m}{2\pi k_B T} \right)^{3/2} \int_{v_0}^{\infty} v_x e^{-\frac{mv_x^2}{2k_B T}} dv_x \int_{-\infty}^{\infty} e^{-\frac{mv_y^2}{2k_B T}} dv_y \int_{-\infty}^{\infty} e^{-\frac{mv_z^2}{2k_B T}} dv_z, \quad (5.12)$$

essendo $\rho = N/V$ la densità di particelle. L'integrale è elementare e si ottiene la legge di Richardson nella forma

$$j_x = \frac{\rho e \bar{v}}{4} e^{-e\phi/k_B T} \quad (5.13)$$

dove $\bar{v} = \sqrt{8k_B T / \pi m}$.

Possiamo notare che la (5.13) si riottiene se nella (5.11) si usa la relazione che lega il potenziale chimico alla densità e alla temperatura, nel limite di gas classico:

$$e^{\beta\mu} = \frac{\rho h^3}{2 (2\pi m k_B T)^{3/2}}, \quad (5.14)$$

in cui il fattore 2 al denominatore tiene conto dello spin. L'equazione (5.14) è vera per $k_B T \gg \epsilon_F$, per cui, essendo le ϵ_F dell'ordine dell'eV, la legge (5.13) non sembra adatta a descrivere l'effetto termoionico.

Nella condizione di degenerazione $k_B T < \epsilon_F$, realizzata negli esperimenti, con buona approssimazione si può porre $\mu \approx \epsilon_F$ e quindi $W = e\phi - \epsilon_F$.

Gli andamenti in funzione della temperatura previsti nelle (5.11) e (5.13) differiscono principalmente nel coefficiente del fattore $1/k_B T$ che compare nell'esponenziale: la "correzione" dovuta alle potenze $T^{1/2}$ o T^2 influisce poco nella rapidità di variazione della corrente. Il confronto con i dati serve quindi a determinare se l'aumento esponenziale della corrente termoionica all'aumentare della temperatu-

ra è guidato da $e\phi$ (regime classico) o da $e\phi - \varepsilon_F$ (regime quantistico). Questo è possibile perché si può determinare l'altezza della barriera di potenziale ($e\phi$) in maniera indipendente, per esempio studiando la diffrazione degli elettroni sul metallo. Come ci si aspetta, il risultato è che i dati sono ben descritti dalla legge (5.11) con $W \approx e\phi - \varepsilon_F \neq e\phi$; ad esempio, per il tungsteno si ha $e\phi \approx 13.5$ eV, $\varepsilon_F \approx 9$ eV e $W \approx 4.5$ eV.

5.1.3 Il problema del potenziale di interazione

Il problema tipico della meccanica statistica viene presentato nel seguente modo: in un recipiente di volume V sono contenute N particelle di massa m che interagiscono con un potenziale a coppie dipendente dalla distanza $U(r)$; il sistema è descritto dall'Hamiltoniana

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \sum_{i,j} U(|\mathbf{q}_i - \mathbf{q}_j|); \quad (5.15)$$

sapendo che il sistema è in equilibrio termico a temperatura T determinare l'equazione di stato, l'energia interna etc.

Conoscendo $U(r)$, almeno nel caso di gas abbastanza diluiti, possiamo trovare un'equazione di stato approssimata

$$\frac{P}{k_B T} = \rho + b_2(T)\rho^2 + b_3(T)\rho^3 + \dots \quad (5.16)$$

dove $\rho = N/V$ e i coefficienti del viriale $b_2(T), b_3(T), \dots$ sono calcolabili esplicitamente a partire da $U(r)$; ad esempio (vedi la sezione Appendice F nel capitolo delle Appendici)

$$b_2(T) = 2\pi \int_0^\infty r^2 (1 - e^{-\beta U(r)}) dr. \quad (5.17)$$

Nella realtà non è quasi mai possibile affrontare il problema come fin qui presentato (che sostanzialmente è la versione da esercizio di esame). Infatti, anche nei casi a temperatura abbastanza alta in cui sia possibile usare la meccanica statistica classica, $U(r)$ non è conosciuto. Notiamo che il potenziale di interazione ha sempre un'origine quantistica e dovrebbe (in linea di principio) essere calcolato a partire dall'equazione di Schrödinger. Solo in casi particolari (liquidi semplici con molecole sferiche) è possibile determinare (con buona accuratezza) $U(r)$ dai principi primi.

Nella realtà in genere si procede nel seguente modo: si ipotizza una forma specifica per $U(r)$ che contiene dei parametri da determinare, ad esempio per i liquidi semplici è molto usato il potenziale di Lennard-Jones, vedi Fig. 5.2,

$$U(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (5.18)$$

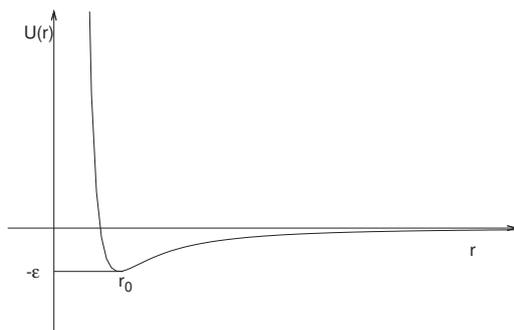


Figura 5.2 Potenziale di Lennard-Jones (5.18). Il parametro σ determina la distanza di equilibrio: $r_0 = 2^{1/6} \sigma$, e ε la profondità della buca: $U(r_0) = -\varepsilon$.

la coda r^{-6} ha un'origine chiara (le interazioni dipolo-dipolo), la parte repulsiva r^{-12} è puramente fenomenologica.

Partendo dall'equazione di Schrödinger e usando l'approssimazione di Born-Oppenheimer o metodi simili (come la tecnica di Car-Parrinello), in alcuni casi particolari, è possibile determinare il valore dei parametri σ e ε . In sistemi più complessi si utilizzano potenziali quasi completamente empirici, ad esempio per i sali fusi si usa la forma proposta da Tosi e Fumi

$$U(r) = Ae^{-r/\lambda} - \frac{B}{r^6} - \frac{C}{r^8} \quad (5.19)$$

dove i parametri A, B, C e λ (come anche ε e σ nel caso del potenziale di Lennard-Jones) sono determinati dalle equazioni di stato ricavate sperimentalmente dalle relazioni tra P e ρ a varie temperature.

La strategia tipicamente usata per determinare il potenziale di interazione è la seguente:

- a) si assume una forma opportuna per $U(r)$, ad esempio tipo Lennard-Jones e Tosi-Fumi;
- b) a partire dal potenziale $U(r)$ vengono calcolati i coefficienti del viriale $b_2(T), \dots$, che dipendono funzionalmente dai parametri;
- c) i parametri del potenziale (A, B, C e λ , oppure ε e σ) sono determinati dal confronto dei coefficienti del viriale calcolati con i dati sperimentali delle equazioni di stato.

5.2 Contare gli atomi senza poterli vedere: il moto Browniano

5.2.1 Un po' di storia

Nell'estate del 1827 il botanico scozzese R. Brown osservò per la prima volta, con l'aiuto del microscopio, il fenomeno ora noto col nome di moto Browniano (MB):

grani di polline in sospensione in acqua si muovono continuamente in modo rapido e irregolare, vedi Fig. 5.3.

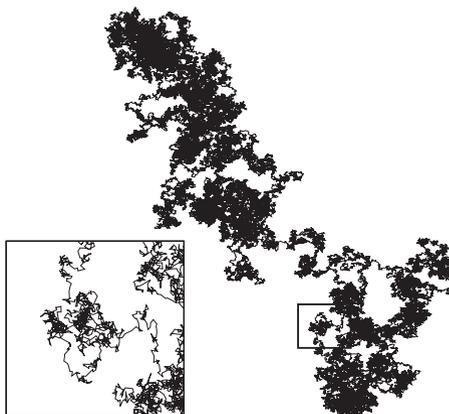
Il lavoro, dal titolo chilometrico: *A brief account of microscopical observations made in the month of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies*, inizialmente venne frainteso e Brown dovette insistere non poco per sottolineare l'origine non organica del fenomeno. Per decenni l'importanza del MB venne sistematicamente sottovalutata; tra le poche eccezioni possiamo citare M. Faraday (il fisico sperimentale più autorevole dell'epoca) che ripeté l'esperimento confermando il risultato. Alcuni tuttavia intuirono l'importanza del fenomeno, ad esempio verso la metà dell'ottocento il fisico padovano Giovanni Cantoni notava che *il moto Browniano ci fornisce una delle più belle e dirette dimostrazioni sperimentali dei fondamentali principi della teoria meccanica del calore, manifestando quell'assiduo stato vibratorio che esser deve e nei liquidi e nei solidi ancor quando non si muta in essi la temperatura*.

Verso la fine del XIX secolo L.G. Gouy effettuò esperimenti sistematici variando la temperatura, la viscosità del liquido e la grandezza del grano, suggerendo che lo studio del MB potesse essere un modo per esaminare la teoria cinetica. H. Poincaré, che conosceva i risultati di Gouy, nel suo influente libro *Il valore della scienza* attirò l'attenzione sull'importanza del MB.

All'inizio del XX secolo il moto Browniano venne finalmente riconsiderato e si capì che, ben lungi dall'essere un fenomeno marginale, costituisce il punto di svolta per la comprensione di un problema antico e nobile: l'esistenza degli atomi.

Il 1905 fu l'*annus mirabilis* di Albert Einstein che, in rapida successione, scrisse una serie di articoli che hanno rivoluzionato la fisica introducendo la relatività ristretta e fornendo spiegazioni per l'effetto fotoelettrico e per il moto Browniano. I primi due lavori sono ben noti, il terzo, pur essendo meno noto degli altri, non è certamente meno importante o innovativo. Einstein non era particolarmente inte-

Figura 5.3 Moto Browniano in un piano; nel quadrato è mostrato l'ingrandimento di una parte della traiettoria.



ressato al fenomeno in sé, ma piuttosto voleva mostrare che *in accordo con la teoria molecolare del calore, corpi di dimensioni visibili al microscopio sospesi in un liquido sono dotati di movimenti di tale ampiezza da poter essere facilmente osservati. È possibile che il fenomeno che discuto sia simile al così detto moto Browniano.* In altre parole voleva trovare un'evidenza sperimentale conclusiva dell'esistenza degli atomi.

5.2.2 La teoria di Einstein nella versione di Langevin

Einstein ragionò nel seguente modo: il granello di polline è molto più grande delle molecole del fluido, in qualche modo deve comportarsi come un corpo macroscopico e quindi quando si muove sente una forza d'attrito; inoltre a causa degli urti con le molecole dopo un certo tempo il granello avrà la stessa energia cinetica media di ogni molecola². La seconda assunzione è veramente ardita, tanto per dare un'idea delle dimensioni in gioco è come supporre che una mongolfiera su cui sbattono moscerini alla fine avrà la stessa energia cinetica media di un moscerino. È importante sottolineare il coraggio di Einstein dato che le due ipotesi sono apparentemente contraddittorie: nella prima si assume che il granello si comporti come un corpo macroscopico, nella seconda si richiede l'equipartizione dell'energia, una tipica proprietà della descrizione microscopica.

L'approccio seguito da Einstein sarà discusso nella sezione Appendice D del capitolo delle Appendici mentre qui seguiremo l'impostazione di Langevin. L'idea è che il moto di un grano di polline è determinato da due contributi:

- a) la forza "media" d'attrito, dovuta alla presenza del fluido;
- b) gli urti con le molecole.

Considerando solo il moto nella direzione x , per semplicità di notazione, si ha quindi:

$$m \frac{dv}{dt} = F_a + F_f . \quad (5.20)$$

Il contributo dell'attrito è dato dalla legge di Stokes $F_a = -6\pi\eta Rv$ dove η è la viscosità del fluido, R il raggio del grano (assunto sferico) di massa m e v la sua velocità. Se si trascura F_f (cioè il contributo degli urti con le molecole) un grano di massa m ha un tempo caratteristico di rilassamento $\tau = m/(6\pi\eta R)$. A temperatura ambiente, per un grano di raggio un micron sospeso in acqua, $\tau = O(10^{-7})$ s, un tempo piccolo ma molto più grande dei tempi di collisione molecolari che sono $O(10^{-11})$ s. Il grande rapporto fra questi due tempi caratteristici e l'elevato numero di molecole che urtano nell'unità di tempo con il grano permettono di modellizzare F_f come un rumore bianco indipendente dalla posizione, cioè un processo stocastico

² Ricordiamo che nella meccanica statistica classica, se T è la temperatura di un sistema in equilibrio, la distribuzione di Maxwell prevede che per le molecole l'energia cinetica media valga $(3/2)k_B T$, indipendentemente dalla massa (equipartizione dell'energia).

gaussiano caratterizzato da media nulla e correlazione temporale a delta:

$$\langle F_f \rangle = 0 \quad \langle F_f(t)F_f(t') \rangle = c \delta(t - t') ; \quad (5.21)$$

la costante c è determinata richiedendo l'equipartizione dell'energia $\langle v^2 \rangle = k_B T / m$, con T temperatura del fluido.

A questo punto è facile ricavare la relazione di Einstein-Smoluchowski³. Ricordando che

$$v = \frac{dx}{dt} \quad \text{e} \quad \frac{d^2 x^2}{dt^2} = 2v^2 + 2x \frac{dv}{dt} , \quad (5.22)$$

moltiplicando ambo i membri dell'eq. (5.20) per x e mediando si ottiene

$$\frac{m}{2} \frac{d^2}{dt^2} \langle x^2 \rangle - m \langle v^2 \rangle = -3\pi\eta R \frac{d}{dt} \langle x^2 \rangle + \langle xF_f \rangle ; \quad (5.23)$$

introducendo la variabile $z = d\langle x^2 \rangle / dt$, usando l'equipartizione dell'energia e tenendo presente che $\langle xF_f \rangle = 0$, si ha

$$\frac{m}{2} \frac{dz}{dt} + 3\pi\eta R z = k_B T . \quad (5.24)$$

Risolvendo l'equazione precedente otteniamo

$$z(t) = \frac{k_B T}{3\pi\eta R} + A e^{-t/\tau} , \quad (5.25)$$

dove il valore della costante A , che dipende dalle condizioni iniziali, è irrilevante per il comportamento a tempi lunghi. Di conseguenza, possiamo scrivere

$$\lim_{t \rightarrow \infty} \frac{1}{2t} \langle [x(t) - x(0)]^2 \rangle = D = \frac{k_B T}{6\pi\eta R} . \quad (5.26)$$

Inoltre Einstein ha mostrato⁴ che su scale di tempo lunghe rispetto a τ la densità di probabilità per la posizione della particella Browniana, $p(x, t)$, evolve secondo l'equazione di diffusione

$$\frac{\partial}{\partial t} p(x, t) = D \frac{\partial^2}{\partial x^2} p(x, t) ; \quad (5.27)$$

per condizioni iniziali molto localizzate $p(x, 0) = \delta(x - x(0))$ è facile verificare che la soluzione è

$$p(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-(x-x(0))^2/4Dt} . \quad (5.28)$$

³ Marian Smoluchowski studiò il MB sia sperimentalmente che teoricamente con un approccio basato sulla teoria cinetica.

⁴ La sua derivazione è sostanzialmente quella che oggi viene seguita per ricavare l'equazione di Fokker-Planck, vedi l'Appendice D.

La relazione (5.26) mostra quindi come, studiando le proprietà di diffusione di particelle Browniane, si è in grado di determinare la costante di Boltzmann e di conseguenza il numero di Avogadro attraverso la costante dei gas.

5.2.3 L'equazione di Langevin a tempi discreti

A qualche lettore la derivazione della (5.26) può sembrare un po' disinvolta. In effetti i calcoli, che seguono fedelmente l'articolo di Langevin, sono semplici ma formali (nel senso negativo del termine) anche se il risultato finale è corretto. Una trattazione più rigorosa richiede l'uso dei processi stocastici a tempi continui. Comunque una versione a tempi discreti dell'eq. (5.20) è sufficiente per ricavare in modo chiaro il risultato fondamentale. Indichiamo con v_n la velocità della particella browniana al tempo $n\Delta t$ dove Δt è piccolo (poi sarà preso il limite $\Delta t \rightarrow 0$), l'equazione di Langevin si può scrivere nella forma⁵

$$v_{n+1} = av_n + bw_n \quad (5.29)$$

dove w_n è una variabile aleatoria ed a e b sono costanti da scegliere in modo opportuno, per ora poniamo $0 < a < 1$ (chiaramente $a \simeq 1 - \Delta t/\tau$). Per semplicità di calcolo assumiamo che le $\{w_n\}$ siano variabili indipendenti con distribuzione $N(0, 1)$, ovvero gaussiana a media nulla e varianza 1; assumiamo inoltre che la distribuzione della velocità iniziale, v_0 , sia $N(\langle v_0 \rangle, \sigma_0^2)$ e indipendente da $\{w_n\}$. Ricordando che la combinazione lineare di variabili gaussiane è ancora gaussiana, le ipotesi fatte e la (5.29) implicano che anche v_n è una variabile gaussiana; è quindi sufficiente determinare $\langle v_n \rangle$ e σ_n^2 . Con calcoli semplici si ottengono le relazioni di ricorrenza:

$$\langle v_{n+1} \rangle = a \langle v_n \rangle \quad \sigma_{n+1}^2 = a^2 \sigma_n^2 + b^2 \quad (5.30)$$

che possono essere risolte senza difficoltà

$$\langle v_n \rangle = a^n \langle v_0 \rangle \quad \sigma_n^2 = \sigma_\infty^2 + a^{2n}(\sigma_0^2 - \sigma_\infty^2), \quad (5.31)$$

con $\sigma_\infty^2 = \langle v^2 \rangle = b^2/(1 - a^2)$.

Quindi la variabile v_n ha distribuzione $N(\langle v_n \rangle, \sigma_n^2)$, dove $\langle v_n \rangle$ e σ_n^2 tendono al loro valore asintotico con un "tempo caratteristico" (in unità Δt) $n_c = 1/|\ln a|$. Notare che se vogliamo che $\langle v^2 \rangle$ sia una quantità finita allora si deve richiedere $b \propto \sqrt{\Delta t}$ e subito si capisce che la variabile $bw_n/\Delta t$, che nel limite $\Delta t \rightarrow 0$ diventa F_f/m , è un oggetto un po' piuttosto "selvaggio".

Consideriamo ora l'evoluzione della posizione: $x_{n+1} = x_n + v_n \Delta t$ e calcoliamo il coefficiente di diffusione D . Poiché siamo interessati al comportamento a tempi lunghi, e il rilassamento della distribuzione di probabilità delle velocità è rapido, possiamo assumere che le velocità siano distribuite secondo $N(0, \langle v^2 \rangle)$. Scrivendo

⁵ W. Pauli nel suo libro *Statistical Mechanics* (Dover Ed. 2000) discute un modello a tempo discreto per il moto Browniano, attribuito a H.A. Lorentz, praticamente uguale a (5.29).

lo spostamento della particella browniana nella forma

$$x_n - x_0 = \Delta t \sum_{j=0}^{n-1} v_j \quad (5.32)$$

si ottiene

$$\langle [x_n - x_0]^2 \rangle = \Delta t^2 \{ n \langle v^2 \rangle + 2 \sum_{j=0}^{n-1} \sum_{k=1}^{n-j-1} \langle v_j v_{j+k} \rangle \} \quad (5.33)$$

un semplice calcolo mostra che $\langle v_j v_{j+1} \rangle = a \langle v^2 \rangle$ e più in generale $\langle v_j v_{j+k} \rangle = a^k \langle v^2 \rangle$, quindi la sommatorie nella (5.33) sono riconducibili a somme di serie geometriche. A questo punto è facile calcolare il coefficiente di diffusione

$$D = \lim_{t \rightarrow \infty} \frac{1}{2n\Delta t} \langle [x_n - x_0]^2 \rangle \quad \text{con } t = n\Delta t; \quad (5.34)$$

ricordando che $a \simeq 1 - \Delta t/\tau$ e usando la (5.33) si ottiene la (5.26):

$$D = \tau \langle v^2 \rangle = \frac{k_B T}{6\pi\eta R}. \quad (5.35)$$

5.2.4 Gli esperimenti di Perrin

Jean Perrin, un raffinato fisico sperimentale, ha avuto il grande merito di verificare in laboratorio, in modo estremamente accurato, le previsioni teoriche di Einstein. Le difficoltà pratiche che dovette affrontare furono enormi, per noi quasi inimmaginabili: all'epoca in un laboratorio bisognava fare tutto da soli a partire dalla costruzione di sferette con massa e raggio desiderati. Per i suoi esperimenti Perrin utilizzò emulsioni di gommagutta trattata con alcool e centrifugata per ottenere granuli di diametro noto. Osservando con un microscopio grani di diametro molto diverso (da 0.2 a 5.5 μm) Perrin⁶, dopo una lunga serie di esperimenti, riuscì a verificare la relazione (5.26) e nel 1926 gli venne conferito il premio Nobel.

5.2.4.1 La formula barotropica

Una volta accettata l'idea che granelli visibili seguano le stesse leggi statistiche delle molecole, ci si aspetta che emulsioni di granelli seguano le leggi dei gas. Perrin ha verificato sperimentalmente la validità della formula barotropica immergendo dei granelli di gommagutta di dimensioni dell'ordine del micron e di massa M (dell'ordine di $10^{-12} - 10^{-11}$ g) in acqua a temperatura T e osservando al microscopio la variazione della densità di granelli in funzione dell'altezza. Se la densità dei gra-

⁶ Perrin confrontando i suoi risultati sperimentali con la teoria trovò un'impresione numerica che venne prontamente corretta da Einstein.

nelli è bassa, e quindi le particelle browniane sono praticamente non interagenti, si possono confrontare i risultati con la formula barotropica dei gas perfetti in campo esterno:

$$\rho(z) = \rho(0) \exp \left[-\frac{Mgz}{k_B T} \right]. \quad (5.36)$$

Anche in questo caso i risultati di Perrin confermano l'idea di Einstein sulla validità della meccanica statistica anche per corpi macroscopici e costituiscono un'altra misura del numero di Avogadro.

5.2.5 L'importanza storica e concettuale del moto Browniano

La rilevanza della relazione di Einstein-Smoluchowski è nella possibilità di collegare la grandezza D , che è determinata da quantità macroscopiche osservabili sperimentalmente ($\langle [x(t) - x(0)]^2 \rangle$, η , R , T etc.), con il numero di Avogadro N_A (ricordiamo che $k_B = \mathcal{R}/N_A$ ove \mathcal{R} è la costante dei gas). Dalla (5.26) abbiamo una connessione non ambigua tra il mondo microscopico (le molecole) e quello macroscopico (il granello) ed inoltre un modo per *contare le molecole*, cioè determinare N_A da una misura di D . Possiamo dire che l'osservazione del moto irregolare di una particella browniana abbastanza grande da essere visibile al microscopio (diciamo del diametro di un micron) fornisce una lente di ingrandimento sul mondo atomico.

Gli esperimenti condotti da J. Perrin chiusero definitivamente la polemica sulla reale esistenza degli atomi tra Boltzmann da una parte e la scuola energetista (capeggiata da Ostwald e Mach) dall'altra. Questo venne "ufficializzato" da Arrhenius che nel 1911 in una conferenza a Parigi, riassumendo i lavori di Einstein, Smoluchowski e Perrin sul MB, dichiarò: *dopo questo, non sembra più possibile dubitare che la teoria molecolare, introdotta dai filosofi dell'antichità, Leucippo e Democrito, descrive la realtà almeno nei suoi aspetti essenziali.*

Dov'è il punto geniale? Nello studio del grano di polline è stato ipotizzato che per la forza ci siano due contributi: uno di tipo macroscopico, dato dalla legge di Stokes, e uno di origine microscopica originato dagli urti con le molecole. Al granello, che ha una massa molto maggiore di quella delle molecole del fluido, è chiesto qualcosa di apparentemente contraddittorio: comportarsi in modo tale da essere contemporaneamente in accordo con la teoria cinetica (l'equipartizione dell'energia) e con l'idrodinamica (legge di Stokes). Questa è l'idea ardita e geniale di Einstein.

È interessante il fatto che lo scopo di Einstein non fosse quello di spiegare il MB del quale (a differenza di Smoluchowski) aveva una conoscenza superficiale; il suo scopo era la ricerca di una manifestazione chiara dell'esistenza degli atomi che potesse essere osservata sperimentalmente.

5.3 Le fluttuazioni sono piccole ma rivelabili

Nei sistemi macroscopici le fluttuazioni sono percentualmente piccole, come è ben evidente dalla nota formula per la varianza dell'energia nell'insieme canonico che

per un sistema di N particelle a temperatura T si scrive

$$\langle (E - \langle E \rangle)^2 \rangle = N c_V k_B T^2, \quad (5.37)$$

dove c_V è il calore specifico per particella calcolato a volume costante. Da questa relazione segue facilmente che le fluttuazioni relative dell'energia sono $O(N^{-1/2})$, quindi estremamente piccole, praticamente trascurabili e, almeno apparentemente, non rivelabili. Tuttavia abbiamo visto nel caso del moto Browniano come le fluttuazioni pur piccole possano essere osservate, ed hanno avuto un ruolo fondamentale per l'affermazione dell'ipotesi atomistica.

Mostriamo come la (5.37) possa essere generalizzata al caso di altre quantità (oltre l'energia) di interesse termodinamico e come le fluttuazioni, benchè percentualmente piccole, siano (almeno indirettamente) misurabili e costituiscano un aspetto importante della meccanica statistica.

Assumiamo che lo stato macroscopico sia descritto da n variabili (che includono anche l'energia) $\alpha_1, \dots, \alpha_n$ che sono funzioni dello stato microscopico \mathbf{X} : $\alpha_j = g_j(\mathbf{X})$ con $j = 1, \dots, n$. Indichiamo con \mathcal{P} i parametri che determinano la densità di probabilità dello stato \mathbf{X} , ad esempio nell'insieme canonico $\mathcal{P} = (T, N, V)$, nel grand canonico $\mathcal{P} = (T, \mu, V)$, nell'insieme isobarico a temperatura fissata $\mathcal{P} = (T, N, P)$. La densità di probabilità delle $\{\alpha_j\}$ è

$$p(\alpha_1, \dots, \alpha_n) = \int \rho(\mathbf{X}, \mathcal{P}) \prod_{j=1}^n \delta(\alpha_j - g_j(\mathbf{X})) d\mathbf{X}$$

dove $\rho(\mathbf{X}, \mathcal{P})$ è la densità di probabilità dello stato \mathbf{X} nell'insieme caratterizzato dai parametri termodinamici \mathcal{P} . È immediato ottenere

$$p(\alpha_1, \dots, \alpha_n) = \exp -\beta \left[F(\alpha_1, \dots, \alpha_n | \mathcal{P}) - F(\mathcal{P}) \right]$$

dove $F(\mathcal{P})$ è l'“energia libera” del sistema con parametri \mathcal{P} mentre $F(\alpha_1, \dots, \alpha_n | \mathcal{P})$ è l'energia libera del sistema in cui sono fissati i parametri \mathcal{P} e le variabili macroscopiche $\alpha_1, \dots, \alpha_n$.

Ad esempio nell'insieme canonico ove $\mathcal{P} = (T, N, V)$ si ha

$$F(\alpha_1, \dots, \alpha_n | \mathcal{P}) = -k_B T \ln \int \prod_{j=1}^n \delta(\alpha_j - g_j(\mathbf{X})) e^{-\beta H(\mathbf{X})} d\mathbf{X}.$$

Possiamo identificare la quantità

$$\delta S(\alpha_1, \dots, \alpha_n) = -\frac{1}{T} \left[F(\alpha_1, \dots, \alpha_n | \mathcal{P}) - F(\mathcal{P}) \right]$$

con la differenza di entropia tra lo stato con parametri $(\alpha_1, \dots, \alpha_n, \mathcal{P})$ e quello con \mathcal{P} . Abbiamo quindi

$$p(\alpha_1, \dots, \alpha_n) = \exp \frac{\delta S(\alpha_1, \dots, \alpha_n)}{k_B} \quad (5.38)$$

questa relazione viene chiamata principio di Boltzmann-Einstein. L'uso del termine entropia per δS è sicuramente corretto nel caso microcanonico; più in generale δS è la differenza di un opportuno potenziale termodinamico.

Nei sistemi macroscopici ci si aspetta che le fluttuazioni rispetto allo stato di equilibrio termodinamico siano (per le variabili estensive) percentualmente piccole. È quindi giustificato sviluppare la $\delta S(\alpha_1, \dots, \alpha_n)$ in serie di Taylor intorno ai valori medi di $\{\alpha_j\}$, che sono proprio i valori all'equilibrio termodinamico $\{\alpha_j^*\}$:

$$\delta S(\alpha_1, \dots, \alpha_n) \simeq -\frac{1}{2} \sum_{i,j} \delta \alpha_i A_{ij} \delta \alpha_j$$

dove

$$\delta \alpha_i = \alpha_j - \alpha_j^*, \quad A_{ij} = -\left. \frac{\partial^2 S}{\partial \alpha_i \partial \alpha_j} \right|_{\alpha^*}.$$

Le piccole fluttuazioni intorno ai valori di equilibrio termodinamico sono descritte da una distribuzione gaussiana multivariata:

$$p(\alpha_1, \dots, \alpha_n) \simeq \sqrt{\frac{\det \mathbf{A}}{(2\pi k)^n}} \exp -\frac{1}{2k} \sum_{i,j} \delta \alpha_i A_{ij} \delta \alpha_j$$

da cui

$$\langle \delta \alpha_i \delta \alpha_j \rangle = k_B [\mathbf{A}^{-1}]_{ij}. \quad (5.39)$$

Le relazioni precedenti sono la base della cosiddetta teoria delle fluttuazioni di Einstein.

Notare che gli elementi di matrice A_{ij} sono calcolati all'equilibrio termodinamico. La matrice \mathbf{A} deve essere definita positiva (cioè con autovalori strettamente positivi); questa proprietà ha un preciso significato fisico: la differenza di entropia (rispetto all'equilibrio termodinamico) deve essere negativa. La (5.37) è un caso particolare della (5.39); un altro esempio importante sono le fluttuazioni della magnetizzazione per particella:

$$\langle (\delta m)^2 \rangle = \frac{k_B}{N} T \chi \quad (5.40)$$

dove χ è la suscettività magnetica per particella.

È interessante il fatto che la varianza della fluttuazione di una data quantità è sempre proporzionale a un'opportuna funzione di risposta. Nella (5.37) appare il calore specifico $c_V = (\partial \langle E \rangle / \partial T) / N$ che misura come cambia l'energia media per particella al variare della temperatura, nella (5.40) interviene la suscettività $\chi = \partial \langle m \rangle / \partial B$ che determina come cambia la magnetizzazione media con il campo magnetico.

Le equazioni (5.37) e (5.40) sono casi particolari di un teorema più generale noto come teorema di fluttuazione-risposta o fluttuazione-dissipazione. Dalle (5.37) e (5.40) si può notare che le fluttuazioni percentuali sono piccole per $N \gg 1$. Questo potrebbe far pensare che nei sistemi macroscopici le fluttuazioni siano irrilevanti e

non misurabili. Al contrario sono concettualmente molto importanti: come notato da Einstein: *se fosse possibile misurare le fluttuazioni di energia del sistema, si avrebbe una determinazione esatta della costante fondamentale k_B e quindi del numero di Avogadro*⁷. Il teorema fluttuazione-risposta, di cui (5.37) e (5.40) sono casi particolari, implica che le fluttuazioni sono (indirettamente) misurabili in quanto connesse a funzioni di risposta (come il calore specifico e la suscettività).

5.4 Le dimensioni delle molecole

Vogliamo mostrare ora come, una volta fatta l'assunzione che un gas sia costituito da atomi (molecole), anche senza grandi sforzi sia possibile determinare le loro "dimensioni" e le energie tipiche in gioco, usando un po' di meccanica statistica e i dati termodinamici sperimentali,

5.4.1 Qualche calcolo semplice

Consideriamo un gas molto diluito, con densità ρ , a temperatura T ; per semplicità possiamo assumere che le interazioni siano solo quelle dovute all'impenetrabilità tra sfere dure di massa m e raggio R . Utilizzando la distribuzione di probabilità di Maxwell e l'equazione di Boltzmann discussa nel Capitolo 4, è possibile calcolare in modo accurato il cammino libero medio e la viscosità. Per i nostri scopi saranno sufficienti semplici considerazioni fisiche.

Cominciamo con il cammino libero medio ℓ , cioè la distanza media percorsa da un atomo tra due urti. Immaginiamo che un solo atomo si muova e che tutti gli altri siano fermi e distribuiti in modo uniforme. Possiamo stimare il cammino libero medio notando che in un cilindro di raggio $2R$ e altezza L ci sono, in media $N^* = 4\pi R^2 L \rho$ particelle, quindi un atomo ne incontra un altro dopo aver percorso una distanza media $\ell = L/N^*$:

$$\ell \sim \frac{1}{\rho R^2}. \quad (5.41)$$

Ricordando l'equipartizione dell'energia, abbiamo che, in un gas a temperatura T , la velocità tipica di una particella vale $v_t \sim \sqrt{k_B T/m}$ e, per il tempo medio di collisione τ^* , si ha

$$\tau^* \sim \frac{\ell}{v_t} \sim \frac{1}{\rho R^2 \sqrt{k_B T/m}}. \quad (5.42)$$

A questo punto è facile ottenere una stima del coefficiente di diffusione D della singola molecola. Consideriamo la sua velocità $v(t)$ in una direzione, $v(t)$ cambia solo a causa delle collisioni e può essere vista come un processo stocastico con un tempo caratteristico dell'ordine del tempo medio di collisione τ^* , quindi (per i

⁷ Questo aspetto è stato discusso in dettaglio per il moto Browniano.

dettagli vedi la sezione Appendice D delle Appendici)

$$D \sim v_t^2 \tau^* \sim \frac{\sqrt{k_B T/m}}{\rho R^2} . \quad (5.43)$$

Si noti che questa espressione è diversa dalla (5.26) trovata da Einstein per un granello in un liquido, infatti stiamo trattando la diffusione in un gas.

Ora siamo pronti per la stima della viscosità η che appare nell'equazione di Navier-Stokes per l'evoluzione del campo di velocità $\mathbf{u}(\mathbf{x}, t)$, che descrive a livello macroscopico il gas:

$$\rho_m \frac{\partial}{\partial t} \mathbf{u}(\mathbf{x}, t) + \dots = \eta \Delta \mathbf{u}(\mathbf{x}, t) \quad (5.44)$$

dove $\rho_m = m\rho$ è la densità di materia e i puntini indicano termini non lineari dell'equazione che nel limite di piccoli numeri di Reynolds⁸ sono trascurabili.

Abbiamo ricordato che Einstein ha mostrato che il coefficiente di diffusione D appare nell'equazione che regola l'evoluzione temporale della concentrazione di particelle:

$$\frac{\partial}{\partial t} C(\mathbf{x}, t) = D \Delta C(\mathbf{x}, t) ; \quad (5.45)$$

a questo punto possiamo notare che le equazioni (5.44) e (5.45) hanno la stessa struttura matematica e la stessa origine fisica (le collisioni tra le molecole), è naturale concludere che

$$\frac{\eta}{m\rho} = \text{cost.} D , \quad (5.46)$$

e quindi

$$\eta \sim m\rho D \sim \frac{\sqrt{mk_B T}}{R^2} . \quad (5.47)$$

Abbiamo quindi, cosa non intuitiva, che la viscosità è indipendente dalla densità.

Un'analisi accurata a partire dall'equazione di Boltzmann mostra che le stime precedenti sono corrette, a parte fattori numerici $O(1)$:

$$\begin{aligned} \ell &= \frac{1}{4\sqrt{2}\pi\rho R^2} , & \tau^* &= \frac{1}{4\sqrt{2}\pi\rho R^2 \bar{v}} , \\ D &= \frac{1}{3} \bar{v}^2 \tau^* , & \eta &= m\rho D . \end{aligned} \quad (5.48)$$

Se invece di un gas di sfere dure si considera il caso più realistico con un'interazione attraverso un potenziale a coppie $U(r)$, le (5.48) sono ancora valide una volta effettuata la sostituzione $4\pi R^2$ con la sezione d'urto totale σ_T (che dipende da $U(r)$).

⁸ Il numero di Reynolds è definito come $Re = UL\rho_m/\eta$ dove U e L sono, rispettivamente, la velocità e la lunghezza tipiche del campo di velocità del fluido a grande scala. Re è piccolo in regime laminare e grande per un fluido turbolento.

5.4.2 Dalla termodinamica ai potenziali di interazione

Lo sviluppo del viriale (5.16) permette la determinazione delle caratteristiche essenziali del potenziale di interazione attraverso il confronto con i dati sperimentali relativi alle equazioni di stato. Consideriamo un semplice potenziale modello $U(r)$, vedi Fig. 5.4, che ha le proprietà fondamentali dei potenziali intermolecolari, cioè è repulsivo a piccola scala, attrattivo a scala intermedia e trascurabile a grandi distanze: $U(r) = \infty$ se $r < r_0$, $U(r) = -\varepsilon$ se $r_0 < r < r_1$ e $U(r) = 0$ per $r > r_1$. Il calcolo del primo coefficiente del viriale è immediato:

$$b_2(T) = 2\pi \int_0^\infty r^2 (1 - e^{-\beta U(r)}) dr = \frac{2}{3} \pi r_1^3 - \frac{2}{3} \pi (r_1^3 - r_0^3) e^{-\beta \varepsilon}. \quad (5.49)$$

Dalle equazioni di stato (pressione in funzione della densità a varie temperature) ottenute in laboratorio possiamo determinare $b_2(T)$ e un confronto con l'espressione precedente permette la stima di r_0 , r_1 e ε .

Per alte temperature $T \gg \varepsilon/k_B$ è possibile un confronto con l'equazione fenomenologica di van der Waals

$$\frac{P + A\rho^2}{k_B T} \left(\frac{1}{\rho} - B \right) = 1. \quad (5.50)$$

Per il coefficiente del viriale si ha

$$b_2(T) \simeq b - a\beta, \quad b = \frac{2}{3} \pi r_0^3, \quad a = \frac{2}{3} \pi (r_1^3 - r_0^3) \varepsilon \quad (5.51)$$

e l'equazione di stato è

$$\frac{P}{k_B T} \simeq \rho + (b - a\beta) \rho^2. \quad (5.52)$$

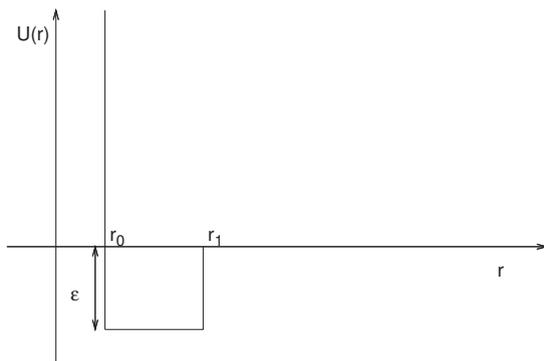


Figura 5.4 Modello semplificato di potenziale intermolecolare: trascurabile per distanze $r > r_1$, attrattivo per distanze $r_0 < r < r_1$, repulsivo a piccole distanze $r < r_0$.

È immediato verificare che ponendo $a = A$ e $b = B$ la (5.50) e (5.52) coincidono fino all'ordine ρ^2 . Abbiamo quindi che l'equazione di van der Waals, originariamente derivata con considerazioni fenomenologiche, può essere ottenuta attraverso lo sviluppo del viriale e i suoi coefficienti A e B hanno un'interpretazione in termini di dimensioni (r_0 ed r_1) ed energia (ϵ) caratteristiche del potenziale di interazione.

È notevole, e confortante, il fatto che i risultati ottenuti con il potenziale modello non cambiano molto se si usa il potenziale Lennard-Jones. Ad alta temperatura ($\epsilon/k_B T \ll 1$) per $r < \sigma$ si ha $f(r) = e^{-\beta U(r)} - 1 \simeq -1$, mentre per $r > \sigma$ si ha $f(r) \simeq -\beta U(r)$ quindi $b_2(t) \simeq (b - \beta a)$ con

$$b = \frac{2}{3}\pi\sigma^3 \quad , \quad a = \frac{16}{9}\epsilon\sigma^3 \quad (5.53)$$

risultati del tutto simili a quelli in (5.51).

5.4.3 Osservazioni conclusive

Nelle pagine precedenti abbiamo visto che diversi risultati teorici combinati con osservazioni sperimentali permettono la stima di quantità microscopiche come il numero di Avogadro N_A (o, equivalentemente, la costante di Boltzmann $k_B = \mathcal{R}/N_A$), la massa m e il raggio R della molecola, l'energia tipica ϵ associata al potenziale di interazione.

Riassumiamo brevemente:

- A) Dal coefficiente di diffusione del moto Browniano nei liquidi abbiamo N_A .
- B) Dalla viscosità nei gas otteniamo \sqrt{m}/R^2 .
- C) Dal secondo coefficiente del viriale abbiamo R^2 ed ϵ .

Anche senza scomodare la moderna tecnologia, con un uso intelligente della meccanica statistica e dei dati di laboratorio, si riesce ad avere un quadro coerente del mondo atomico.

Concludiamo questo capitolo ricordando Democrito che, nel dialogo tra la Ragione ed i Sensi (uno dei pochi frammenti pervenutici), riassume in modo magistrale il dilemma tra l'ipotesi atomistica e la nostra incapacità di sperimentare direttamente gli atomi.

Dice la Ragione:

Solo in apparenza una cosa è dolce o amara, solo in apparenza è calda o fredda, solo in apparenza ha un colore; in realtà esistono solo gli atomi e lo spazio vuoto.

Rispondono i Sensi:

Povero intelletto! Tu che stai prendendo la tua evidenza da noi, stai cercando di spodestarci? La tua vittoria sarà la tua rovina!

Se Democrito avesse conosciuto i risultati della meccanica statistica e del moto Browniano, nel dialogo la Ragione avrebbe potuto (a buon diritto) concludere reclamando il suo trionfo:

Dal caldo e dal freddo dimostro che esistono gli atomi, li posso contare e misurarne le dimensioni.

5.5 Letture consigliate

- L. Peliti *Appunti di Meccanica Statistica* (Bollati Boringhieri, Torino 2009);
- L.E. Reichl *A modern Course in Statistical Physics* (Wiley, 1998);
- G. M. Kremer *An Introduction to the Boltzmann Equation and Transport* (Springer-Verlag, 2010).

I lavori di A. Einstein sul moto browniano sono stati tradotti in inglese e raccolti nel volume

- A. Einstein *Investigations on the Theory of the Brownian Motion* (Dover Publ., 1956).

Per una ricostruzione appassionata delle vicende del moto browniano e la determinazione del numero di Avogadro

- J. Perrin *Les Atomes* (Alcan, Paris 1913), traduzione italiana *Gli Atomi* (Editori Riuniti, Roma 1981).

L'edizione italiana è ormai fuori commercio e quasi introvabile, speriamo che qualche Editore provveda alla ristampa.

Per un'ottima rassegna, in parte anche storica, sul moto browniano con una dettagliata analisi dei lavori di Einstein, Smoluchowski e Perrin si veda

- S. Chandrasekhar *Stochastic problems in physics and astronomy* Rev. Mod. Phys, **15**, 1 (1943).

Appendice A

Osservazioni su esponenziali e potenze

Gli argomenti, essenzialmente qualitativi, esposti nel Capitolo 1 per presentare le idee di base per una interpretazione della Termodinamica dal punto di vista della Meccanica, sono fondati sui peculiari comportamenti dei volumi negli spazi di alta dimensionalità. Abbiamo notato che tali comportamenti, a loro volta, sono conseguenza delle proprietà delle funzioni esponenziali del tipo $\exp(Nx)$ con $N \gg 1$ e $x = O(1)$. In questa appendice vogliamo dare di ciò un'illustrazione quantitativa, mettendo in evidenza la notevole differenza tra la funzione potenza e quella esponenziale.

A.1 Le funzioni

Disponendo di N quantità $\{y_i\}$, dello stesso ordine di grandezza (che prendiamo come scala di riferimento e che assumiamo essere di $O(1)$, ad esempio $y_i \in (0, 10)$), tra tante operazioni che si possono eseguire, due sono particolarmente semplici: possiamo considerarne la somma o il prodotto, ottenendo

$$\sum_i y_i = N \left(\frac{\sum_i y_i}{N} \right) \equiv Nx, \quad (\text{A.1})$$

$$\prod_i y_i = \exp \left(\sum_i \ln y_i \right) = \exp N \left(\frac{\sum_i \ln y_i}{N} \right) \equiv \exp N \bar{x}, \quad (\text{A.2})$$

con x e \bar{x} quantità di ordine 1. Ovviamente la dipendenza da N è a potenza nel caso della somma ed esponenziale nel caso del prodotto. È interessante studiare il comportamento di funzioni di questo tipo, nel caso in cui sia $N \gg 1$ (per fissare le idee, stabiliamo $N = O(10^{20})$).

Consideriamo, come esempio, le seguenti funzioni decrescenti ‘tipo potenza’ e ‘tipo esponenziale’

$$f_p(x) = \left(\frac{1}{Nx}\right)^n \quad \text{e} \quad f_e(x) = \exp[-Nx]; \quad (\text{A.3})$$

con x di ordine 1. Prendiamo $x_0 = 1$ come punto di riferimento e calcoliamo quanto cambia il valore delle funzioni passando da x_0 a $x_1 = x_0 + \delta > x_0$, analizzando il rapporto $f(x_1)/f(x_0) \equiv r$.

Per le potenze si ha:

$$r_p = \left(\frac{x_0}{x_1}\right)^n = \left(\frac{1}{1 + (\delta/x_0)}\right)^n; \quad (\text{A.4})$$

per l’esponenziale:

$$r_e = \exp[-N(x_1 - x_0)] = \exp[-N\delta]. \quad (\text{A.5})$$

Consideriamo il caso di uno spostamento relativo ‘piccolo’: $\delta/x_0 \ll 1$ (con piccolo intenderemo un numero il cui ordine di grandezza può essere 10^{-1} , 10^{-2} o anche 10^{-6} , stabiliamo $\delta/x_0 = 10^{-3}$) per cui si ha

$$r_p \cong \left(1 - (\delta/x_0)\right)^n \cong 1 - n\left(\frac{\delta}{x_0}\right), \quad (\text{A.6})$$

dove l’ipotesi $\delta/x_0 \ll 1$ ha permesso di fare due approssimazioni al primo ordine, la prima sulla serie geometrica

$$\frac{1}{1-z} = \sum_{k=0}^{\infty} z^k \quad (|z| < 1), \quad (\text{A.7})$$

la seconda sulla potenza di un binomio

$$(1+z)^n = \sum_{k=0}^n \binom{n}{k} z^k. \quad (\text{A.8})$$

Troviamo quindi che spostandosi ‘di poco’ da x_0 , f_p cambia ‘poco’ essendo r_p quasi 1. In effetti dalla (A.4) si capisce che f_p in x_1 ha valore confrontabile con quello in x_0 anche se la variazione relativa non è piccola, per esempio se x_1 è 2 o 3 volte x_0 . Le funzioni a potenza cambiano lentamente sulla scala definita dalla variabile x .

Per l’esponenziale si ha invece (usando i valori proposti sopra)

$$r_e = \exp[-10^{17}], \quad (\text{A.9})$$

cioè praticamente $r_e = 0$: anche spostandosi ‘di poco’ (sulla scala della x) la funzione f_e cambia ‘enormemente’.

È allora interessante chiedersi quali sono i valori di δ per cui $f_e(x_1)$ è ancora confrontabile con $f_e(x_0)$, cioè per cui si ha $r_e \approx 1$. Questo succede se $N\delta \approx 1$ ovvero $\delta \approx 1/N$, che in questo esempio significa $\delta \approx 10^{-20}$: i cambiamenti dell'esponenziale risultano 'piccoli' se la variabile x subisce variazioni che possiamo dire 'infinitesime' (se sono solo 'piccole' f_e cambia enormemente).

È opportuno sottolineare che stiamo discutendo dei *cambiamenti* dei valori delle funzioni; in particolare il valore di f_e in x_0 , assunto come riferimento, in sé non ha alcun significato particolare: quello che abbiamo trovato è che, se δ è 'piccolo' (e positivo), $f_e(x_0)$ risulta enormemente più grande di $f_e(x_0 + \delta)$, ma chiaramente è enormemente più piccolo di $f_e(x_0 - \delta)$ (avendo scelto di studiare una funzione decrescente).

In termini di variazioni relative delle funzioni studiate, dato che

$$\frac{\Delta f}{f} = \frac{f(x_1) - f(x_0)}{f(x_0)} = r - 1, \quad (\text{A.10})$$

si ha che per piccole variazioni di x ($\delta/x_0 = 10^{-3}$) la variazione relativa delle funzioni potenza è piccola dello stesso ordine di grandezza

$$\frac{\Delta f_p}{f_p} = r_p - 1 \cong -n \left(\frac{\delta}{x_0} \right); \quad (\text{A.11})$$

mentre quella dell'esponenziale non è piccola: è praticamente -1 , essendo r_e praticamente 0. In situazioni del genere (δ 'piccolo' ma non 'infinitesimo') per calcolare Δf_p è lecito usare uno sviluppo perturbativo con pochi termini (anche 1), mentre per Δf_e no.

Consideriamo, per esempio, la funzione $g(x) = Nx$ e l'esponenziale $h(x) = \exp[Nx] = \exp[g(x)]$. Nel passaggio $x_0 \rightarrow x_1 = x_0 + \delta$ si ha

$$g(x_1) = N(x_0 + \delta) = g(x_0) + N\delta = g(x_0) + g'(x_0)\delta = g(x_0) + \Delta g \quad (\text{A.12})$$

e, in questo caso molto semplice, $\Delta g = g'(x_0)\delta$, per ogni δ e senza nessuna approssimazione. Ma si ha

$$h(x_1) = \exp[N(x_0 + \delta)] = h(x_0) \exp[N\delta] \quad (\text{A.13})$$

mentre uno sviluppo al primo ordine in x dà

$$\begin{aligned} h(x_1) &= h(x_0) + h'(x_0)\delta = \exp[Nx_0] + N\delta \exp[Nx_0] = \\ &= h(x_0) (1 + N\delta), \end{aligned} \quad (\text{A.14})$$

espressione che fornisce il risultato corretto solo se $N\delta \ll 1$ cioè se δ è 'infinitesimo'.

Più in generale, se $g(x) = N\phi(x)$ e $h(x) = \exp[g(x)] = \exp[N\phi(x)]$, con $\phi(x)$ una funzione 'tranquilla' (continua con tutte le derivate che servono, con valori sulla

scala delle x), e se $x_0 \rightarrow x_1 = x_0 + \delta$, con $\delta/x_0 \ll 1$ si ha

$$\frac{\Delta g}{g(x_0)} = \frac{\Delta \phi}{\phi(x_0)} \cong \frac{\phi'(x_0) \delta}{\phi(x_0)} ; \quad (\text{A.15})$$

cioè la variazione relativa di g è dello stesso ordine della variazione relativa di x e si ha

$$\Delta g \cong g(x_0) \frac{\phi'(x_0) \delta}{\phi(x_0)} = N \phi'(x_0) \delta = g'(x_0) \delta . \quad (\text{A.16})$$

Per la $h(x)$ si ha invece

$$\begin{aligned} \frac{\Delta h}{h(x_0)} &= \exp\{N [\phi(x_0 + \delta) - \phi(x_0)]\} - 1 = \\ &= \exp\{\Delta g\} - 1 \cong \exp\{N \phi'(x_0) \delta\} - 1 , \end{aligned} \quad (\text{A.17})$$

per cui, se δ è piccolo, ma non infinitesimo, assumendo che sia $\phi'(x_0) > 0$

$$\Delta h \cong h(x_0) \exp\{N \phi'(x_0) \delta\} \gg h(x_0) \quad (\text{A.18})$$

e non è utilizzabile l'approssimazione

$$\Delta h \cong h'(x_0) \delta = h(x_0) N \phi'(x_0) \delta \quad (\text{A.19})$$

che, come si vede dalla (A.17), è corretta solo se $N \delta \ll 1$, cioè se δ è non solo piccolo ma infinitesimo.

A.2 Gli integrali

Quanto trovato ha ripercussioni notevoli nel calcolo degli integrali. Supponiamo di dover calcolare

$$I = \int_{x_0}^{x_1} f(x) dx . \quad (\text{A.20})$$

Nel caso $f(x) = f_p(x)$ si ha

$$I_p = \int_{x_0}^{x_1} \frac{dx}{(Nx)^n} = \frac{1}{N^n (n-1)} \left(\frac{1}{x_0^{n-1}} - \frac{1}{x_1^{n-1}} \right) \quad (n > 1) , \quad (\text{A.21})$$

che si può anche scrivere

$$I_p = \frac{N^{-n}}{(n-1)x_0^{n-1}} \left[1 - \left(\frac{x_0}{x_1} \right)^{n-1} \right] . \quad (\text{A.22})$$

Abbiamo già visto che $f_p(x)$ decresce lentamente, la (A.22) ci mostra infatti che la dipendenza dell'integrale da x_1 si fa sentire su un intervallo abbastanza ampio di valori. Solo per x_1 abbastanza grande rispetto a x_0 il termine $(x_0/x_1)^{n-1}$ è trascurabile rispetto a 1, nel qual caso

$$I_p \cong \frac{N^{-n}}{(n-1)x_0^{n-1}} = \int_{x_0}^{\infty} f_p(x) dx \quad (\text{A.23})$$

Per esempio, assumendo $x_0 = 1$, se $x_1 = 10$ e $n = 2$ la correzione a 1 è del 10%, se $n = 3$ è dell'1%, ordini di grandezza 'piccoli' ma non trascurabili.

Nel caso $f(x) = f_e(x)$ si ha invece

$$I_e = \int_{x_0}^{x_1} \exp[-Nx] dx = \frac{\exp[-Nx_0]}{N} [1 - \exp[N(x_1 - x_0)]]; \quad (\text{A.24})$$

quindi anche se x_1 è solo di 'poco' più grande di x_0 , come nel solito esempio $x_1 = x_0 + 10^{-3}$ (e $N = 10^{20}$), la (A.24) mostra che la dipendenza da x_1 è già trascurabile e

$$I_e \cong \frac{\exp[-Nx_0]}{N} = \int_{x_0}^{\infty} f_e(x) dx. \quad (\text{A.25})$$

A causa della rapida variazione di $f_e(x)$, l'integrale I_e riceve contributo significativo solamente dai punti in un intervallo di ordine $1/N$ (per noi, infinitesimo) a partire da x_0 : i punti al di fuori di questo intervallo danno contributo praticamente nullo. Quindi, se l'intervallo di integrazione non è infinitesimo ma, sulla scala delle x , è anche solamente 'piccolo', l'integrale praticamente non dipende da x_1 e può essere esteso fino all'infinito.

Naturalmente il ruolo primario svolto da x_0 è dovuto all'aver considerato, in questo esempio, una f_e decrescente; nel caso di una f_e crescente il valore di I_e sarebbe stato determinato dai punti in un intervallo infinitesimo con estremo superiore in x_1 .

Le considerazioni fatte su I_e valgono più in generale per integrali del tipo

$$I = \int_a^b \alpha(x) \exp[N\phi(x)] dx, \quad (\text{A.26})$$

dove ϕ , come sopra, è una funzione 'tranquilla' insieme ad α . Possiamo considerare tre casi.

- a) Nell'intervallo $[a, b]$ $\phi'(x) > 0$: I sarà determinato dall'andamento dell'integrando in un intervallo 'infinitesimo' prima di b (dove l'esponenziale è massimo).
- b) Nell'intervallo $[a, b]$ $\phi'(x) < 0$: I sarà determinato dall'andamento dell'integrando in un intervallo 'infinitesimo' dopo a (dove l'esponenziale è massimo).
- c) All'interno di $[a, b]$ esiste un punto di massimo per ϕ : $\phi'(x_m) = 0$ con $a < x_m < b$. In questo caso (che è quello che si considera di solito) I sarà determinato dall'andamento dell'integrando in un intorno 'infinitesimo' di x_m . Infatti l'intervallo di

integrazione può essere diviso in 3 parti e si può riscrivere

$$I = \int_a^{x_m-\varepsilon} \alpha(x) \exp [N \phi(x)] dx + \int_{x_m-\varepsilon}^{x_m+\varepsilon} \alpha(x) \exp [N \phi(x)] dx + \int_{x_m+\varepsilon}^b \alpha(x) \exp [N \phi(x)] dx \quad (\text{A.27})$$

dove 2ε è un intervallo ‘piccolo’ intorno a x_m . Il primo e il terzo integrale, essendo determinati dai valori di $\exp [N \phi(x)]$ che risultano massimi nei rispettivi intervalli d’integrazione, saranno di ordine $\exp [N \phi(x_m \pm \varepsilon)]$ e quindi completamente trascurabili rispetto al secondo, di ordine $\exp [N \phi(x_m)]$. A questo punto, si può scrivere

$$I = \int_a^b \alpha(x) \exp [N \phi(x)] dx \cong \int_{x_m-\varepsilon}^{x_m+\varepsilon} \alpha(x) \exp [N \phi(x)] dx \equiv I_\varepsilon \quad (\text{A.28})$$

dove, avendo scelto ε ‘piccolo’, I_ε in effetti non dipende da ε (il suo valore essendo determinato da un intervallo ‘infinitesimo’ intorno a x_m) mentre $\alpha(x)$ e $\phi(x)$ cambiano poco. Si procede quindi approssimando α con il suo valore in x_m e $\phi(x)$ con uno sviluppo al secondo ordine intorno a x_m :

$$I \cong I_\varepsilon \cong \int_{x_m-\varepsilon}^{x_m+\varepsilon} \alpha(x_m) \exp \left[N \left(\phi(x_m) + (1/2) \phi''(x_m)(x-x_m)^2 \right) \right] dx. \quad (\text{A.29})$$

L’ultima osservazione da fare è che, dato che I_ε non dipende da ε i limiti d’integrazione si possono estendere da $-\infty$ a $+\infty$. Si arriva a un integrale Gaussiano (ricordarsi che, per ipotesi, $\phi''(x_m) < 0$) da cui si ottiene

$$I = \alpha(x_m) \exp [N \phi(x_m)] \sqrt{\frac{2\pi}{N|\phi''(x_m)|}}. \quad (\text{A.30})$$

Questa è l’espressione standard del metodo (di Laplace) per calcolare integrali di tipo c). Ragionando allo stesso modo si è in grado di ricavare le espressioni adatte a quelli degli altri due tipi. Ad esempio, se in $[a, b]$ $\phi' < 0$ scriviamo

$$I = \int_a^b \alpha(x) \exp [N \phi(x)] dx \cong \int_a^{a+\varepsilon} \alpha(x) \exp [N \phi(x)] dx, \quad (\text{A.31})$$

con ε ‘piccolo’, in modo che il risultato sia praticamente insensibile al suo valore. Approssimiamo poi le funzioni ‘tranquille’ in questo intervallo, sostituendo a α il suo valore in a e a ϕ il suo sviluppo al primo ordine (che stavolta è $\neq 0$):

$$I \cong \alpha(a) \exp [N \phi(a)] \int_a^{a+\varepsilon} \exp [N \phi'(a)(x-a)] dx. \quad (\text{A.32})$$

Per ultimo estendiamo l'intervallo di integrazione fino a ∞ e otteniamo (si ricordi che $\phi' < 0$)

$$I \cong \alpha(a) \exp [N\phi(a)] \left(-\frac{1}{N\phi'(a)} \right). \quad (\text{A.33})$$

A.3 Volumi

Una conseguenza sorprendente del comportamento degli esponenziali, per noi che abbiamo una intuizione fondata su esperienze tridimensionali, si ha, per esempio, quando si considera il rapporto tra i volumi di sfere di raggio diverso, in spazi di dimensione D , con D molto grande, per esempio $D = 10^{20}$ (come lo spazio delle fasi di un sistema macroscopico).

Definiamo (iper-)sfera di raggio R l'insieme dei punti nello spazio D -dimensionale con distanza minore o uguale a R da un centro comune. Il volume della sfera si può scrivere

$$V_D(R) = C_D R^D, \quad (\text{A.34})$$

dove C_D non dipende da R . Il rapporto tra i volumi di una sfera di raggio $R' = R - \delta < R$ e una di raggio R è quindi dato da

$$r_D = \frac{V_D(R - \delta)}{V_D(R)} = \left(\frac{R - \delta}{R} \right)^D = \exp \left[D \ln \left(1 - \frac{\delta}{R} \right) \right]. \quad (\text{A.35})$$

I volumi sono confrontabili se r_D è una quantità di ordine 1, o 'piccola' nel significato usato sopra, cioè

$$D \ln \left(1 - \frac{\delta}{R} \right) \approx -1 \quad (\text{A.36})$$

$$\ln \left(1 - \frac{\delta}{R} \right) \approx -\frac{1}{D}. \quad (\text{A.37})$$

Poiché stiamo considerando $1/D = 10^{-20}$, che è una quantità infinitesima, si può usare l'approssimazione $\ln(1+x) \cong x$ (valida per $x \ll 1$) e si ricava la condizione

$$\frac{\delta}{R} \approx \frac{1}{D}. \quad (\text{A.38})$$

Abbiamo ottenuto che se la differenza relativa tra R' e R è 'infinitesima' i volumi delle 2 sfere sono confrontabili. Se la differenza relativa è solamente 'piccola', per esempio se si riduce il raggio di un millesimo: $R' = 0,999R$, si ha

$$r_D = \left(\frac{R'}{R} \right)^D = 10^{D \lg [0,999]} \approx 10^{-10^{16}} \quad (\text{A.39})$$

cioè $V_D(R')$ è completamente trascurabile *rispetto* a $V_D(R)$ (ma, a sua volta, è enormemente più grande rispetto ai volumi ‘interni’). Il volume di una (iper-)sfera ad alta dimensione è praticamente concentrato in uno strato infinitesimo al di sotto della sua superficie.

Considerazioni simili si possono fare anche per gli (iper-)cubi in D dimensioni (i cui volumi sono L^D , in funzione del lato L), o per qualunque famiglia di (iper-)solidi convessi: quasi tutto il volume è sulla superficie.

Appendice B

Legge dei grandi numeri, piccole e grandi fluttuazioni

In questa Appendice diamo qualche dettaglio su alcuni teoremi limite della teoria delle probabilità, esaminando diverse possibili situazioni interessanti.

B.1 Dalla legge binomiale al teorema del limite centrale

Iniziamo considerando un esperimento in cui un particolare evento a ha probabilità p di verificarsi e studiamo la distribuzione di probabilità della frequenza di apparizione, v , in una successione di N ripetizioni indipendenti dell'esperimento. Per comodità di scrittura a volte scriveremo $1 - p = q$.

Su una successione lunga N l'evento a può verificarsi n volte, in un certo ordine, con probabilità $p^n q^{N-n}$. Questa, moltiplicata per il numero di modi in cui si può scegliere in quali delle N ripetizioni capitano gli n eventi (numero dato dal coefficiente binomiale) fornisce la probabilità che a si verifichi n volte in un ordine qualunque:

$$P_N(n) = p^n q^{N-n} \binom{N}{n} = p^n q^{N-n} \frac{N!}{n!(N-n)!}. \quad (\text{B.1})$$

Ricordando che tra fattoriali e funzioni Gamma vale la relazione $k! = \Gamma(k+1)$, usando l'approssimazione di Stirling nella forma $\ln k! \approx k \ln k - k$, si ottiene

$$P_N(n) \approx \exp \left[N \left(\frac{n}{N} \ln p + \frac{N-n}{N} \ln q \right) \right] \times \exp \left[N \left(-\frac{n}{N} \ln \frac{n}{N} - \frac{N-n}{N} \ln \frac{N-n}{N} \right) \right], \quad (\text{B.2})$$

si noti che, essendo gli argomenti dei logaritmi quantità minori di 1, il primo esponenziale (che è la probabilità di singola sequenza) ha argomento negativo, mentre il secondo (che dà la molteplicità delle sequenze con assegnata frequenza n/N) ha

argomento positivo. Passando alla variabile $v = n/N$ e ricombinando i due fattori, si ha:

$$P_N(v) \approx \exp \left[-N \left(v \ln \frac{v}{p} + (1-v) \ln \frac{1-v}{1-p} \right) \right] \equiv \exp \left[-N \times I(v||p) \right]. \quad (\text{B.3})$$

Nella (B.3) si è introdotta $I(v||p)$, funzione di v contenente p come parametro, detta “entropia relativa” o “divergenza di Kullback-Leibler”, che è una quantità *non negativa* e che vale zero solo per $v = p$, come si può verificare usando la proprietà di concavità della funzione logaritmo.

Si osservi che se $N \rightarrow \infty$ la variabile v assume un’infinità di valori equispaziati in $[0, 1]$ e si può considerare praticamente una variabile continua. Anche se quella di partenza è una probabilità, come la (B.1) o la (B.2) che sono definite per valori interi di n , quelle a cui si arriva sono delle densità di probabilità.

Nel limite $N \rightarrow \infty$, la probabilità degli insiemi di sequenze con $v \neq p$ (e quindi $I(v||p) \neq 0$) va a 0, e le sole sequenze la cui probabilità totale resta diversa da 0, e quindi tende a 1, sono quelle con $v = p$. In questo limite la distribuzione in frequenza (B.3) è rappresentabile con una densità data da $\delta(v - p)$, che è un modo di enunciare la legge dei grandi numeri.

Per valori $1 \ll N < \infty$ (per noi $N = O(N_0) \sim 10^{24}$) le frequenze cui corrispondono classi di sequenze con probabilità totali significativamente diverse da zero sono quelle per cui $NI(v||p) = O(1)$, cioè $I(v||p) = O(1/N) \approx 0$, e quindi le frequenze interessanti sono quelle con $v \approx p$. Possiamo eseguire uno sviluppo di Taylor partendo da p , punto in cui $I(v||p)$ ha un minimo che vale zero, per cui si ha

$$I(v||p) = \frac{I''(v=p)}{2} (v-p)^2 + \frac{I'''(v=p)}{6} (v-p)^3 + \dots \quad (\text{B.4})$$

Essendo le derivate della I quantità di ordine 1: $I''(v=p) = 1/(pq)$, $I'''(v=p) = 1/q^2 - 1/p^2$, considerando frequenze tali che sia

$$|v-p| = O\left(\frac{1}{\sqrt{N}}\right), \quad (\text{B.5})$$

si ha

$$I(v||p) = \frac{I''(v=p)}{2} O\left(\frac{1}{N}\right) + \frac{I'''(v=p)}{6} O\left(\frac{1}{N\sqrt{N}}\right) + \dots \quad (\text{B.6})$$

L’intervallo di valori definito dalla (B.5), seleziona le *piccole fluttuazioni* per la v . In questa zona, come si vede dalla (B.6), lo sviluppo al second’ ordine per la $I(v||p)$ è una buona approssimazione e le probabilità corrispondenti sono ben approssimate da una gaussiana

$$P_N(v) \approx \exp \left[-N \left(\frac{I''(v=p)}{2} (v-p)^2 \right) \right] = \exp \left[-N \left(\frac{(v-p)^2}{2pq} \right) \right], \quad (\text{B.7})$$

che, correttamente normalizzata, si riscrive

$$P_N(v) \approx \frac{1}{\sqrt{2\pi pq/N}} \exp \left[- \left(\frac{(v-p)^2}{2pq/N} \right) \right]; \quad (\text{B.8})$$

essa è caratterizzata da valor medio p e varianza $\sigma_N^2 = pq/N$ e traduce il teorema del limite centrale applicato al caso della variabile $v = n/N$.

Al di fuori del range delle piccole fluttuazioni, per esempio se $|v-p| = O(1)$, si entra nel dominio delle *grandi fluttuazioni*. Qui uno sviluppo di Taylor con pochi termini non è più utilizzabile e, tipicamente, $I(v||p) = O(1)$ cosicché le classi di sequenze corrispondenti hanno probabilità totale data dalla (B.3), cioè esponenzialmente piccola in N .

Una volta in possesso dello scenario relativo alle probabilità delle sequenze, in funzione di v , si possono riformulare i risultati sopra esposti in termini di ogni altra variabile funzione della frequenza come, per esempio, la variabile da cui siamo partiti: $n = vN$. Eseguendo il cambiamento di variabili si ha:

$$P_N(n) \approx \frac{1}{\sqrt{2\pi pqN}} \exp \left[- \left(\frac{(n-Np)^2}{2pqN} \right) \right], \quad (\text{B.9})$$

cioè una gaussiana centrata in Np con varianza Npq , che descrive la statistica delle piccole fluttuazioni della variabile n , quelle tali che

$$|n - Np| = O(\sqrt{N}). \quad (\text{B.10})$$

Notare che anche le grandi fluttuazioni di n , quelle corrispondenti a $|v-p| = O(1)$ e quindi con $|n - Np| = O(N)$, sono regolate dalla (B.3).

B.2 Oltre il limite centrale

Tutto quanto precedentemente discusso, nel caso di un esperimento che abbia solo due possibili esiti, può essere ripetuto nel caso in cui in ogni esperimento elementare si possano presentare (o l'apparato sperimentale è in grado di distinguere) d eventi a_1, a_2, \dots, a_d con probabilità $p_1, p_2, \dots, p_d \equiv \{p\}$. Su ogni successione di N esperimenti si contano le volte che si è presentato ognuno degli eventi $k_1, k_2, \dots, k_d \equiv \{k\}$ e si determinano le frequenze di apparizione $v_i = k_i/N$. La probabilità delle sequenze in cui gli eventi a_i si sono verificati k_i volte è allora

$$P_N(\{k\}) = p_1^{k_1} p_2^{k_2} \dots p_d^{k_d} \frac{N!}{k_1! k_2! \dots k_d!}, \quad (\text{B.11})$$

espressione che, usando l'approssimazione di Stirling e passando alle frequenze, diventa

$$P_N(\{v\}) \approx \exp \left[-N \times I(\{v\} \parallel \{p\}) \right], \quad (\text{B.12})$$

in cui si è posto

$$I(\{v\} \parallel \{p\}) = \sum_{i=1}^d v_i \ln \frac{v_i}{p_i}. \quad (\text{B.13})$$

Come la $I(v \parallel p)$ definita nella (B.3), anche l'entropia relativa di $\{v\}$ rispetto a $\{p\}$, $I(\{v\} \parallel \{p\})$, è una funzione non negativa di $\{v\}$ con minimo pari a zero in $\{v\} = \{p\}$. A partire dalla (B.12), procedendo analogamente a come fatto sopra, si ottengono:

- a) la legge dei grandi numeri: nel limite di sequenze infinitamente lunghe, quelle in cui $v_i \rightarrow p_i$, $\forall i$, hanno probabilità totale che tende a 1, la probabilità di tutte le altre tende a 0;
- b) la distribuzione gaussiana per le piccole fluttuazioni delle frequenze: se $|v_i - p_i| = O(1/\sqrt{N})$, $\forall i$, sviluppando al second' ordine intorno al minimo si ha

$$I(\{v\} \parallel \{p\}) = \frac{1}{2} \sum_{i=1}^d \frac{1}{p_i} (v_i - p_i)^2 + \dots \quad (\text{B.14})$$

e quindi

$$P_N(\{v\}) \approx \prod_i \frac{1}{\sqrt{2\pi p_i/N}} \exp \left[- \left(\frac{(v_i - p_i)^2}{2 p_i/N} \right) \right], \quad (\text{B.15})$$

con il vincolo $\sum_i v_i = 1 = \sum_i p_i$ ¹;

- c) il decadimento esponenzialmente veloce in N in presenza di grandi fluttuazioni, ovvero se per qualche i si ha $|v_i - p_i| = O(1)$.

Supponiamo adesso che gli eventi elementari a_1, a_2, \dots, a_d siano caratterizzati per mezzo di una variabile numerica x : all'evento a_i è associato il valore x_i , la cui probabilità è quindi p_i . Su ogni successione lunga N definiamo la variabile $Y_N = \sum_{i=1}^N x^{(i)}$ e ci chiediamo con quale probabilità si presenta il valore \tilde{Y} . Dato che sulle sequenze vale la relazione $\sum_{i=1}^N x^{(i)} = \sum_{j=1}^d x_j k_j$, si ha:

$$P(Y_N = \tilde{Y}) = \sum_{\{k\}: \sum_j x_j k_j = \tilde{Y}} P_N(\{k\}), \quad (\text{B.16})$$

dove P_N è dato dalla (B.11). La presenza del vincolo nella (B.16) rende questo calcolo più complicato di quelli precedenti e suggerisce di usare un metodo alterna-

¹ Si verifichi che la (B.7) si può scrivere anche nella forma:

$$\exp \left[- \frac{N}{2} \left(\frac{(v_1 - p)^2}{p} + \frac{(v_2 - q)^2}{q} \right) \right] \quad \text{con} \quad v_1 + v_2 = 1.$$

tivo, che passa attraverso la definizione della *funzione caratteristica* di una variabile casuale (in questo caso la x)

$$\phi_x(t) = \sum_{j=1}^d e^{itx_j} p_j = \langle e^{itx} \rangle. \tag{B.17}$$

La funzione caratteristica costituisce uno strumento analitico molto potente per la dimostrazione della legge dei grandi numeri e del teorema del limite centrale. In particolare, ponendo $\langle x \rangle = \mu$ e $\langle (x - \langle x \rangle)^2 \rangle = \sigma^2$, si ha

- nel limite $N \rightarrow \infty$ la distribuzione di $y = Y_N/N$ tende a $\delta(y - \mu)$ (legge di grandi numeri);
- nello stesso limite, la distribuzione di $z = (Y_N - N\mu)/\sqrt{N}$ è una gaussiana centrata in 0 con varianza σ^2 (teorema del limite centrale).

L'utilizzo delle funzioni caratteristiche permette inoltre una immediata generalizzazione al caso continuo in cui il risultato di ciascun esperimento possa assumere i valori di una variabile continua distribuita con una densità $p(x)$. In questo caso, sempre per la somma $Y_N = \sum_{i=1}^N x^{(i)}$, valgono i risultati precedenti nell'ipotesi che esistano (finiti) il valor medio μ e la varianza σ^2 della x .

B.3 Grandi deviazioni

Impostando un calcolo combinatorio, nelle precedenti sezioni B.1 e B.2 abbiamo ricavato le equazioni (B.1) e (B.11) che forniscono le probabilità $P_N(n)$ e $P_N(\{k\})$. Abbiamo poi visto che il decadimento delle probabilità degli eventi “estremi” (quelli sulle “code” delle distribuzioni) è regolato dalle funzioni $I(v||p)$ e $I(\{v\}||\{p\})$.

Queste sono esempi particolari delle cosiddette *funzioni di Cramer*. È interessante e utile ritrovare questo risultato per mezzo di considerazioni di carattere generale, che possano quindi applicarsi a uno schema probabilistico qualunque.

Illustriamo brevemente qui di seguito i passaggi essenziali che conducono al risultato e che mostrano come l'andamento delle code di una distribuzione sia legato alla funzione generatrice dei momenti.

- 1) Per iniziare consideriamo una variabile x , che può assumere d valori x_1, x_2, \dots, x_d , con probabilità p_1, p_2, \dots, p_d ; indichiamo con x_{max} il massimo tra le x_i . Ricordiamo la definizione della *funzione generatrice dei momenti*² $M(t) = \sum_{j=1}^d e^{tx_j} p_j = \langle e^{tx} \rangle$, che è una funzione convessa ($M''(t) \geq 0$), e la disuguaglianza di Markov

$$P(|x| \geq x_0) \leq \frac{1}{x_0^k} \langle |x|^k \rangle \quad (x_0 > 0). \tag{B.18}$$

² Così chiamata perchè la sua derivata k -ima in $t = 0$ è $\langle x^k \rangle$, momento di ordine k della x .

- 2) Scegliamo un α tale che risulti $\langle x \rangle < \alpha < x_{max}$ e ci chiediamo qual è la probabilità che x sia maggiore o uguale ad α . Possiamo scrivere

$$\begin{aligned} P(x \geq \alpha) &= P(x - \alpha \geq 0) = \\ &= P(e^{t(x-\alpha)} \geq 1) \leq \langle e^{t(x-\alpha)} \rangle = e^{-t\alpha} M(t), \end{aligned} \quad (\text{B.19})$$

avendo applicato la (B.18), con $k = 1$, alla variabile positiva $e^{t(x-\alpha)}$. La disuguaglianza scritta vale per ogni t , per cui

$$P(x \geq \alpha) \leq \min_t [e^{-t\alpha} M(t)] \equiv \rho(\alpha \| \{p\}), \quad (\text{B.20})$$

dove abbiamo indicato esplicitamente che la funzione ρ contiene una dipendenza dalla distribuzione di probabilità considerata.

- 3) Essendo $[e^{-t\alpha} M(t)]$ una funzione positiva, ponendo $\ln M(t) = C(t)$, si ha

$$\min_t [e^{-t\alpha} M(t)] = \exp\left(\min_t [C(t) - t\alpha]\right) = e^{\ln \rho}, \quad (\text{B.21})$$

per cui la (B.20) si può riscrivere come

$$P(x \geq \alpha) \leq \exp[\ln \rho(\alpha \| \{p\})] = \exp[L(\alpha \| \{p\})], \quad (\text{B.22})$$

dove $L(\alpha \| \{p\}) = \min_t [C(t) - t\alpha]$ indica la trasformata di Legendre di $C(t)$ che, in teoria della probabilità, è nota come la *funzione generatrice dei cumulanti*.

- 4) La disuguaglianza (B.22) vale per variabili qualunque. Se la si applica al caso di una variabile somma di variabili indipendenti e identiche $Y_N = \sum_{i=1}^N x^{(i)}$, per la quantità $P(Y_N \geq N\alpha > N\langle x \rangle)$ vale il *teorema di Chernoff*:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln P(Y_N \geq N\alpha) = \ln \rho(\alpha \| \{p\}) = L(\alpha \| \{p\}), \quad (\text{B.23})$$

cioè: per grandi valori di N la disuguaglianza è praticamente un'uguaglianza:

$$P(Y_N \geq N\alpha) \approx \exp[NL(\alpha \| \{p\})]. \quad (\text{B.24})$$

La formula (B.24) è il risultato cercato, a proposito del quale, facciamo le seguenti osservazioni.

– Dato che, per sua definizione, $\rho(\alpha \| \{p\}) < 1$ e quindi $L(\alpha \| \{p\}) < 0$, a volte, è comodo scrivere

$$P(Y_N \geq N\alpha) \approx \exp[-N\Lambda(\alpha \| \{p\})], \quad (\text{B.25})$$

avendo posto $L(\alpha \| \{p\}) = -\Lambda(\alpha \| \{p\}) = -\max_t [t\alpha - C(t)]$.

– Grazie alle proprietà delle funzioni esponenziali, illustrate nell'appendice A, ci aspettiamo che (se δ è una quantità di ordine α) si abbia

$$P(N\alpha < Y_N < N(\alpha + \delta)) = P(Y_N \geq N\alpha) - P(Y_N \geq N(\alpha + \delta)) \approx P(Y_N \geq N\alpha), \quad (\text{B.26})$$

cioè: come il volume di un'ipersfera è concentrato immediatamente sotto la sua superficie, la probabilità dell'evento $[Y_N \geq N\alpha]$ è dominata da quella dell'evento $[Y_N = N\alpha]$; per cui si può scrivere

$$P(Y_N \cong N\alpha) \approx P(Y_N \geq N\alpha) \cong \exp[-N\Lambda(\alpha||\{p\})]. \quad (\text{B.27})$$

– Se si considera una variabile x che può assumere due valori: $x = 1$ con probabilità p , $x = 0$ con probabilità $q = 1 - p$, sulle sequenze di lunghezza N si ha $Y_N = \sum_{i=1}^N x^{(i)} = n$, dove n è la stessa variabile studiata nella sottosezione B.1. Si può verificare che, in questo caso, la definizione data per la funzione $\Lambda(\alpha||\{p\})$ conduce al risultato $\Lambda(\alpha||p) = I(\alpha||p)$.

– Possiamo adesso tralasciare la condizione di partenza $\alpha > \langle x \rangle$ (che ci serviva in quanto interessati alle code) e studiare la funzione $\Lambda(\alpha||\{p\}) = \max_t [t\alpha - C(t)]$.

A causa delle proprietà di convessità di $M(t)$ e quindi di $C(t)$, dalla definizione segue che anche $\Lambda(\alpha)$ è una funzione convessa del suo argomento, oltre a essere positiva. Considerazioni elementari mostrano che essa ha un minimo in $\alpha^* = \langle x \rangle$, punto in cui la funzione vale zero: come nelle sezioni precedenti, questo segnala che per Y_N/N vale la legge dei grandi numeri e la sua densità tende a $\delta((Y_N/N) - \langle x \rangle)$.

Uno sviluppo al second'ordine intorno a α^* conduce al teorema del limite centrale, ovvero all'andamento gaussiano delle piccole fluttuazioni di Y_N

$$P(Y_N = \tilde{Y}) = \frac{1}{\sqrt{2\pi N/\Lambda''(\alpha^*)}} \exp \left[- \left(\frac{(\tilde{Y} - N\langle x \rangle)^2}{2N/\Lambda''(\alpha^*)} \right) \right]. \quad (\text{B.28})$$

– La relazione (B.27) continua a valere anche se x è una variabile a valori continui con densità $p(x)$, per la quale esista la generatrice dei momenti $M(t)$, almeno in un intorno di $t = 0$.

– I risultati illustrati in questa sottosezione si possono applicare alle grandi fluttuazioni delle variabili Y_N della precedente sottosezione B.2.

Appendice C

Un altro modo di scrivere la densità nel microcanonico

Nel secondo capitolo abbiamo introdotto la densità microcanonica per un sistema la cui energia è definita con un'incertezza Δ : la probabilità di trovare il sistema in un volume infinitesimo $d\mathbf{X}$ del suo spazio delle fasi è $d\mathbf{X}/\Gamma_{\Delta}(E)$. In questa appendice ricaviamo la densità microcanonica nel caso di energia esattamente definita.

Secondo il teorema di Liouville, per un sistema hamiltoniano i volumi nello spazio delle fasi si mantengono costanti durante l'evoluzione. Se un dato insieme di punti occupa una regione di volume $V(t_0)$ a un tempo t_0 , lo stesso insieme al tempo t si troverà in una regione di volume $V(t) = V(t_0)$. L'uguaglianza, in particolare, vale per i volumi infinitesimi $dV(t) = dV(t_0)$.

Nel capitolo 3 abbiamo notato che anche la densità degli stati

$$\omega(E, V, N) = \lim_{dE \rightarrow 0} \frac{\int_{H(\mathbf{X}) < E + dE} d\mathbf{X} - \int_{H(\mathbf{X}) < E} d\mathbf{X}}{dE} = \frac{\partial \Sigma(E, V, N)}{\partial E} \quad (\text{C.1})$$

è un invariante del moto, ma essa è relativa agli stati che occupano non un volume ma una superficie (di energia costante) nello spazio delle fasi. Studiamo più da vicino questa quantità.

Calcoliamo $\omega(E, V, N)$ come derivata di

$$\Sigma(E, V, N) = \int_{H(\mathbf{X}) < E} d\mathbf{X} \quad (\text{C.2})$$

riscrivendo l'integrale in modo più opportuno. Introduciamo nuove variabili che tengano conto del ruolo particolare dell'energia. Il cambiamento di variabili è tale che in ogni punto dello spazio delle fasi per l'elemento di volume si ha $d\mathbf{X} = dX_1 dX_2 \dots dX_{6N} = d\sigma dn$, dove $d\sigma$ indica l'elemento infinitesimo della superficie di energia costante passante per il punto dato e dn è l'elemento di linea lungo la normale alla superficie¹. Il versore lungo la normale \hat{n} ha come componenti i coseni

¹ Per esempio, nel caso (non hamiltoniano) di \mathbf{X} tridimensionale e $H(\mathbf{X})$ dipendente solo da $|\mathbf{X}|$, passando a coordinate sferiche (r, θ, ϕ) si ha $d\mathbf{X} = [r^2 \sin(\theta) d\theta d\phi] [dr] \equiv [d\sigma] [dn]$.

direttori $\cos(\hat{n}, X_i)$ che, in questo caso, si possono scrivere anche

$$\cos(\hat{n}, X_i) = \frac{\partial H / \partial X_i}{|\nabla H|}. \quad (\text{C.3})$$

Poiché $dX_i = dn \cos(\hat{n}, X_i)$, possiamo ricavare la relazione che lega la distanza infinitesima tra superfici vicine (dn) e la loro differenza energetica (dE):

$$dE = \sum_i \frac{\partial H}{\partial X_i} dX_i = dn \sum_i \frac{\partial H}{\partial X_i} \cos(\hat{n}, X_i) = dn \frac{(\nabla H)^2}{|\nabla H|}, \quad (\text{C.4})$$

da cui otteniamo $dn = dE / |\nabla H|$ e quindi

$$d\mathbf{X} = dE \frac{d\sigma}{|\nabla H|}. \quad (\text{C.5})$$

In pratica, con questo cambiamento di variabili abbiamo diviso lo spazio delle fasi negli strati infinitesimi di energia fissata, cosicché l'equazione (C.2) diventa

$$\Sigma(E, V, N) = \int_0^E dE' \int_{\sigma(E')} \frac{d\sigma}{|\nabla H|}, \quad (\text{C.6})$$

dove l'integrale in $d\sigma$ per ogni fissato valore E' va fatto sulla corrispondente superficie di energia data e dove abbiamo assunto che le energie partano da zero. Il risultato di questa integrazione sarà una funzione di E' (oltre che di V e N) che, per la definizione (C.1), è $\omega(E', V, N)$:

$$\omega(E, V, N) = \frac{\partial \Sigma}{\partial E} = \int_{\sigma(E)} \frac{d\sigma}{|\nabla H|}. \quad (\text{C.7})$$

La relazione trovata esprime la densità degli stati come un integrale su una superficie e vale, in effetti, per una qualunque funzione nello spazio delle fasi che sia abbastanza regolare, in modo che sia possibile usare i suoi valori per "affettare" lo spazio.

Ciò che caratterizza il caso dell'energia è che, se questa si conserva, allora non solo $\omega(E, V, N)$ si mantiene costante nel tempo ma, come segue dalla (C.5) e dal teorema di Liouville, anche $d\sigma / |\nabla H|$ resta costante durante il moto. Se l'estensione di un insieme J sulla superficie di energia E è pesata con $d\sigma / |\nabla H|$, la misura dell'insieme

$$\mathcal{M}(J) = \int_J \frac{d\sigma}{|\nabla H|}, \quad (\text{C.8})$$

resterà costante nel tempo, seguendo l'evoluzione dei punti di J ; mentre in generale la quantità $\int_{J(t)} d\sigma$ dipenderà dal tempo. Come $d\mathbf{X}$ è una misura invariante nello spazio delle fasi così $d\sigma / |\nabla H|$ è una *misura invariante* sulla superficie di energia fissata.

Appendice D

La teoria di Einstein del moto Browniano

Nel suo approccio al moto Browniano Einstein utilizzò due risultati che all'epoca erano ben noti:

- a) la legge di Stokes per la forza d'attrito che agisce su una sfera di raggio R che si muove con velocità v in un fluido con viscosità η , che per semplicità di notazione scriviamo per un moto in una dimensione:

$$F_S = -6\pi R\eta v .$$

- b) la legge di Van' t Hoff per la pressione osmotica.

L'idea, semplice e rivoluzionaria allo stesso tempo, fu di assumere che la meccanica statistica valga anche per oggetti macroscopici (della dimensione del micron).

Consideriamo particelle di massa m immerse in un liquido in un campo di gravità costante diretto lungo l'asse z . Dalla formula barotropica¹ abbiamo che la densità di particelle è:

$$\rho(z) = \rho(0)e^{-mgz/k_B T} . \quad (D.1)$$

La forza totale che agisce su ogni particella è:

$$F_S - mg$$

che si annulla quando

$$v = v^* = -\frac{mg}{6\pi\eta R} ,$$

la velocità tende al suo valore asintotico con un tempo caratteristico $\tau = m/(6\pi R\eta)$ che è estremamente piccolo (da un punto di vista macroscopico). Alla quota z abbiamo una densità di corrente di particelle (verso il basso)

$$j_-(z) = v^* \rho(z) = -\frac{mg}{6\pi\eta R} \rho(z) . \quad (D.2)$$

¹ Si può ottenere lo stesso risultato usando la legge di Van't Hoff per la pressione osmotica.

Poiché la densità non è costante le particelle diffondono da concentrazioni maggiori a concentrazioni minori secondo l'equazione

$$\frac{\partial}{\partial t}\rho(z,t) = D \frac{\partial^2}{\partial z^2}\rho(z,t), \quad (\text{D.3})$$

che può essere riscritta come

$$\frac{\partial}{\partial t}\rho(z,t) - \frac{\partial}{\partial z}\left(D \frac{\partial}{\partial z}\rho(z,t)\right) = 0,$$

ove D è il coefficiente di diffusione finora incognito. Dall'equazione di diffusione nella forma precedente abbiamo una densità di corrente (verso l'alto) $j_+(z)$ dovuta alla migrazione di particelle da concentrazioni maggiori verso quelle minori:

$$j_+(z) = -D \frac{\partial}{\partial z}\rho(z,t),$$

che, tenendo conto della (D.1), diventa

$$j_+(z) = D \frac{mg}{k_B T}\rho(z,t). \quad (\text{D.4})$$

In una situazione di stazionarietà i due flussi di devono bilanciare $j_+ + j_- = 0$, usando la (D.2) e la (D.4) si ottiene

$$D = \frac{k_B T}{6\pi\eta R}.$$

È facile verificare che il risultato non cambia se invece del caso con forza di gravità costante $-mg$ si considerano particelle colloidali sottoposte ad una generica forza $f(z) = -dU(z)/dz$. Il lettore è invitato a ripetere i calcoli e a ritrovare il risultato finale.

La derivazione di Einstein dell'equazione di diffusione

La procedura che usò Einstein per ottenere l'equazione di diffusione (D.3) è quella comunemente usata oggi per ricavare l'equazione di Fokker-Planck, di cui la (D.3) è un caso particolare. Per completezza la riportiamo.

Indicando con $\phi(s, \Delta t)$ la densità di probabilità di avere uno spostamento s in un intervallo di tempo Δt e assumendo che le proprietà statistiche di s non dipendono del punto di partenza abbiamo

$$\rho(z, t + \Delta t) = \int \rho(z - s, t) \phi(s, \Delta t) ds;$$

per Δt piccolo si ha

$$\rho(z, t) + \frac{\partial \rho(z, t)}{\partial t} \Delta t = \rho(z, t) \int \phi(s, \Delta t) ds - \frac{\partial \rho(z, t)}{\partial z} \int s \phi(s, \Delta t) ds + \frac{1}{2} \frac{\partial^2 \rho(z, t)}{\partial z^2} \int s^2 \phi(s, \Delta t) ds + \dots$$

Notando che $\int \phi(s, \Delta t) ds = 1$ e $\int s \phi(s, \Delta t) ds = 0$, indicando con $\langle s^2(\Delta t) \rangle = \int s^2 \phi(s, \Delta t) ds$, trascurando i termini di ordine superiore indicati dai puntini e introducendo

$$D = \lim_{\Delta t \rightarrow 0} \frac{\langle s^2(\Delta t) \rangle}{2\Delta t},$$

abbiamo la (D.3).

Per una derivazione più accurata si deve richiedere che

$$\lim_{\Delta t \rightarrow 0} \frac{\langle s^n(\Delta t) \rangle}{\Delta t} = 0,$$

per $n \geq 3$. Nell'equazione di Fokker-Planck in genere lo spostamento s dipende anche dal punto di partenza z quindi la densità di probabilità è della forma $\phi(s, z, \Delta t)$ e quindi c'è un termine in più:

$$\frac{\partial}{\partial t} p(z, t) = -\frac{\partial}{\partial z} (a(z)p(z, t)) + D \frac{\partial^2}{\partial z^2} p(z, t),$$

dove

$$a(z) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int s \phi(s, z, \Delta t) ds,$$

che è assente (per simmetria) nel caso che stiamo considerando.

Appendice E

Relazione di Taylor-Kubo per il coefficiente di diffusione

Nel Capitolo 5 abbiamo visto che per il modello a tempi discreti dell'equazione di Langevin si ha $\langle v_0 v_k \rangle = d^k \langle v^2 \rangle$; nel limite continuo la funzione di correlazione $\langle v(t)v(0) \rangle$ è proporzionale a $e^{-t/\tau}$ e si ottiene una relazione tra il coefficiente di diffusione, la velocità quadratica media ed il tempo caratteristico:

$$D = \langle v^2 \rangle \tau . \quad (\text{E.1})$$

La formula precedente vale in generale se si intende con τ un “tempo caratteristico” definito in modo non ambiguo dalla funzione di correlazione. Per semplicità di notazione consideriamo un caso $1d$ a tempo continuo

$$x(t) - x(0) = \int_0^t v(t') dt'$$

facendo il quadrato e mediando si ottiene

$$\langle [x(t) - x(0)]^2 \rangle = \int_0^t \int_0^t \langle v(t_1)v(t_2) \rangle dt_1 dt_2 .$$

Se la funzione di correlazione decade a zero più velocemente di t^{-1} allora è facile mostrare che

$$\langle [x(t) - x(0)]^2 \rangle \simeq 2t \int_0^\infty \langle v(t')v(0) \rangle dt' , \quad (\text{E.2})$$

in altre parole

$$D = \int_0^\infty \langle v(t')v(0) \rangle dt' . \quad (\text{E.3})$$

Per ricavare la (E.2) una volta effettuato il cambio di variabili $(t_1, t_2) \rightarrow (t' = t_1, t'' = t_2 - t_1)$, usando la proprietà $\langle v(t')v(t' + t'') \rangle = \langle v(t')v(t' - t'') \rangle$ (conseguenza della stazionarietà) abbiamo

$$\int_0^t \int_0^t \langle v(t_1)v(t_2) \rangle dt_1 dt_2 = 2 \int_0^t \left[\int_0^{t-t''} \langle v(t')v(t' + t'') \rangle dt' \right] dt'' ,$$

se $\langle v(t')v(t'+t'') \rangle$ decade più rapidamente di $1/t''$ allora per t grande si ha

$$\int_0^{t-t''} \langle v(t')v(t'+t'') \rangle dt'' \simeq \int_0^\infty \langle v(t')v(t'+t'') \rangle dt''$$

e quindi segue la (E.2).

Una definizione naturale di tempo caratteristico è

$$\tau = \frac{1}{\langle v^2 \rangle} \int_0^\infty \langle v(t')v(0) \rangle dt',$$

da cui segue (E.1).

Come esercizio il lettore può ripetere i calcoli per il caso a tempo discreto

$$x_{n+1} = x_n + v_n,$$

mostrando che, se $\langle v_j v_0 \rangle$ decade più velocemente di j^{-1} allora per $n \gg 1$ si ha

$$\langle [x_n - x_0]^2 \rangle \simeq 2n \left(\frac{1}{2} \langle v^2 \rangle + \sum_{j=1}^\infty \langle v_j v_0 \rangle \right).$$

La formula (E.3), e la sua analoga a tempi discreti

$$D = \lim_{n \rightarrow \infty} \frac{\langle (x_n - x_0)^2 \rangle}{2n} = \frac{1}{2} \langle v^2 \rangle + \sum_{j=1}^\infty \langle v_j v_0 \rangle,$$

è chiamata relazione di Taylor-Kubo.

Appendice F

Lo sviluppo del viriale per gas debolmente interagenti

Consideriamo un gas di N particelle in un volume V , con Hamiltoniana

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} u_{ij}$$

dove

$$u_{ij} = U(|\mathbf{q}_i - \mathbf{q}_j|).$$

Per trovare l'equazione di stato

$$P = -\frac{\partial F}{\partial V}$$

è necessario calcolare (almeno approssimativamente) l'energia libera $F = -k_B T \ln Z$ e quindi la funzione di partizione

$$Z = \int e^{-\beta H} d\mathbf{p}_1 \dots d\mathbf{p}_N d\mathbf{q}_1 \dots d\mathbf{q}_N = (2\pi m k_B T)^{3N/2} Q(V, N, T)$$

con

$$Q = \int e^{-\beta \sum_{i<j} u_{ij}} d\mathbf{q}_1 \dots d\mathbf{q}_N = \int \prod_{i<j} (1 + f_{ij}) d\mathbf{q}_1 \dots d\mathbf{q}_N,$$

avendo posto $f_{ij} = e^{-\beta u_{ij}} - 1$.

Se il gas è molto diluito, cioè la distanza tipica tra particelle vicine $\delta = (V/N)^{1/3}$ è molto maggiore della portata del potenziale $U(r)$, allora il contributo dell'interazione sarà piccolo rispetto al termine del gas perfetto:

$$Q = \int \prod_{i<j} (1 + f_{ij}) d\mathbf{q}_1 \dots d\mathbf{q}_N = \int (1 + \sum_{i<j} f_{ij} + \dots) d\mathbf{q}_1 \dots d\mathbf{q}_N$$

dove i punti di sospensione indicano termini tipo $f_{ij} f_{i'j'}$, $f_{ij} f_{i'j'} f_{i''j''}$ e così via, che per un gas diluito danno contributi trascurabili. Tenendo presente che con N oggetti

si possono fare $N(N-1)/2$ coppie, per $N \gg 1$ abbiamo

$$Q = V^N + \frac{N^2}{2} V^{N-2} \int f_{12}(|\mathbf{q}_1 - \mathbf{q}_2|) d\mathbf{q}_1 d\mathbf{q}_2 + \dots$$

Usando la variabile $\mathbf{r} = \mathbf{q}_2 - \mathbf{q}_1$ e sfruttando la simmetria radiale del potenziale si ha

$$Q = V^N + 2\pi N^2 V^{N-1} \int_0^\infty r^2 f(r) dr + \dots = V^N - N^2 V^{N-1} b_2(T) + \dots$$

dove $f(r) = e^{-\beta U(r)} - 1$ e

$$b_2(T) = -2\pi \int_0^\infty r^2 f(r) dr = 2\pi \int_0^\infty r^2 (1 - e^{-\beta U(r)}) dr. \quad (\text{F.1})$$

Abbiamo quindi

$$\ln Q = \ln V^N \left[1 - \frac{N^2}{V} b_2(T) + \dots \right] \simeq N \ln V - \frac{N^2}{V} b_2(T) + \dots$$

e, calcolando la pressione, otteniamo l'equazione di stato

$$P = k_B T \frac{\partial}{\partial V} \ln Q \simeq k_B T \left[\frac{N}{V} + \frac{N^2}{V^2} b_2(T) + \dots \right]$$

$$\frac{P}{k_B T} = \rho + b_2(T) \rho^2 + \dots \quad (\text{F.2})$$

Il gas di Tonks

Discutiamo ora un modello molto semplificato per un gas interagente in cui è possibile scrivere in forma esatta l'equazione di stato; questa soluzione è interessante in quanto permette un confronto con lo sviluppo del viriale. Il gas di Tonks è un sistema unidimensionale costituito da N bastoncini rigidi di lunghezza b che si muovono liberamente nell'intervallo $(0, L)$ e non possono sovrapporsi. Indicando con q_i l'origine dell' i -mo bastoncino, poniamo

$$0 < q_1 < q_2 < \dots < q_{i-1} < q_i < q_{i+1} < \dots < q_N < L,$$

con la condizione $q_{i+1} - q_i > b$. Il moto libero senza sovrapposizione è realizzato assumendo un potenziale $U(q_{i+1} - q_i) = \infty$ se $q_{i+1} - q_i < b$ e zero altrimenti.

Per la Q si ha

$$Q = \int_{(N-1)b}^L dq_N \int_{(N-2)b}^{q_N-b} dq_{N-1} \dots \int_b^{q_3-b} dq_2 \int_0^{q_2-b} dq_1.$$

con il cambiamento di variabili $y_n = q_n - (n - 1)b$ otteniamo

$$Q = \int_0^{L-(N-1)b} dy_N \int_0^{y_N} dy_{N-1} \dots \int_0^{y_3} dy_2 \int_0^{y_2} dy_1$$

$$= \frac{1}{N!} (L - (N - 1)b)^N = \frac{1}{N!} L^N (1 - \rho b)^N,$$

dove $\rho = N/L$ è la densità che, ovviamente, deve essere minore di $1/b$.

L'equazione di stato è

$$\frac{P}{k_B T} = \rho + \frac{b\rho^2}{(1 - b\rho)} = \frac{\rho}{(1 - b\rho)},$$

ed è una relazione esatta per ogni densità consentita. È interessante considerare il caso diluto ($\rho \ll 1/b$), effettuando uno sviluppo al primo ordine in $b\rho$: il risultato che si ottiene è lo stesso che si avrebbe con lo sviluppo del viriale (F.2), con b_2 non dipendente dalla temperatura

$$\frac{P}{k_B T} \simeq \rho + b\rho^2,$$

infatti in questo caso, ricordando che il sistema è unidimensionale, abbiamo

$$b_2 = \int_0^\infty (1 - e^{-\beta U(r)}) dr = b.$$

Appendice G

Le catene di Markov in due parole

Le catene di Markov¹ costituiscono il più semplice esempio non banale di processo stocastico a tempi discreti e stati discreti. Un processo stocastico x_t che al tempo discreto t può trovarsi in uno tra M possibili stati X_1, \dots, X_M è una catena di Markov (CM) se lo stato futuro dipende solo dal suo stato presente², in formule:

$$\text{Prob}(x_t = i_t | x_{t-1} = i_{t-1}, \dots, x_{t-n} = i_{t-n}) = \text{Prob}(x_t = i_t | x_{t-1} = i_{t-1}),$$

per ogni t , ove $i_t = 1, \dots, M$. In altre parole la transizione dallo stato $x_{t-1} = X_i$ a $x_t = X_j$ avviene con probabilità

$$\text{Prob}(x_t = j | x_{t-1} = i) = P_{i \rightarrow j} = W_{ji}.$$

Per semplicità consideriamo solo il caso di catene temporalmente omogenee, per le quali le probabilità di transizione non dipendono dal tempo t . Notare che gli elementi di matrice W_{ji} non possono essere completamente arbitrari, dovendo soddisfare le seguenti relazioni

$$W_{ji} \geq 0 \quad , \quad \sum_{j=1}^M W_{ji} = 1 \quad (\forall i).$$

La matrice di transizione $\{W_{ji}\}$ è in qualche modo il DNA della catena di Markov e ne determina tutte le sue proprietà.

Consideriamo l'evoluzione temporale del vettore $\mathbf{P}(t) = (P_1(t), P_2(t), \dots, P_M(t))$, ove $P_i(t)$ è la probabilità di essere al tempo t nello stato i ; è immediato scrivere

$$P_j(t+1) = \sum_{i=1}^M W_{ji} P_i(t) \tag{G.1}$$

¹ Così chiamate in onore del matematico russo Andrei Andreievich Markov(1856-1922) il primo a introdurre e studiare questo tipo di processi stocastici.

² Ci si potrebbe domandare perché non considerare anche i casi in cui il futuro al tempo $t+1$ dipende dagli stati al tempo t , quello a $t-1$, fino al tempo $t-r+1$. Si può vedere che questo tipo di processi è facilmente riconducibile al caso $r=1$, considerando una nuova catena in cui lo stato y_t è il vettore $(x_t, x_{t-1}, \dots, x_{t-r+1})$.

e quindi

$$\mathbf{P}(t) = \hat{W}^t \mathbf{P}(0) .$$

È naturale domandarsi se al crescere di t il vettore $\mathbf{P}(t)$ converge a un limite \mathbf{P}^* , se ciò accade allora

$$\mathbf{P}^* = \hat{W} \mathbf{P}^* , \tag{G.2}$$

è chiaro il significato di \mathbf{P}^* : le sue componenti sono le probabilità invarianti (o stazionarie).

Notare che può esistere una probabilità invariante \mathbf{P}^* che soddisfa la (G.2) anche se $\mathbf{P}(t)$ non converge verso un limite. Consideriamo un random walk su un anello con 4 siti in cui la probabilità di saltare a destra o a sinistra è $1/2$; è immediato verificare che $\mathbf{P}^* = (1/4, 1/4, 1/4, 1/4)$, tuttavia $\mathbf{P}(t)$ non tende a \mathbf{P}^* . Infatti se al tempo $t = 0$, $x_0 = 1$ con certezza cioè la probabilità è $(1, 0, 0, 0)$ allora è facile mostrare che al tempo $t = 1$ (e a tutti i tempi dispari) si ha $\mathbf{P}(t) = (0, 1/2, 0, 1/2)$, mentre ai tempi pari $\mathbf{P}(t) = (1/2, 0, 1/2, 0)$; il lettore è invitato a considerare il caso con un numero pari arbitrario di siti.

Il problema della convergenza a una probabilità stazionaria è, nell'ambito delle CM, l'analogo della domanda che s'era posto Boltzmann per la densità di probabilità. Nelle CM la risposta è molto più facile, vale il seguente teorema:

Teorema Appendice G.1. *Se esiste un intero n tale che per ogni coppia (i, j) si ha una probabilità non nulla di andare da j a i in n passi, cioè*

$$[W^n]_{ji} > 0 ,$$

allora esiste un'unica probabilità invariante e la convergenza è esponenzialmente veloce:

$$\mathbf{P}(t) = \hat{W}^t \mathbf{P}(0) = \mathbf{P}^* + O(|\alpha_2|^t) \rightarrow \mathbf{P}^* ,$$

dove α_2 è il secondo autovalore³ della matrice \hat{W} .

Tali CM sono ergodiche, cioè la media temporale lungo un random walk effettuato con le probabilità di transizione è asintoticamente uguale alla media con le probabilità $\{\mathbf{P}_i^*\}$.

Poiché $|\alpha_2| < 1$ scrivendo

$$|\alpha_2|^t = e^{-t \ln(1/|\alpha_2|)}$$

abbiamo che il tempo caratteristico è

$$\tau_c = \frac{1}{\ln(1/|\alpha_2|)} .$$

³ Per questo tipo di matrice è possibile dimostrare (teorema di Perron-Frobenius) che il primo autovalore è non degenere e si ha $\alpha_1 = 1 > |\alpha_2| \geq |\alpha_3| \dots$

Come esempio consideriamo una catena con due stati in cui

$$P_{1 \rightarrow 1} = 1 - p, P_{1 \rightarrow 2} = p, P_{2 \rightarrow 1} = p, P_{2 \rightarrow 2} = 1 - p,$$

usando la (G.1) si ha

$$P_1(t+1) = p + (1-2p)P_1(t).$$

La soluzione dell'equazione precedente è

$$P_1(t) = \frac{1}{2} + (1-2p)^t \left(P_1(0) - \frac{1}{2} \right) \rightarrow \frac{1}{2},$$

quindi $P_1^* = P_2^* = 1/2$ indipendentemente da p ; il tempo caratteristico è

$$\tau_c = -\frac{1}{\ln|2p-1|};$$

notare che $\tau_c \rightarrow \infty$ se $p \rightarrow 1$ oppure $p \rightarrow 0$, mentre $\tau_c \rightarrow 0$ quando $p \rightarrow 1/2$.

Catene di Markov reversibili

Una classe molto importante di CM sono le catene reversibili, cioè quelle per le quali la probabilità di una “traiettoria” $(i_1, i_2, \dots, i_{n-1}, i_n)$ è uguale a quella per la “traiettoria inversa” $(i_n, i_{n-1}, \dots, i_2, i_1)$. Questo accade se vale la condizione di *bilancio dettagliato*

$$P_i^* W_{ji} = P_j^* W_{ij}. \tag{G.3}$$

Il significato è intuitivo: per ogni coppia di stati (i, j) il flusso da i a j , $P_i^* P_{i \rightarrow j}$ è uguale al flusso da j a i , $P_j^* P_{j \rightarrow i}$.

È immediato verificare che se vale la (G.3) allora

$$Prob(i_1, i_2, \dots, i_{n-1}, i_n) = P_{i_1}^* P_{i_1 \rightarrow i_2} P_{i_2 \rightarrow i_3} \dots P_{i_{n-1} \rightarrow i_n} =$$

$$Prob(i_n, i_{n-1}, \dots, i_2, i_1) = P_{i_n}^* P_{i_n \rightarrow i_{n-1}} P_{i_{n-1} \rightarrow i_{n-2}} \dots P_{i_2 \rightarrow i_1}.$$

Per dare un'idea intuitiva del significato del bilancio dettagliato consideriamo il random walk su un anello con M siti con probabilità di transizione non simmetriche:

$$P_{n \rightarrow n \pm 1} = \frac{1}{2} \pm \Delta, \quad 0 \leq |\Delta| < \frac{1}{2}, \tag{G.4}$$

è facile vedere che se M è dispari, allora asintoticamente $\mathbf{P}(t)$ tende a $\mathbf{P}^* = (1/M, \dots, 1/M)$ indipendentemente dal valore di Δ . Si ha il bilancio dettagliato solo se $\Delta = 0$, mentre per $\Delta \neq 0$ si ha una direzione preferenziale (una corrente media non nulla).

Appendice H

Tempi di ricorrenza

Per completezza riportiamo il teorema di ricorrenza di Poincaré:

Teorema Appendice H.1. *Dato un sistema Hamiltoniano in uno spazio delle fasi limitato Ω , e un insieme $A \subset \Omega$, tutte le traiettorie che partono da $x \in A$ ritorneranno in A dopo un certo tempo e questo avverrà infinite volte, eccetto per condizioni iniziali in un insieme di misura nulla.*

La dimostrazione è piuttosto semplice. Indichiamo con \mathcal{S}^t l'operatore di evoluzione del sistema: se $x(0)$ è la condizione iniziale allora $x(t) = \mathcal{S}^t x(0)$. Indichiamo con $B_0 \subseteq A$ l'insieme dei punti che non ritornano in A . Ci sarà un tempo t_1 tale che $B_1 = \mathcal{S}^{t_1} B_0$ non si sovrappone con A e quindi $B_0 \cap B_1 = \emptyset$. In modo analogo ci saranno tempi $t_N > t_{N-1} > \dots > t_2 > t_1$ tali che $B_n \cap B_k = \emptyset$ per $n \neq k$ dove $B_n = \mathcal{S}^{(t_n - t_{n-1})} B_{n-1} = \mathcal{S}^{t_n} B_0$. Questo si capisce poiché assumendo $C = B_n \cap B_k \neq \emptyset$, per esempio per $n > k$, si ha una contraddizione con l'ipotesi che i punti in B_0 non ritornano in A . Gli insiemi $D_1 = \mathcal{S}^{-t_n} C$ e $D_2 = \mathcal{S}^{-t_k} C$, sono entrambi contenuti in B_0 , e D_2 può essere espresso come $D_2 = \mathcal{S}^{(t_n - t_k)} \mathcal{S}^{-t_n} C = \mathcal{S}^{(t_n - t_k)} D_1$, quindi i punti in D_1 ritornano in B_0 dopo un tempo $t_n - t_k$, in contrasto con l'ipotesi.

Consideriamo ora l'insieme $\bigcup_{n=1}^N B_n$, usando il fatto che gli insiemi $\{B_n\}$ non si sovrappongono e, per il teorema di Liouville, $\mu(B_n) = \mu(B_0)$, si ha¹

$$\mu\left(\bigcup_{n=1}^N B_n\right) = \sum_{n=1}^N \mu(B_n) = N\mu(B_0).$$

Poiché $\mu(\bigcup_{n=1}^N B_n)$ deve essere limitata (dal volume dello spazio delle fasi), e N può essere arbitrariamente grande, l'unica possibilità è che $\mu(B_0) = 0$.

Applicando il risultato dopo ogni ritorno in A si capisce che ogni traiettoria (con l'eccezione di un insieme di misura nulla) ritorna infinite volte in A .

Notiamo che nella dimostrazione abbiamo usato solo il teorema di Liouville, è quindi immediato capire che il teorema non vale solo per i sistemi Hamiltoniani ma in tutti quelli in cui si ha conservazione del volume dello spazio delle fasi. In

¹ In pratica $\mu(A)$ non è altro che il volume (nello spazio delle fasi) dell'insieme A .

effetti il teorema vale anche nei sistemi in cui non si ha la conservazione del volume dello spazio delle fasi, purché esista una misura invariante. La dimostrazione è ovvia: basta ripetere passo passo il procedimento usato per i sistemi Hamiltoniani semplicemente usando per la $\mu(\cdot)$ la misura invariante. Notare che per la validità del teorema non è necessario che il sistema sia caotico, basta l'ergodicità.

Lemma di Kac

Il teorema di ricorrenza non dice nulla su quanto tempo sia necessario per avere un ritorno. A questa domanda risponde il lemma di Kac:

Teorema Appendice H.2. *In un sistema ergodico, il tempo medio di ritorno in una regione A è dato da*

$$\langle \tau_A \rangle = \frac{\tau_c}{\mu(A)}, \quad (\text{H.1})$$

dove τ_c è un tempo caratteristico e $\mu(A)$ è la probabilità della regione A .

Il risultato precedente è una conseguenza diretta dell'ergodicità. Ne diamo una dimostrazione nel caso di una catena di Markov ergodica con probabilità invarianti $\{P_k^*\}$; in questo caso se al tempo 0 il sistema è nello stato m il tempo medio di ritorno è

$$\langle T_m \rangle = \frac{1}{P_m^*}. \quad (\text{H.2})$$

Indichiamo con τ_1 il tempo che il sistema impiega a tornare per la prima volta nello stato iniziale m , τ_2 il tempo necessario per il secondo ritorno nello stato iniziale etc. Dopo \mathcal{N} ritorni sarà passato un tempo

$$t_{\mathcal{N}} = \sum_{n=1}^{\mathcal{N}} \tau_n$$

ed il sistema è stato nello stato m la frazione di tempo

$$f_{\mathcal{N}} = \frac{\mathcal{N}}{t_{\mathcal{N}}}.$$

Per l'ergodicità si ha

$$\lim_{\mathcal{N} \rightarrow \infty} f_{\mathcal{N}} = P_m^*, \quad \lim_{\mathcal{N} \rightarrow \infty} \frac{t_{\mathcal{N}}}{\mathcal{N}} = \langle T_m \rangle,$$

e quindi si ottiene la (H.2).

Appendice I

Un tributo a Ludwig Boltzmann, il grande visionario

Ludwig Boltzmann (1844-1906) nacque a Vienna il 20 febbraio del 1844 nella notte tra martedì grasso e mercoledì delle ceneri; un fatto che lui scherzando riteneva all'origine della sua alternanza tra stati di entusiasmo e periodi di depressione che lo portarono infine al suicidio a Duino (vicino Trieste).

Su Boltzmann circola una leggenda infondata: la sua depressione sarebbe nata dal mancato riconoscimento accademico e la negativa accoglienza delle sue teorie. È vero che le sue idee vennero da alcuni (primo fra tutti Mach) aspramente contrastate. Falso che non fossero riconosciute come importanti: Boltzmann fu eletto membro delle principali accademie, ottenne lauree e dottorati ad honorem, inviti a conferenze, offerte di cattedre dalle più prestigiose università tedesche (all'epoca le più importanti del mondo).

Ludwig Boltzmann (LB) è un animo inquieto, uno scienziato generoso alla ricerca di una tranquillità che non riesce a trovare. Cambia più volte università: nel 1869 è professore a Graz, nel 1873 a Vienna, nel 1876 ancora a Graz, nel 1887 accetta una cattedra a Berlino (ma poi cambia idea e non prende servizio)¹, nel 1890 è professore a Monaco, nel 1894 torna a Vienna, nel 1900 è a Lipsia, nel 1902 ancora a Vienna: è un ritorno trionfale², addirittura titolare di due cattedre (una di fisica teorica l'altra di filosofia della scienza). Questa volta l'imperatore Francesco Giuseppe decisamente irritato pretende una lettera con un impegno formale a non emigrare nuovamente.

Si potrebbe pensare che Boltzmann fosse un opportunista a caccia di posizioni prestigiose e ben pagate, ma ci sono evidenze che fosse molto a disagio nel mondo accademico, poco dotato nell'arte dell'intrigo³ e sicuramente non particolarmente

¹ A quanto pare era spaventato dall'ambiente troppo formale di Berlino. Nella lettera di rinuncia sostenne, con notevole faccia tosta, che non si sentiva abbastanza preparato in matematica per ricoprire una cattedra di fisica teorica.

² La lezione inaugurale del corso di filosofia della scienza viene seguita da oltre 600 persone, addirittura la più grande aula dell'università non è sufficiente e molte rimangono in piedi, ne parlano tutti i giornali e riceve le congratulazione dell'imperatore.

³ In una lettera alla madre Boltzmann scrisse di *essere più bravo ad integrare* [le equazioni] *che ad intrigare*.

desideroso di onori, ad esempio rifiutò il titolo nobiliare offertogli dall'imperatore dicendo che *il nostro nome borghese è andato bene per i miei antenati e andrà bene per i miei figli e nipoti*.

Nonostante i tanti riconoscimenti ricevuti ed il grande successo dei suoi allievi⁴ Boltzmann, almeno nell'ultima parte della vita, aveva spesso la sensazione di essere isolato ed incompreso. È significativo quanto scrive nella prefazione delle *Lezioni sulla teoria dei gas* (il suo libro più importante): *Sono cosciente di essere solo un individuo che lotta debolmente contro la corrente del tempo. Ma posso sempre contribuire, in modo tale che quando la teoria dei gas sarà ripresa non ci sia troppo da essere riscoperto*.

Nonostante le avversità fu sempre molto combattivo, sono rimasti epocali i suoi scontri con gli antiatomisti; il più famoso fu nel 1895 a Lubecca al congresso degli scienziati di lingua tedesca. Così Sommerfeld a distanza di anni descrive lo scontro: *Il campione dell'energetismo era Helm; dietro di lui stava Ostwald e dietro entrambi la filosofia di Mach (che non era presente di persona). L'avversario era Boltzmann, assecondato da Felix Klein. La battaglia tra Boltzmann e Ostwald somigliava molto a un duello tra un toro ed uno smilzo torero. Tuttavia, in quella occasione il toro sconfisse il toreador ad onta dell'agilità di quest'ultimo. I ragionamenti di Boltzmann risultarono convincenti. Noi, giovani matematici, eravamo tutti dalla parte di Boltzmann. Pur amici Ostwald e Boltzmann continuarono la loro dura polemica con una serie di articoli, nel 1909 Ostwald finalmente riconoscerà che aveva torto, Mach non cambiò mai idea. Se fosse vissuto solo pochi anni in più LB avrebbe assistito al trionfo completo della sua ricerca*.

Boltzmann era in contatto epistolare con i principali fisici del suo tempo (come Helmholtz, Lorentz, Planck e Ostwald) e visitò più volte gli Stati Uniti. Del suo viaggio in California nel 1905 ha lasciato un divertente resoconto *Viaggio di un professore tedesco nell'Eldorado*. Una parte non piccola del libretto è dedicato al cibo ed alle bevande. Nella prima pagina racconta l'ultimo pranzo a Vienna prima della partenza, scrive *mangiai maiale arrosto, crauti e patate, con qualche bicchiere di birra ... La mia memoria con i numeri in genere è buona, ma nel contare i bicchieri di birra ho qualche difficoltà*. Ebbe qualche problema con la cucina californiana: in una cena ufficiale venne servita una pietanza che giudicò più adatta all'alimentazione delle oche, anzi *un'oca viennese non si sarebbe degnata di mangiarla*.

Boltzmann ammirava molto lo stile di vita americano, ben più libero e democratico di quello dell'impero asburgico, però non sopportava il puritanesimo *il vino viene nascosto, come fa uno studente col suo sigaro. Questo è quello che chiamano libertà*. Notò con grande disappunto che a chi chiedeva indicazioni per un negozio di vino venivano riservate espressioni di grande disapprovazione, come se avesse chiesto informazioni su "quelle certe signorine". Nella traversata di ritorno, lontano dalla puritana America, si consolò evitando l'acqua e bevendo solo vino bianco del Reno.

⁴ Tra cui: W. Nernst, S. Arrhenius (entrambi Nobel per la chimica) e P. Ehrenfest.

L'attività scientifica di LB fu in gran parte dedicata alla costruzione di una teoria statistica capace di descrivere i comportamenti dei sistemi macroscopici, ma ci furono anche altri risultati meno noti, e certo non trascurabili. Ripercorriamo brevemente i contributi di LB alla fisica.

I primi lavori di LB riguardavano l'elettromagnetismo. Fu tra i primi a comprendere l'importanza delle equazioni di Maxwell, in particolare la connessione tra ottica ed elettromagnetismo: tra le sue prime pubblicazioni ci sono dettagliati esperimenti sulla relazione tra la costante dielettrica e l'indice di rifrazione. Dopo i famosi esperimenti di H. Hertz sulle onde elettromagnetiche LB si appassionò così tanto che ripeté gli esperimenti, dedicò alcuni dei suoi corsi all'elettromagnetismo e pubblicò le sue lezioni (in due volumi nel 1891 e 1893).

Nel 1872 LB presenta quella che è universalmente nota come equazione di Boltzmann per l'evoluzione della densità di probabilità delle velocità di un gas rarefatto, ricavando in modo dinamico la distribuzione di Maxwell ed il celebre teorema H. In questo articolo per la prima volta appare un'equazione di evoluzione per una densità di probabilità; possiamo dire che è il primo esempio in cui la probabilità assume un ruolo veramente importante nelle scienze della natura⁵.

Boltzmann è stato uno degli iniziatori della "meccanica ereditaria", cioè lo studio dell'effetto della memoria della storia passata sulla deformazione di materiali come i vetri. Nel suo lavoro del 1874 LB introduce due funzioni $\phi(t - \tau)$ e $\psi(t - \tau)$ che tengono conto, al tempo t , dei contributi ereditari dovuti alla storia ai tempi precedenti $t - \tau$. Questi contributi compaiono in equazioni integrali che verranno poi trattate in modo più ampio all'inizio del XX secolo da Picard e Volterra.

Nel 1877 LB pubblica l'interpretazione probabilistica della termodinamica, la celebre relazione $S = k_B \ln W$.

Nel 1884 dimostra che l'energia totale della radiazione emessa da un corpo nero è proporzionale alla quarta potenza della temperatura, confermando la legge proposta dal suo maestro Josef Stefan sulla base di alcuni dati sperimentali. La dimostrazione è tecnicamente semplice (vedi dopo) ma è uno dei primi esempi di uso della termodinamica in sistemi diversi dai gas.

Il pensiero dominante dell'attività di Boltzmann fu la conciliazione tra la meccanica da una parte e la termodinamica dall'altra, cioè trovare un ponte, tecnico e concettuale, tra questi due mondi. Il percorso fu lungo e tortuoso, possiamo riassumere la grande visione di Boltzmann nei due punti seguenti:

- I) Introduzione di idee probabilistiche e loro interpretazione in termini fisici.
- II) Una relazione che fornisca un legame tra il mondo macroscopico (termodinamica) e quello microscopico (dinamica).

Il punto I) è estremamente delicato ed ancora oggi oggetto di studio, l'idea di Boltzmann fu di sostituire la media temporale con una media da calcolare con un'opportuna densità di probabilità. Questa congettura è chiamata *ipotesi ergodica*. Nel caso l'ipotesi sia soddisfatta allora la probabilità di una regione \mathcal{A} dello spazio delle fasi

⁵ L'equazione di Boltzmann, oltre a costituire uno dei capitoli più eleganti e profondi della fisica teorica, trova applicazioni in molti campi scientifici e tecnologici, dai semiconduttori alla dinamica dei gas rarefatti in ambito aereospaziale.

non è altro che la frazione di tempo che il sistema durante la sua evoluzione (su un tempo molto lungo) passa in \mathcal{A} .

La relazione che fornisce il ponte tra la termodinamica ed il mondo microscopico è incisa sulla tomba di Boltzmann:

$$S = k_B \ln W ,$$

ove S è l'entropia (una quantità termodinamica) del corpo macroscopico e W è il numero di stati microscopici (quindi una quantità di tipo meccanico) che realizzano la stessa configurazione macroscopica. La legge precedente è una delle grandi conquiste della scienza (al livello di $F = ma$ e $E = mc^2$) e racchiude in nuce gran parte della meccanica statistica.

L'approccio di Boltzmann è di tipo probabilistico, questo comporta che non c'è la completa certezza del secondo principio della termodinamica, che perde (almeno a livello concettuale) il suo status di legge assoluta, per diventare un fatto probabilistico. Dopo un lungo percorso costellato da controversie e discussioni, alcune animate da intenti costruttivi (come quella con l'amico Loschmidt), altre feroce-mente critiche (quelle con Zermelo, Ostwald e Mach), LB riesce ad intuire (dando anche qualche argomento) che la reversibilità meccanica non è affatto in contrasto con l'irreversibilità termodinamica. In particolare mostra che per i corpi macroscopici il teorema di ricorrenza di Poincaré praticamente non ha alcuna rilevanza a causa dell'enorme tempo necessario per avere la ricorrenza. Tanto per dare un'idea: per osservare un comportamento vietato dalla termodinamica, ad esempio una fluttuazione dell'ordine del 10% della densità in un centimetro cubo di aria, è necessario attendere qualcosa come $10^{10^{19}}$ anni, un intervallo temporale spaventosamente grande (l'età dell'universo è solo $O(10^{10})$ anni).

Boltzmann capì che nella conciliazione tra meccanica e termodinamica un ruolo importante è giocato dalle condizioni iniziali (cioè posizioni e velocità delle particelle del gas): quelle che portano ad un comportamento irreversibile sono molto più numerose di quelle dalle quali si ha un comportamento reversibile. In questo modo il secondo principio è una sorta di risultato probabilistico, che per i corpi macroscopici (a causa dell'enorme numero di particelle che contengono) è praticamente certo.

Boltzmann ebbe anche ampi interessi filosofici, e benchè pensatore non sistematico (il suo impegno principale era ovviamente per la fisica)⁶, alcuni aspetti del suo pensiero sono estremamente interessanti. Ad esempio anticipò l'idea di Kuhn sulle rivoluzioni scientifiche ed i paradigmi:

L'uomo della strada può immaginare che nuove nozioni e cause di fenomeni siano aggiunte gradualmente a quelle fondamentali già esistenti.... Questa opinione però è errata, e lo sviluppo della fisica teorica è stato sempre per salti...

⁶ Comunque negli ultimi anni della sua vita tenne anche il corso di filosofia della scienza che era stato di Mach.

A volte è stato frainteso, ad esempio Popper, che pur apprezzava Boltzmann, lo accusa di aver costruito teorie non falsificabili tacciandolo di idealismo. La critica è relativa al calcolo di Boltzmann sul tempo di ritorno $T_R \sim C^N \tau_0$ ove $C > 1$ e τ_0 è un tempo caratteristico. In un corpo macroscopico in cui $N \sim 10^{20}$ si ottengono tempi molto più lunghi dell'età dell'universo, quindi secondo Popper, saremmo di fronte ad una previsione di fatto non falsificabile. L'accusa è ingiusta ed errata: considerando sistemi piccoli con N dell'ordine di qualche decina si può studiare il problema al computer (ed è stato effettivamente fatto) trovando un perfetto accordo con le previsioni.

Purtroppo su Boltzmann esistono ricostruzioni storiche imprecise, ad esempio l'introduzione degli insiemi statistici è spesso attribuita a Gibbs, nella realtà fu un'idea di Boltzmann, alcune assolutamente fuorvianti. Prigogine e Stenger hanno ripetutamente sostenuto che le idee di Boltzmann erano incoerenti se non errate. Niente di più falso: le intuizioni di Boltzmann sono state puntualmente e sistematicamente confermate da una serie di dettagliati lavori di fisica matematica.

Bisogna dire che Boltzmann non si sforzava più di tanto per farsi capire, lo stesso Maxwell, che condivideva la sua impostazione filosofica, notava che *Studiando Boltzmann non sono riuscito a capirlo. Lui non mi capiva per la mia concisione, ma la sua prolissità era, ed è, un intoppo per me.*

Boltzmann scriveva articoli di 80 – 100 pagine con torrenti di equazioni, relegando in poche righe la discussione degli aspetti concettuali più importanti, alle critiche rispondeva che *l'eleganza è di pertinenza dei sarti e dei calzolai.* Comunque il suo stile ebbe anche estimatori: il giovane Albert Einstein, studente del Politecnico di Zurigo, all'inizio del novecento scriveva alla fidanzata dicendo che aveva letto Boltzmann e lo trovava magnifico. Nonostante il suo apprezzamento anche Einstein era ben cosciente delle difficoltà dei lavori di Boltzmann e riconosceva che anche grandi fisici non erano stati in grado di capirlo.

La legge di Stefan-Boltzmann

Riteniamo utile e interessante riportare infine la derivazione della legge di Stefan-Boltzmann, che richiede solamente l'uso dei due principi della termodinamica e della legge per pressione di radiazione fornita dall'elettromagnetismo. Essa costituì un notevole successo della termodinamica, mostrando quanto grande fosse l'ambito di applicabilità dei suoi principi.

Indichiamo con e l'energia per unità di volume del campo elettromagnetico in equilibrio a temperatura T , all'interno di una cavità di volume V ,

$$e = \frac{E}{V} = e(T)$$

e ricordiamo che la pressione di radiazione, nota nell'ambito dell'elettromagnetismo, può essere scritta nella forma

$$P = \frac{1}{3} e .$$

Dal primo principio della termodinamica nella forma

$$dQ = dE + PdV = d(eV) + PdV ,$$

usando la pressione di radiazione, possiamo ricavare dQ in termini di V e T :

$$dQ = \frac{4}{3} e dV + V \frac{de}{dT} dT ,$$

ottenendo per l'entropia

$$dS(T, V) = \frac{dQ}{T} = \frac{4}{3} \frac{e}{T} dV + \frac{V}{T} \frac{de}{dT} dT .$$

Dato che, per il secondo principio della termodinamica, dS è un differenziale esatto, deve essere

$$\frac{\partial}{\partial V} \left(\frac{V}{T} \frac{de}{dT} \right) = \frac{4}{3} \frac{\partial}{\partial T} \left(\frac{e}{T} \right) ,$$

da cui

$$\frac{de}{dT} = 4 \frac{e}{T} ,$$

che può essere facilmente risolta ottenendo la legge di Stefan-Boltzmann:

$$e = C \times T^4 , \tag{I.1}$$

dove C è una costante il cui valore non può essere determinato in ambito classico. La (I.1) è valida anche in ambito quantistico e la costante C davanti a T^4 può essere calcolata (solo) usando la statistica di Bose-Einstein:

$$C = \frac{8\pi^5 k_B^4}{15h^3 c^3} . \tag{I.2}$$

Lecture consigliate per le appendici

Un'introduzione al calcolo delle probabilità espressamente pensata per studenti di fisica:

- G. Boffetta e A. Vulpiani *Probabilità in Fisica* (Springer-Verlag Italia, 2012).

I lavori di A. Einstein sul moto browniano sono stati tradotti in inglese e raccolti nel volume

- A. Einstein *Investigations on the Theory of the Brownian Motion* (Dover Publ., 1956).

Per la figura di Boltzmann, fisico e filosofo, e la descrizione dell'ambiente in cui è vissuto, un libro bellissimo e fonte inesauribile di informazioni è

- C. Cercignani *Ludwig Boltzmann: the man who trusted atoms* (Oxford University Press 1998).

Un'ottima selezione di scritti divulgativi di Boltzmann, che fornisce un quadro abbastanza completo della sua visione filosofica è

- L. Boltzmann *Modelli matematici, fisica e filosofia* (Bollati Boringhieri 1999).

Per una presentazione (a volte storicamente infedele, ma sempre interessante) delle idee di Boltzmann in meccanica statistica

- P. Ehrenfest and T. Ehrenfest *The conceptual foundation of the statistical approach in mechanics* (Cornell University Press, New York 1956, l'edizione originale in tedesco è apparsa nel 1911).

Indice analitico

- algoritmo di Verlet, 48
- barotropica
 - formula, 80, 107
- binomiale
 - distribuzione, 97
- Boltzmann
 - distribuzione di Maxwell-B., 17, 40–42, 57
 - equazione di, 16, 17, 53–61, 85
 - legge di Stefan-B., 1
 - limite di B.-Grad, 17
 - principio di B.-Einstein, 83
- Browniano
 - moto, 75, 81, 82, 87, 107
- cammino libero medio, 84
- caos
 - deterministico, 65
 - molecolare, 17, 55, 60, 62, 64, 66
- catena di Markov, 49, 61, 117–119, 122
- coefficiente di diffusione, *vedi* diffusione,
 - coefficiente di
- correlazione
 - funzione di, 49, 50, 111
- Cramer
 - funzione di, 101
- densità di probabilità
 - empirica, 17, 42
 - evoluzione di una, 43–46
 - invariante, 65
 - marginale, 38
 - stazionaria, 44–46
- deviazioni
 - grandi, 101–103
- diffusione
 - coefficiente di, 79, 84, 85, 87, 108, 111
- distribuzione
 - stazionaria, 57
- Ehrenfest
 - modello di, 61–65
- Einstein
 - principio di Boltzmann-E., 83
 - relazione di E.-Smoluchowski, 78, 81
 - teoria del moto Browniano di, 107–108
 - teoria delle fluttuazioni di, 83
- emissione termoionica, 72–74
- energia libera, 39
- ensembles*, 43–46, 64
- entropia, 29–33, 36–37, 57, 128
 - relativa, 98, 100
- ergodico
 - ipotesi ergodica, 12, 13, 19–28, 33, 46, 125
- Fermi, Pasta e Ulam (FPU)
 - esperimento numerico di, 46, 50
- fluttuazioni
 - teoria di Einstein delle, 83
- funzione
 - caratteristica, 101
 - Gamma di Eulero, 24
- Helmholtz
 - teorema di, 31–33
- Kac
 - lemma di, 11, 23, 59, 64
- KAM
 - teorema, 46
- Khinchin, 26, 28, 46
- Kubo
 - relazione di Taylor-K., 112
- Lanford

- teorema di, 17
- Langevin
 - e il moto Browniano, 77
- lemma di Kac, *vedi* Kac, lemma di
- limite centrale
 - teorema del, 7, 97, 101, 103
- linee spettrali
 - allargamento delle, 71
- Liouville
 - equazione di, 45
 - operatore di, 45
 - teorema di, 15, 105, 106, 121
- Loschmidt, 3, 57, 59
- Lyapunov
 - disuguaglianza di, 27
 - esponente di, 15
- Markov
 - catena di, 49, 61, 117–119, 122
 - disuguaglianza di, 101
- matrice di transizione, 117
- Maxwell
 - distribuzione di M.-Boltzmann, *vedi* Boltzmann, distribuzione di Maxwell-B.
- modello di Ehrenfest, *vedi* Ehrenfest, modello di
- Monte Carlo
 - metodo, 46, 49–50
- numeri
 - legge dei grandi, 7–9, 17, 27, 42, 98, 100, 101, 103
- paradosso
 - della reversibilità, *vedi* reversibilità, paradosso della
 - della ricorrenza, *vedi* ricorrenza, paradosso della
- Perrin
 - e il moto Browniano, 80–81
- Perron-Frobenius
 - operatore di, 46
 - teorema di, 118
- Poincaré, 15, 58, 76
- potenziale di interazione, 20, 55, 70, 74–75, 86–87
- probabilità
 - stazionaria, 61, 118
- probabilità di transizione, 49, 117, 119
- relazione
 - di Einstein-Smoluchowski, 78, 81
 - di Taylor-Kubo, 112
- reversibile
 - catena di Markov, 119
- reversibilità
 - paradosso della, 3, 14, 57, 59
- ricorrenza
 - paradosso della, 3, 11, 14, 57, 59
 - tempo di, 11, 12, 24, 59, 64, 121
 - teorema di, 3, 122, 126
- Schwartz
 - disuguaglianza di, 27
- simulazioni numeriche, 46–50
- Smoluchowski
 - relazione di Einstein-S., *vedi* relazione, di Einstein-Smoluchowski
- Stefan
 - legge di S.-Boltzmann, 1, 125, 127–128
- Stirling
 - approssimazione di, 24, 30, 41, 97
- Taylor
 - relazione di T.-Kubo, 112
- teorema
 - del limite centrale, 7, 97–99, 101, 103
 - di Chernoff, 102
 - di fluttuazione-dissipazione, 83
 - di fluttuazione-risposta, 83
 - di Helmholtz, 31–33
 - di Lanford, 17
 - ergodico, 22
 - H, 55–57, 66
- Tonks
 - gas di, 114–115
- transizione
 - probabilità di, 49, 117, 119
- trasformazione canonica, 48
- Verlet
 - algoritmo di, 48
- viriale
 - coefficienti del, 74, 86
 - sviluppo del, 86, 113–114
- viscosità, 77, 85, 87, 107
- Zermelo, 3, 57, 59

“Finito di stampare nel mesh di settembre 2015”