PARTE I: PLANCK EINSTEIN E POINCARÉ, KUBO E FERMI

Prefazione alla prima parte

In questa parte ci occupiamo della prima fase di creazione della meccanica quantistica, con la crisi del principio di equipartizione dell'energia, e la sofferta soluzione mediante la quantizzazione dell'oscillatore armonico, che fu introdotta da Planck tra il 19 ottobre e il 14 dicembre 1900 per "spiegare" la legge del corpo nero, e fu poi utillizzata da Einstein nel 1907 per "spiegare" il calore specifico dei solidi.

In questo ambito, un risultato chiaro e definitivo sulla linea del *classical program* di Einstein sarebbe il calcolo classico del calore specifico in funzione della temperatura per un solido concreto, ad esempio il Fluoruro di Litio, che riproducesse in maniera molto buona i risultati sperimentali, analogamente a quanto è avvenuto per lo spettro dei cristalli ionici nell'infrarosso e sarà illustrato nella seconda parte. Questo risultato sul calore specifico non è stato ancora ottenuto. Tuttavia si sono avuti notevoli progressi. Anzitutto a livello numerico, con il classico lavoro di Fermi, Pasta ed Ulam (FPU) del 1954, in cui fu messo in luce che a basse temperature le dinamica classica, presentando moti non completamente caotici, potrebbe non garantire l'applicabilitá degli ordinari metodi della meccanica statistica. Poi a livello analitico, con la dimostrazione che moti ordinati aventi rilevanza per la meccanica statistica, come quelli osservati nel lavoro FPU, esistono anche al limite termodinamico (ovvero al limite di infinite particelle con fissata energia specifica e densità), contro il diffuso pregiudizio che al limite termodinamico possano esistere solo moti caotici. Questi fatti lasciano ancora aperta la domanda se sia proprio vero che la meccanica classica fallisca qualitativamente nel determinare le proprietà termodinamiche del corpo nero, e quelle dei solidi a basse temperature.

Dopo avere illustrato il problema FPU, ripercorreremo lo sviluppo storico del problema dell'equipartizione dell'energia, attraverso i lavori classici di Boltzmann, di Planck e di Einstein. In particolare ricorderemo come le prime critiche

alla validità del principio di equipartizione nella fisica classica. o almeno al suo significato, risalgano a Boltzmann stesso. Poi daremo un'esposizione critica di un fondamentale lavoro di Poincaré del 1912 nel quale sarebbe stata provato che la quantizzazione è necessaria se si vuole ottenere con metodi statistici la legge di Planck per l'energia media di un oscillatore. Risulta che tale necessità riposa su una ipotesi la cui validità sembra essere almeno dubbia, sulla base del grande risultato di Kolmogorov del 1954. Nella presente versione delle note non è ancora stato possibile illustrare questo punto, e viene data soltanto una esposizione abbastanza dettagliata del procedimento statistico seguito da Poincaré, basandosi sulla sua ipotesi (ci ripromettiamo di ritornare presto su questo punto). Comunque, risulta che il teorema di Poincaré può essere letto secondo l'atteggiamento generale che Einstein aveva già proposto nel suo contributo alla conferenza Solvay del 1911. Ovvero: se per semplicità di descrizione si ammette valida l'ipotesi di Poincaré, allora la legge di Planck può ottenersi solo quantizzando l'energia del sistema. In tale senso, con le parole anticipate da Einstein in un contesto analogo, *"tutto va come se l'energia del sistema fosse quantizzata"*. La quantizzaione sarebbe dunque un modo comodo per rappresentare teoricamente le osservazioni, descrivibili però, almeno in linea di principio, anche in modo classico. Illustreremo infine i metodi di Green–Kubo in relazione alla formula meccanico–statistica del calore specifico, la quale conferma come, in una descrizione classica, l'esistenza di moti ordinati diminuisca il valore del calore specifico, e quindi anche quello dell'energia interna termodinamica, rispetto a quello previsto dal principio di equipartizione.

Nota biografica su Planck. In quel fatidico volgere del secolo, nell'anno 1900 in cui fondò la meccanica quantistica, Planck aveva 42 anni, essendo nato a Kiel il 23 aprile 1858. La vita di Planck fu accompagnata da tragedie. Egli aveva quattro figli. Il primo gli morì da militare nella prima guerra mondiale. Poi venivano due gemelle. Una morì di parto, e la seconda ebbe la stessa sorte, dopo avere sposato il cognato. L'ultimo figlio, con cui condivideva la passione per il piano e l'organo, fu contattato dai cospiratori in occasione del putsch contro Hitler. A seguito del fallimento del putsch, fu impiccato.

Einstein su Planck. L'unico grande fisico veramente vicino a Planck nell'atteggiamento scientifico fu (oltre a Nernst) Einstein. Questi aveva di lui un stima grandissima, come testimoniato dall'obituario che egli scrisse in occasione della morte di Planck (avvenuta a Göttingen il 4 ottobre 1947), e che qui riproduciamo: [1]

"*Quante varietà di stili nel tempio della scienza! E come diversi sono gli uomini che lo frequentano e diverse le forze morali che ve li hanno condotti. Più di uno si dedica alla scienza con la gioia di rendersi conto delle proprie superiori facoltà intellettuali: per lui la scienza è lo sport preferito che gli permette di vivere una vita intensa e di appagare le sue ambizioni. Ve ne sono poi anche molti, i quali unicamente a scopo utilitario vogliono portare le loro offerte all'attività del cervello.*

[1] A. Einstein, *Come io vedo il mondo*, Newton Compton, pag 36. Citato da P. Campogalliani.

Basterebbe che un angelo divino scacciasse dal tempio gli uomini di queste due categorie e l'edificio rimarrebbe vuoto in modo inquietante, se non ci restassero ancora alcuni uomini del passato e del presente: di questo numero fa parte il nostro Planck ed è questa la ragione per cui lo amiamo."

Capitolo 1

Fermi, Pasta ed Ulam (FPU), 1954: critica su basi dinamiche del principio di equipartizione

Abbiamo sottolineato nella prefazione generale come nelle presenti note si intenda ripercorrere i punti di crisi della della fisica classica che hanno condotto alla fondazione della MQ (meccanica quantistica), considerarandoli come argomenti di ricerca attuale. L'idea centrale è che in quei punti critici i problemi matematici che si presentano in una trattazione classica sono così formidabili, che al momento non è chiaro quali siano esattamente le predizioni classiche. È chiaro che se si riuscisse a mostrare che le previsioni classiche sono molto più vicine ai risultati sperimentali di quanto comunemente si ritiene, l'intero problema delle relazioni tra meccanica classica e MQ andrebbe rivisto.

L'esempio del lavoro FPU è quanto mai adatto per illustrare questo modo di procedere. Qui si tratta di decidere se sia giustificato applicare il principio di equipartizione dell'energia alle basse temperature. Un principio che sembrerebbe seguire dalle leggi della fisica classica, ma invece è in palese contraddizione a tutte le temperature con la legge del corpo nero, e a basse temeperature con i dati sperimentali sui calori specifici delle molecole poliatomiche, e dei solidi. Quest'ultimo problema è quello che sostanzialmente avevano di mira Fermi, Pasta ed Ulam nel 1954.[1]

[1]Una rassegna recente sul problem FPU può trovarsi nel lavoro A. Carati, L. Galgani, A. Maiocchi, F. Gangemi, R. Gangemi, *The FPU Problem as a Statistical-mechanical Counterpart of the KAM Problem, and its Relevance for the Foundations of Physics*, pubblicato sulla rivista moscovita Regular and Chaotic Dynamics, 23 (6), 704-719 (2018); doi: 10.1134/S1560354718060060.

1.1 Premessa: Il lavoro FPU come lavoro attuale di ricerca sulle relazioni tra MQ e fisica classica: la meccanica statistica e il teorema KAM

La ragione per cui Fermi giunse a compiere questo celebre lavoro FPU, forse l'ultimo lavoro prima della sua morte, è che, al pari di Einstein, anche egli non amava l'interpretazione ortodossa della MQ. Questo fatto è poco noto, e vale la pena di sottolinearlo. A testimonianza di questo atteggiamento di Fermi esiste anzitutto il lavoro FPU che ora illustreremo (oltre ai tre lavori del 1923 cui accenneremo). Esiste però anche una testimonianza personale del suo allievo Emilio Segrè. Si tratta di una conversazione che questi ebbe a Berkeley attorno al 1975 con il più anziano dei presenti autori, a proposito del lavoro FPU e dei risultati allora recenti sulle proprietà qualitative dei sistemi dinamici classici riguardanti la coesistenza di moti ordinati e moti caotici (come il lavoro di Kolmogorov del 1954, quelli di Contopoulos (1963), e quello di Hénon ed Heiles del 1964). In tale conversazione Segrè disse esplicitamente che Fermi, al pari di Einstein, non amava l'interpretazione ortodossa. Però, a causa del suo carattere notoriamente molto schivo che lo faceva rifuggire da discussioni di tipo, diciamo così, ideologico,[2] non amava discuterne pubblicamente, e lo confidava soltanto ai suoi amici intimi.[3]

Dunque non dovrebbe sembrare strano che nel suo ultimo lavoro Fermi abbia toccato proprio il punto cruciale nel quale venne per la prima volta introdotta la quantizzazione il 14 dicembra 1900, ovvero la crisi del principio di equipartizione, in relazione all'energia media di un sistema di oscillatori armonici.

Infatti, lungo tutta la sua vita scientifica Fermi ebbe sempre ben fisso nella mente il fatto che il procedimento usato da Boltzmann per dedurre la sua celebre distribuzione di probabilità, (la legge di MB o Maxwell–Boltzmann), da cui segue come un corollario il teorema di equipartizione dell'energia, richiede una giustificazione dinamica. Questa giustificazione dinamica costituisce quello che comunemente è noto come **il problema ergodico**, problema che era stato posto inizialmente da Boltzmann stesso

Ribadiamolo ancora. Il procedimento con cui si deduce la distribuzione di MB (che illustreremo nel prossimo capitolo) ha carattere sostanzialmente cinematico, perché consiste nel determinare la probabilità di un evento mediante un opportuno conteggio coinvolgente tutte le regioni accessibili dello spazio delle fasi, senza tener conto del fatto che in tale spazio si presentino delle barriere, dovute ad esempio all'esistenza di costanti del moto diverse dall'energia del sistema. Questo procedimento è giustificato se il sistema possiede una opportuna

[2] *"Solo problemi concreti potevano dargli una immediata sensazione dell'importanza dei risultati raggiunti"*, da E. Segrè, Nota biografica su Fermi, pag. XXVII.

[3] Esiste tuttavia una testimonianza scritta. Si tratta del suo manoscritto, dal titolo *Notes on quantum mechanics*, delle lezioni tenute a Chicago nel 1954 sulla MQ. Nelle prime pagine, dopo avere introdotto l'equazione di Schrödinger, parlando della interpretazione di Born la qualifica, in maniera apparentemente strana, come *"provisional"*.

proprietà di tipo dinamico (l'ergodicità. appunto), che può essere posseduta da un sistema oppure no. Il problema ergodico consiste appunto nello stabilire se un sistema dinamico abbia tale proprietà, oppure no. Più precisamente, un sistema hamiltoniano è ergodico se su ogni superficie di energia fissata quasi tutte le orbite sono ovunque dense, e inoltre la frazione di tempo che ciascuna di esse passa in ogni regione è proporzionale alla misura di quella regione (almeno se si considerano tempi sufficientemente lunghi). È chiaro allora che la distribuzione di MB non sia giustificata se vi sono nello spazio delle fasi dei vincoli dinamici, delle barriere, che impediscano al sistema di accedere a tutte le regioni di uguale energia. Concretamente, una condizione necessaria per l'ergodicità di un sistema è che esso non ammetta integrali del moto (o costanti del moto) oltre all'energia.[4]

Ora, era stato dimostrato da Poincaré (attorno al 1990) che *in generale* (in un senso che richiederebbe una accurata precisazione) in un sistema hamiltoniano non esistono costanti del moto diverse dall'energia, e dunque in generale i sistemi hamiltoniani sarebbero ergodici. Su questo problema Fermi scrisse ben tre lavori nel 1923, mentre si muoveva tra Roma e Göttingen.

Essi sono riprodotti nel primo volume delle sue opere, con una prefazione di Segrè a pag. 79. Queste sono le parole di Fermi, quando comincia a ricordare il risultato di Poincaré: *"Poincaré ha dimostrato che un sistema di equazioni canoniche normali, oltre all'integrale della conservazione dell'energia, non può in generale ammettere altro integrale primo analitico uniforme e indipendente dal tempo"*.[5] Dunque non esisterebbero altre costanti del moto, oltre l'energia. Si noti come questo risultato sia in un certo senso paradossale. Perché in una gran parte dei sistemi di interesse per la fisica, tra cui (lo vedremo poco più avanti) il modello FPU, si ha a che fare con sistemi che sono, come si dice, perturbazioni di sistemi integrabili, cioè perturbazioni di sistemi che nel caso imperturbato ammettono tante costanti del moto indipendenti quanti sono i gradi di libertà. Il paradosso consiste dunque nel fatto che l'accensione di una perturbazione, piccola quanto si vuole, ridurrebbe il numero di constanti del moto da N, arbitrariamente grande, a solo uno. Sembrerebbe sparire ogni continuità, nonostante che si stiano considerando modelli assolutamente "lisci" (*smooth*), cioè si lavori in ambito analitico (ovvero con sistemi dinamici definiti da una hamiltoniana analitica).

Questo paradosso venne eliminato dal grandissimo lavoro di Kolmogorov del 1954, in cui egli diede un celebre teorema noto come teorema KAM (Komogorov, Arnold e Moser), in cui viene aggiunto il nome di due celebri matematici (uno russo e l'altro tedesco) che ne produssero due nuove dimostrazione nel 1961. Si noti che Moser ammise privatamente che egli credeva che la dimostrazione originale di Kolmogorov fosse sbagliata, e qualche dubbio si trova anche tra le pieghe del lavoro di Arnold. Mentre invece la dimostrazione di Kolmogorov è

[4]Stiamo escludendo i casi in cui esistano costanti del moto come il momento totale o il momento angolare totale. Queste sono costanti del moto per un sistema isolato, ma non lo sono ad esempio per un gas racchiuso in una scatola, a causa delle collisioni delle molecole con le pareti.

[5]J.H. Poincaré, *Méthodes nouvelles de la mécanique celeste*, vol. I, cap. 5.

assolutamente corretta, e addirittura più semplice delle due successive.[6]

Qui occorre ricordare che ad ogni costante del moto (diciamo $F = F(z)$ se z denota le coordinate di una regione dello spazio dellle fasi) corrisponde tutta una famiglia di superfici invarianti (superfici aventi la proprietà che, se contengono un punto, allora contengono anche tutta l'orbita che da esso si origina), definite ciascuna in forma implicita da $F(z) = c$ per un fissato valore di c. Ebbene, già Fermi aveva prodotto un lavoro matematico in cui generalizzava il risultato di Poincaré, nel senso che non solo non esistono altre costanti del moto (e quindi non esistono le famiglie di corrispondenti superfici invarianti), ma addirittura non esiste neppure *una singola* superficie invariante che non sia una superficie di energia costante. Nelle sue parole: *"Ci proponiamo in questo lavoro di generalizzare tale teorema dimostrando che ... all'infuori delle F = cost."* (F denota la funzione hamiltoniana, cioè l'energia) *"non esiste in generale* **neanche una sola** *ipersuperficie ... ".*[7]

Il punto sottile, che permette di fare la connessione con il teorema KAM, è che nel teorema di Poincaré, come in quello di Fermi, si ricercano superfici invarianti singole, o famiglie di tali superfici, definite come "superfici di livello" di una funzione $F(x)$ analitica. Invece le singole superfici invarianti di Kolmogorov non sono definite da una equazione $F(x) = c$ attraverso una funzione F analitica. La funzione può avere al più una regolarità di tipo C^∞ anziché analitico. Tuttavia Kolmogorov recupera una continuità in misura per l'insieme costituito da tali supe4rfici, perché egli aggiunge che tale insieme. pur essendo "strano", con un carattere analogo a quello del celebre insieme "bucato" di Cantor, ha invece una misura relativa che diventa la misura piena (cioè che tende ad 1) quando la perturbazione tende a zero. Dunque si ottiene una continuità "in misura", perché nel caso imperturbato la misura è piena.

Abbiamo voluto ricordare questi aspetti analitici per mettere in luce la complessità matematica del problema delle costanti del moto. Quello che ci preme sottolineare é come ci debba dunque attendere che una analoga complessità si presenti anche nella cossispondente meccanica statistica di sistemi siffatti. Più concretamente: attendersi che in tali sitruazioni un sistema dinmico debba essere trattato dal punto di vista della meccanica statistica mediante il procedimento

[6]Questo fatto stranissimo, si spiega forse con la eccezioanle novità dei concetti coinvolti. Comunque, tutta la comunità scientifica riconosce ora la assoluta e completa priorità del lavoro di Kolmogorov, dopo che esso venne riproposto, nella versione originale e solo completato con abbastanza banali calcoli al livello di studenti del terzo anno di università, nel lavoro G. Benettin, L. Galgani, A. Giorgilli, J.-M. Strelcyn, *A Proof of Kolmogorov Theorem on Invariant Tori using Canonical Transformations Defined by the Lie Method*, Nuovo Cimento B79, 201 (1984).

[7]I lavori di feermi sono di non facile lettura. Una critica riguardante una applicazione alla teoria ergodica fu rivolta da Urbanski, e Fermi subito ne riconobbe la fondatezza. (pag. 97 del primo volume delle opere). Una critica venne fatta anche dal noto matematico C.L. Siegel (il maestro di J. Moser) nel 1956 (due anni dopo la morte di Fermi). Ma a una successiva revisione compiuta nel 1982 (si veda G. Benettin, G. Ferrari, L. Galgani, A. Giorgilli, *An extension of the Poincaré–Fermi theorem on the nonexistence of invariant manifolds in nearly integrable Hamiltonian systems*, Nuovo Cim. **72B**, 137 (1982)), il risultato di Fermi appare corretto, anzi è stato addirittura generalizzato.

di Gibbs sembrerebbe, al minimo, ingenuo. Questo spiega la ragione del titolo di un lavoro citato piú sopra, ovvero: *The FPU Problem as a Statistical-mechanical Counterpart of the KAM Problem, and its Relevance for the Foundations of Physics*.

Qui ci basta avere segnalato che esiste ancora oggi un problema aperto (quale sia la meccanica statistica da usarsi per sistemi dinamici come quello FPU), e inoltre che esisteva un interesse profondissimo di Fermi per tale problema, avendo intuito la profondità degli aspetti matematici coinvolti.. Si noti bene che a Fermi la matematica importava solo per il contributo che può dare alla fisica.[8] Quindi, vogliamo sottolinearlo, tanto interesse in Fermi per questo problema ergodico non era dovuto alla "bellezza" del problema matematico, ma alla sua profonda rilevanza fisica. E tale rilevanza fisica consiste nel poter decidere se sia davvero giustificato il principio di equipartizione, nel quale sembrava fallire la fisica classica, particolarmente alle basse temperature.

Questo è dunque, ne siamo certi, il motivo per il quale, quando dopo la guerra ebbe per la prima volta a disposizione il computer di Los Alamos per poterlo utilizzare a scopi puramente scientifici, Fermi ne approfittò immediatamente per controllare, aldilà di tutti i complicati teoremi matematici, se veramente la meccanica classica conduca alla proprietà di equipartzione dell'energia. E questo doveva essere controllato in una maniera che fosse la più semplice e diretta possibile, che è proprio quello che egli fece, come ora illustriamo.

1.2 Il modello FPU, e i corrispondenti modi normali di oscillazione

Il problema di interesse è dunque la distribuzione di probabilità dell'energia di un oscillatore armonico di data frequenza, e il suo valor medio. Naturalmente, l'energia di un singolo oscillatore isolato, che non interagisca con altri sistemi, è una costante del moto, non dipende dal tempo, mantiene il valore che le è stato dato inizialmente. Abbiamo invece un problema statistico, indotto dalla dinamica, quando si ha un sistema di oscillatori interagenti mutuamente o con altri sistemi. Il caso più semplice è quello in cui si ha un sistema di N oscillatori, con N grande, in mutua interazione, sicché in generale l'energia del singolo oscillatore varierà nel tempo, al punto che si possa addirittura presumere che non esista nessuna costante del moto oltre l'energia totale del sistema. In tal caso la distribuzione di probabilità dell'energia di ogni singolo oscillatore sarà determinata dalla dinamica, una volta assegnati i dati iniziali (posizioni e velocità) di tutti i singoli oscillatori. I dati iniziali conterranno implicitamente la temperatura attraverso

[8] E questo, nonostante il fatto che egli fosse dotato anche per la matematica. Si veda la nota a pag 167 del libro di Levi Civita *The absolute differential calculus*, Dover (New York, 1970). Riportiamo anche da E. Segrè, Nota biografica su Fermi, pag. XXV: *"Anche se ostentava talvolta un certo disprezzo per alcune forme di matematica troppo formali o poco creative . . . , se occorreva Fermi sapeva dare stretto rigore matematico a una qualunque dimostrazione, e spesso accadeva che su un esempio facesse vedere tutte le raffinatezze della critica, ma poi le abbandonava per procedere più speditamente nello svolgere una teoria"*.

il valore dell'energia totale iniziale (che resta costante durante il moto), anche se, in effetti, la relazione tra energia totale e temperatura è tutt'altro che ovvia. Su questo ritorneremo in seguito.

Il modello, descritto in termini di particelle

L'esempio considerato da Fermi et al. è quello del più semplice modello monodimensionale di solido, che è anche il modello discretizzato di corda vibrante che fu studiato per la prima volta da Lagrange a Torino nel 1759 (almeno nell'approssimazione lineare).[9] Si hanno $N+2$ particelle su una retta, con le posizioni x_0, x_{N+1} delle particelle estreme tenute fisse, mentre le posizioni $x_1, x_2, \dots x_N$ delle altre sono le coordinate libere del sistema. Le interazioni sono "a primi vicini", ovvero possono pensarsi come dovute a molle che collegano due punti adiacenti, essendo le molle nonlineari, della forma che diremo subito. Si ha dunque, in opportune unità, una hamiltoniana del tipo

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2} + \sum_{i=1}^{N+1} V(x_i - x_{i-1}) \,,$$

dove l'energia potenziale di ogni molla ha la forma

$$V(y) = \frac{y^2}{2} + \alpha \frac{y^3}{3} + \beta \frac{y^4}{4}$$

con due opportuni parametri α, β. Ad esempio, nel caso $\beta = 0$ le equazioni di moto hanno la forma (per $i = 1, \dots, N$)

$$\ddot{x}_i = x_{i+1} + x_{i-1} - 2x_i + \alpha\big[(x_{i+1} - x_i)^2 - (x_i - x_{i-1})^2\big] \,.$$

Passaggio ai modi normali di oscillazione

Questo per quanto riguarda le particelle costituenti il solido. Veniamo ora agli oscillatori armonici accoppiati. Fin dai tempi di Lagrange era noto che nel caso di forze lineari ($\alpha = \beta = 0$), passando ad opportune nuove variabili $a_1, \dots a_N$, il sistema risulta essere equivalente a un sistema di N oscillatori armonici indipendenti (ovvero disaccoppiati). Si ha dunque una hamiltoniana

$$H_0 = \sum_{j=1}^{N} \frac{1}{2} \left[\dot{a}_j^2 + \omega_j^2 a_j{}^2 \right]$$

con equazioni di moto

$$\ddot{a}_j + \omega_j^2 a_j = 0 \,, \tag{1.2.1}$$

[9] E. Fermi, J. Pasta, S. Ulam, *Studies of non linear problema*, in E. Fermi, *Note e memorie*, lavoro n. 266, vol. II, pag. 977.

e certe frequenze

$$\omega_j = 2 \sin \frac{\pi j}{2(N+1)} \ .$$

Per dimostrare questa cosiddetta *diagonalizzazione del problema*, ovvero, questo *passaggio ai modi normali di oscillazione*, basta eseguire il cambiamento di variabili lineare

$$a_j = \sum_k x_k \sin \frac{\pi k j}{N+1} \ . \tag{1.2.2}$$

Nell'approssimazione lineare $\alpha = \beta = 0$ le equazioni sono duqnue banali, perché ogni oscillatore si muove per conto proprio, e dunque tutte le singole energie E_j sono delle costanti del moto. Ma le nonlinearità inducono un accoppiamento tra i modi normali di oscillazione, sicché le singole energie non sono più delle costanti delle moto, e si potrebbe forse presumere che si abbia una strana proprietà di discontinuità per cui l'energia totale resti l'unica costante del moto, *comunque piccolo sia l'accoppiamento tra gli oscillatori*. Ma vedremo che questo, in effetti, non avviene.

Ora, le equazioni nonlineari del moto per le ampiezze a_j dei modi normali sono alquanto complicate, perché mentre le particelle agiscono ciascuna solo sulle due adiacenti, avviene invece che le ampiezze dei modi si accoppiano "tutte con tutte". In termini delle ampiezze dei modi le equazioni del moto, che avevano la forma (1.2.1) nel caso noninteragente (o imperturbato) $\alpha = \beta = 0$, nel caso ineragente prendono una forma del tipo

$$\ddot{a}_j = -\omega_j^2 a_j + \text{termini di ordine superiore} \ . \tag{1.2.3}$$

Il modo piú semplice per studiare il sistema è di integrare numericamente, con opportuni metodi, le equazioni di moto per le particelle, e compiere di tanto in tanto la trasformazione (1.2.2) che fornisce le ampiezze a_j dei modi normali, le corrispondente velocità e quindi anche le corrispondenti energie

$$E_j = \frac{1}{2}\left[\dot{a}_j^2 + \omega_j^2 {a_j}^2\right] \ .$$

L'energia totale come parametro perturbativo rilevante, e la complessità del problema ergodico. Il teorema di Kolmogorov

Bisogna dunque tenere conto della perturbazione indotta dalle nonlinearità sulle equazioni del moto per le ampiezze degli oscillatori. A tal fine occorre tener presente il fatto che, se l'energia totale è nulla, allora tutte le ampiezze a_j e le corrispondenti velocità $\dot{a}_j$ sono nulle; nel caso di energia nonnulla tali quantità a_j ed $\dot{a}_j$ tendono comunque a zero quando l'energia totale tende a zero. Dunque, al diminuire dell'energia totale in generale la perturbazione, dovuta a termini quadratici o cubici nelle equazioni di moto per le ampiezze, diventa sempre più trascurabile rispetto al termine lineare. In conclusione, l'energia totale agisce da

parametro perturbativo, perché il sistema si riduce sempre più a un sistema di oscillatori disacoppiati quanto più viene diminuita l'energia totale. Ma il sistema disaccoppiato è costituito da N oscillatori indipendenti, e pertanto possiede N costanti del moto, le singole energie E_j dei modi normali di oscillazione.

Si capisce allora quanto sia ingenua la presunzione che il sistema possa essere ergodico (nessuna costante del moto, oltre l'energia totale) per ogni valore arbitrariamente piccolo dell'energia. Si avrebbe infatti una stranissima discontinuità: N costanti del moto per il sistema imperturbato, nessuna costante del moto (oltre l'energia totale) comunque piccola sia la perturbazione. Ci si dovrebbe attendere che si incontri invece una qualche sorta di continuità, e abbiamo giá ricordato come una continuità in misura sia stata dimostrata con il teorema KAM. Questa breve descrizione dovrebbe fare intuire quali siano le difficoltà concettuali e tecniche del problema ergodico. Invece, la soluzione trovata nel lavoro di FPU fu molto semplice ed illuminante.

1.3 La scoperta di FPU: apparente nonequipartizione dell'energia a basse energie secondo la fisica classica, e il problema della caoticità dei moti

Premessa: il problema delle medie temporali

Per andare al cuore del problema nella maniera più semplice possibile. ci resta ancora da richiamare come mai entrino in gioco delle *medie temporali*.

Sappiamo tutti (e lo rivedremo nel prossimo capitolo) che esiste il principio di equipartizione, che appare come un teorema della meccanica classica se il sistema studiato può essre descritto con la statistica di Maxwell–Boltzmann. Secondo tale principio, in un sistema di oscillatori debolmente accoppiati di fissata energia totale E, tutti gli oscillatori dovrebbero avere la medesima energia media

$$\overline{E_j} \simeq \frac{E}{N} , \qquad j = 1, \ldots, N .$$

Inoltre, con un ulteriore argomento, su cui ritorneremo in seguito, si introduce l'identificazione della temperatura mediante la relazione

$$\frac{E}{N} = k_B T ,$$

sicché il principio di equipartizione, nella sua forma che coinvolge la temperatura, e che chiameremo "estesa", si esprime nella forma

$$\overline{E_j} = k_B T .$$

Dunque occorre precisare anzitutto cosa si deve intendere per energia media. Fin dai tempi di Boltzmann, e poi quelli successivi di Einstein e di Poincaré, tutti

erano d'accordo che si debba fare riferimento alla media temporale. Consideriamo un sistema dinamico in uno spazio delle fasi, e denotiamo con z un punto generico di tale spazio, e con $\Phi^t(z)$ il suo evoluto temporale al tempo t secondo la assegnata dinamica. Allora per le energie E_j degli oscillatori, come per ogni altra variabile dinamica, la media temporale *fino al* tempo t è definita

$$\overline{E_j} = \frac{1}{t}\int_0^t E_j(\Phi^s(z))\,\mathrm{d}s \ .$$

Come si vede, la media temporale dipende dal tempo t di osservazione, e dal punto iniziale z. Se il sistema è ergodico, la media temporale ammette limite per $t \to \infty$ e il limite non dipende dal punto iniziale z. Naturalmente, non è affatto chiaro cosa si debba intendere fisicamente per "limite infinito del tempo di osservazione". È questo un punto cruciale, riguardante il cosiddetto *tempo di rilassamento*, che studieremo nell'ultimo capitolo di questa prima parte delle note. Solitamente si ammette che dopo un tempo opportuno (tempo di rilassamento) la media temporale diventi costante, non dipenda più dal tempo.

Ricordiamo tuttavia che era già stato fatto presente da Boltzmann, ripetutamente, che in generale si possano presentare rilassamenti ad equilibri parziali su scale di tempo aventi ordini di grandezza completamente diversi. Questo fu messo in evidenza anche da Poincaré in un bellissimo lavoro,[10] sostanzialmente sconosciuto, in cui studia degli interessanti esempi e parla esplicitamente di *tempi grandi di primo ordine* e *tempi grandi di secondo ordine*.

L'altro elemento (relativo al punto iniziale z) è veramente cruciale, ed è quello che riguarda la possibile esistenza di altre costanti del moto o di altre superfici invarianti: il valore di equilibrio (quello su cui si è eventualmente stabilizzata la media temporale di una generica grandezza) non deve dipendere dal punto iniziale z. Qui si sottintende che si tratta sempre di punti di una certa superficie di energia, relativa ad un fissato valore dell'energia totale E.

Nonequipartizione a basse energie (FPU, 1954) vs equipartizione ad alte energie (Izrailev e Chirikov, 1966)

Detto tutto questo, il principio di equipartizione prevede che le medie temporali di tutte le energie E_j, oltre ad essersi rilassate nel tempo ed avere un valore indipendente dal punto iniziale z, siano anche uguali tra di loro. Per questo si parla di equipartizione (o forse, meglio, di equiripartzione). L'energia (in media temporale) non dipende dal modo normale, in particolare dalla sua frequenza. Si ha democrazia, tutte le frequenze hanno lo stesso *share* di energia.

Possiamo ora andare al cuore del problema, e questo lo facciamo rinunciando per un momento a riportare i fatti come si sono succeduti in ordine storico, mescolando il contributo di FPU e quello successivo, del 1966, di Chirikov (o di

[10] H. Poincaré, *Reflections sur la théorie cinétique des gas*, *J. Phys. Théor. Appl.* 5, 369–403 (1906). in *Oeuvres*. Vol. IX, pag. 586.

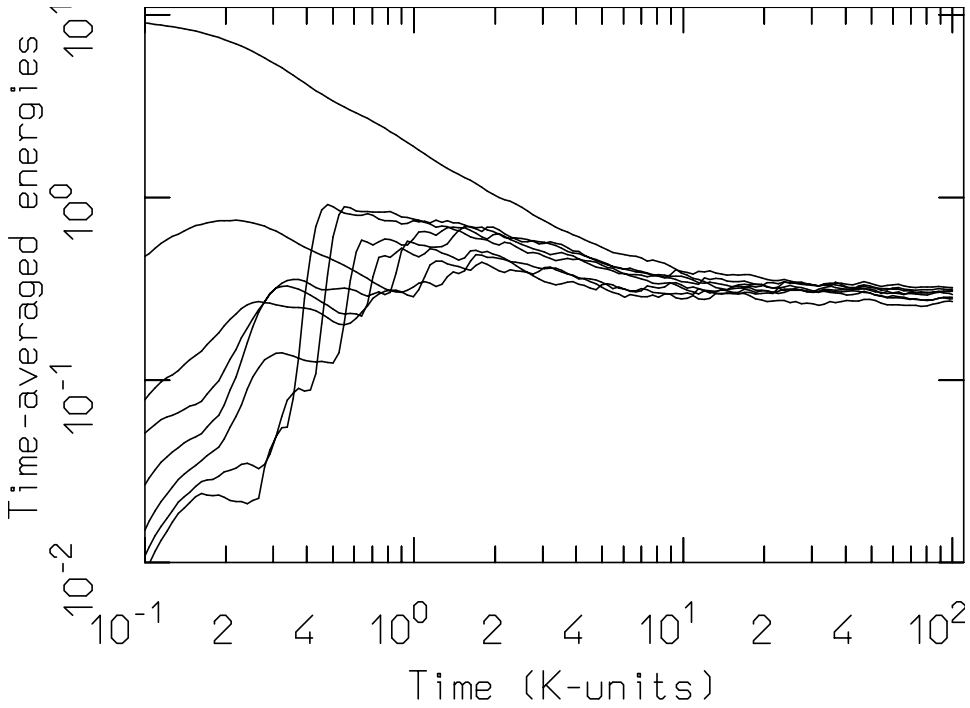

Figura 1.1: Equipartizione delle energie dei modi (in media temoprale) ad alte energie

Izrailev e Chirikov). Guardiamo le due figure (1.1) e (1.2) qui sotto.[11] Entrambe si riferiscono a un sistema FPU con $N = 32$ e valori dei parametri $\alpha = \beta = 1/4$, ma a due diversi valori dell'energia totale E (e dunque del parametro perturbativo): precisamente $E = 10$ nella prima ed $E = 0.05$ (un valore duecento volte più piccolo) nella seconda. Il punto iniziale z in entrambi i casi corrisponde ad assegnare tutta l'energia iniziale solo al primo modo normale (quello con la frequenza minima), con velocità $\dot{a}_j$ tutte nulle. Entrambe le figure riportano, in funzione del tempo (in certe unità che qui non ci importano), le medie temporali delle energie $\overline{E_j}$ di 8 fra i 32 modi normali del sistema.

Nella prima figura, ad energia alta, si vede che, al passare del tempo, tutte le energie medie, partendo dai valori iniziali considerati (nulli per tutti i modi tranne il primo), vengono poi a formare un fascio sempre più stretto, fino a rilassare tutte ad un valore sostanzialmente unico (il valore di equipartizione, appunto). Si potrebbe mostrare che il risultato finale non cambia se si prende un'altra condizione iniziale alla stessa energia, ad esempio tutta l'energia assegnata solo al modo più alto (quello di massima frequenza).

[11] Si veda A. Carati, L. Galgani, A. Giorgilli, *Dynamical Systems and Thermodynamics*, in *Encyclopedia of Mathematical Physics*, Elsevier (Oxford, 2006).

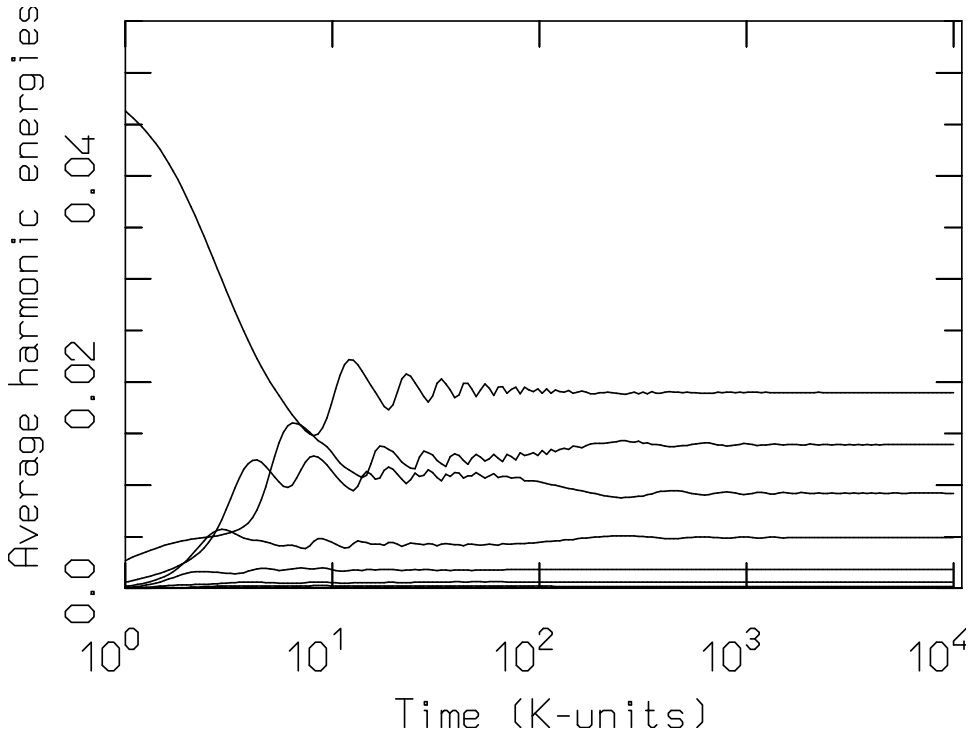

Figura 1.2: Nonequipartizione a basse energie

Invece, nella seconda figura, corrispondente a un valore duecento volte più piccolo dell'energia, le cose sono completamente diverse. Ogni energia va per conto proprio. L'energia, data inizialmente solo al primo modo, passa anche agli altri, ma poco: si potrebbe vedere che le energie degli altri modi decrescono esponenzialmente al crescere della frequenza. La cosa interessantissima è che ciononostante si ha ancora un rilassamento, perché le medie temporali rilassano tutte a valori costanti entro il tempo di osservazione, anche se ciascuna a un valore diverso. È stato dunque raggiunto uno stato di equilibrio (o forse di metaequilibrio? si veda l'ultimo capitolo di questa prima parte), che tuttavia non è più lo stato di MB. Questa è la grande scoperta di Fermi e compagni. A energia sufficientemente bassa esiste ancora una statistica classica, perché si è avuto un rilassamento, ma non è più quella di MB. Nessuno è ancora in grado di descriverla, ma certo non è quella che conduce alla legge di Dulong e Petit per il calore specifico dei solidi. Nelle parole di Ulam, riportate nella sua prefazione al lavoro FPU nei *Collected Papers* di Fermi: *"The results of the computations were interesting and quite surprising to Fermi. He expressed the opinion that they really constituted a little discovery in providing intimations that the prevalent beliefs in the universality of mixing and thermalization in nonlinear systems may not be always justified."*

1.4 Il contributo di Chirikov: la congettura sulla soglia di caoticità al limite termodinamico

Torniamo ora per un momento alla parte storica: la seconda figura riportata sopra è sostanzialmente l'ultima figura del lavoro originale FPU. In molte precedenti figure tali autori riportano l'andamento dei valori istantanei delle energie dei vari modi in funzione del tempo, e solo nell'ultima sono riportate le medie temporali. Ma l'ultima figura del lavoro FPU, sostanzialmente quella qui riportata, è certamente il cuore del loro risultato.

Il lavoro, inizialmente, non ebbe molto seguito, se non per gli studi di Kruskal, Zabusky ed altri, attorno all'anno 1966, in cui essi presero spunto dal lavoro FPU per fare interessantissime digressioni sull'esistenza di solitoni, che aprirono la via a moltissimi lavori riguardanti la dinamica dei sistemi hamiltoniani ad infiniti gradi di libertà, che è uno dei temi di ricerca oggi più studiati in matematica (equazioni di Korteweg de Vries ecc.)

L'intervento di Chirikov: la soglia di caoticità

Il problema fisico che stava a cuore a Fermi venne invece ripreso, come problema rilevante, dal fisico russo Boris Chirikov. Egli si rese subito conto dell'interesse grandissimo del lavoro, per il fatto che esso sembrava sconvolgere le convinzioni di tutta la comunità scientifica (ma non di Einstein e di Fermi) sulle relazioni tra MQ e fisica classica. L'equipartzione non sarebbe una conseguenza della fisica classica a basse energie.

Di fronte a tale situazione Chirikov fece anzitutto anch'egli una rilevante scoperta, complemenatare a quella di FPU, e inoltre pervenne a formulare una congettura che permettesse di sistemare le cose nella maniera tradizionale che tutti desideravano, in modo cioè che il paradosso FPU (così potremmo chiamarlo) venisse eliminato.

La scoperta di Chirikov è implicitamente quella che abbiamo illustrato nella prima figura. Il fatto è che FPU avevano condotto i loro calcoli per un solo, ben fissato, valore dell'energia. Noi siamo certi che Fermi, se fosse vissuto più a lungo, avrebbe provveduto egli stesso a colmare questa lacuna. Il grande merito della nuova scoperta, dunque, va tutto a Chirikov e ai suoi collaboratori, in particolare a Izrailev, che scrisse con lui il lavoro del 1966. Sostanzialmente Chirikov comprese che il paradosso FPU deve scomparire per energie iniziali sufficientemente alte. Diciamo che, avendo fissato il tipo di condizioni iniziali, ad esempio tutta l'energia iniziale data al primo modo, deve esistere una soglia di energia, sopra la quale si ha equipartizione (entro tempi abbastanza brevi), mentre non si ha equipartizione per energie inferiori. La prima figura sopra riportata mostra che questo è sostanzialmente vero. Ad una energia 200 volte più alta di quella relativa alla prima figura, abbastanza rapidamente si ha equipartizione.

La congettura di Chirikov

Si presenta allora tutta una serie di problemi su come dipende la soglia dalla forma dei dati iniziali. Ma di questo qui non ci occupiamo, e veniamo alla grande idea di Chirikov. L'osservazione cruciale è che per la termodinamica statistica non importano le quantità cosiddette estensive, come l'energia totale E del sistema, ma solo quantità intensive, come l'energia specifica E/N. Importa quello che viene chiamato il *limite termodinamico*, cioè il limite $N \to \infty$ in cui però si tengano fissati i valori specifici di energia e volume, E/N, V/N.

Dunque, avendo stabilito che per un sistema di N oscillatori esiste una energia di soglia o energia critica E^c, la quantità di interesse è la corrispondente energia specifica E^c/N. Chirikov allora formula la congettura che

$$E^c/N \to 0 \quad \text{per} \quad N \to \infty \ ,$$

ovvero che nel limite termodinamico il sistema sia sostanzialmente sempre caotico. In questo modo, da grande fisico teorico, egli da una parte comprende il significato profondo dell'esistenza del paradosso FPU, e dall'altra si inventa un modo in cui possa avvenire che tale paradosso sia fisicamente irrilevante. In tale modo tutti possono mettersi il cuore in pace, e dare per assodato che, dove nacque la MQ con i suoi nuovi metodi inconciliabili con i principi classici, veramente la fisica classica fallisce.

1.5 Il dibattito sulla soglia di Chirikov. La nuova teoria perturbativa in senso statistico, valida nel limite termodinamico

Chirikov e i suoi collaboratori diedero alcune indicazioni, sia numeriche che semianalitiche, che sembravano confermare la suddetta congettura. Nel 1971 intervenne la scuola di Milano e Pavia, con un lavoro di Bocchieri, Scotti e Loinger[12] in cui si davano fortissime indicazioni numeriche a favore della persistenza della soglia in energia specifica, che appariva sostanzialmente non cambiare per N che cresce da 10 fino a 100.

Subito dopo venne suggerito esplicitamente che il paradosso FPU potesse indicare la possibiltà di una comprensione classica della legge di Planck,[13] cosa che ebbe una certa risonanza. Intervennero poi altri studiosi da Padova, Roma, Firenze, da Berkeley e dal Brasile, e si accese un ampio dibattico. Ma i risultati sembravano talvolta supportare la persistenza della soglia al crescere di N e altre volte venivano interpretati come indicanti il contrario.

Ci si doveva allora rivolgere a studi analitici, utilizzando i metodi classici della teoria delle perturbazioni, perché la persistenza di moti ordinati al crescere

[12] Bocchieri, Scotti, Bearzi, Loinger
[13] Galgani Scotti, Cercignani Galgani Scotti,

del parametro perturbativo è, potrebbe dirsi, il cuore stesso della teoria perturbativa. Una teoria, questa, che svolge "da sempre" un ruolo fondamentale nella meccanica celeste e in meccanica analitica, e ha avuto un ruolo rilevante anche per i fondamenti della meccanica quantistica, negli anni di poco precedenti il 1925. Ad esempio nel circolo di Göttingen il capo (Born) aveva affidato al giovane Heisenberg il compito di studiare il classicissimo libro di Poincaré in tre volumi *Les méthodes nouvelles de la mécanique celeste*. Nell'ambito di problemi del tipo FPU, dopo la familiarizzazione della comunità scientifica con i progressi di Kolmogorov e quelli successivi culminati nel teorema di Nekhoroshev,[14] il massimo sforzo di utilizzare metodi di tipo classico venne compiuto da D. Bambusi.[15] Egli fu in grado di ottenere un risultato nel limite $N \to \infty$, ma al costo di dovere tenere finita l'energia totale E, ovvero avere energia specifica E/N nulla. Dunque i metodi classici della teoria delle perturbazioni, nella forma fino ad allora conosciuta, non si applicano al limite termodinamico. Questo fatto fu comunemente interpretato come una indicazioone che in tale limite la soglia specifica E^c/N di Chirikov sarebbe nulla. Al limite termodinamico si avrebbero dunque solo moti caotici, e quindi equipartizione, e sarebbe così confermato il fallimento della fisica classica. Tuttavia, in un colloquio svoltosi a Mosca nel 1974 tra Kolmogorov (nel suo appartamento all'Università) e Galgani e Scotti, Kolmogorov disse esplicitamente che non vedeva alcun ostacolo alla permanena di moti ordinati al limite termodinamico. E questo fatto fu sempre tenuto presente nel gruppo milanese. Occorreva compiere un nuovo progresso qualitativo nella teoria perturbativa.

Il passo significativo venne in effetti compiuto, e fu il seguente. Fu compreso che si possano applicare i metodi perturbativi al limite termodinamico, non nel senso tradizionale della meccanica celeste, ma in senso statistico. Il punto è che nei metodi tradizionali si mira a controllare i dati iniziali uniformemente in tutto lo spazio delle fasi (ad esempio a fissata energia). Ma era ben noto che esistono dei dati inziali molto speciali in cui le perturbazioni sono tanto grandi da non permettere delle stime uniformi al crescere di N. La nuova idea fu allora di ricercare un risultato più debole, che fosse però significativo per la meccanica statistica. Si tratta di rinunciare a controllare tutti i dati iniziali, controllando solo il comportamento di opportuni valori medi, calcolati ad esempio usando la misura di Gibbs. In tal modo fu possibile ottenere risultati che valgono, in energia specifica, uniformemente in N per grandi N (e quindi anche nel limite $N \to \infty$), e garantiscono una forma di stabilità dei moti su tempi finiti. Ad esempio, nel senso che fino a quei tempi restano diverse da zero le autocorrelazioni temporali di significative funzioni.[16] Sulla rilevanza di tali autocorrelazioni temporali torneremo nell'ultimo capitolo di questa prima parte. Successivamente, tale metodo perturbativo di tipo statistico fu applicato a vari modelli simili al

[14] Nekhoroshev, Benettin Galgani Giorgilli, Galgani Giorgilli Martinoli Vanzini
[15] D. Bambusi
[16] A- Carati, *J. Stat. Phys.* **128**, 1057 (2007).

modello FPU. e al modello FPU stesso.[17]

1.6 La fase attuale. Rilassamento all'equipartizione per tempi sufficientemente lunghi, ma non ad uno stato di equilibrio termico. Deduzione della della termodinamica come problema aperto

Diamo qui un cenno rapidissimo ai risultati degli ultimi anni. Il primo risultato è che il problema di Chirikov è ormai superato, perché viene infine accettato che l'equipartizione nel classico problema FPU viene raggiunta dopo tempi sufficientemente lunghi, che dipendono dall'energia specifica E/N e divergono all'annullarsi di E/N. Questo fu visto dapprima a Milano, e in seguito confermato in maniera molto precisa a Padova.[18]

Dunque nel modello FPU si raggiunge, dopo un tempo sufficiente, uno stato di apparente equilibrio, con equipartizione dell'energia. D'altra parte, come vedremo più avanti, l'equipartizione dell'energia, la cui crisi ha dato inizio alla MQ con la legge di Planck, è una condizione necessaria ma non sufficiente perché la corretta statistica indotta dalla dinamica coincida, in tale stato di apparente equilibrio, con quella di equilibrio (o di Gibbs). Ci si confronta qui con il problema ergodico e con il problema ad esso connesso di stabilire quale sia la statistica indotta dalla dinamica. Infatti, in un recentissimo lavoro è stato mostrato[19] che il modello realistico di tipo FPU che descrive il cristallo ionico di LiF, citato nella prefazione generale e illustrato nella seconda parte di queste note,[20] non è ergodico almeno fino a scale di tempi geologiche. In particolare non è giustificata la consueta identificazione meccanico–statistica della temperatura, almeno a basse temperature. e quindi non è noto come si debba formulare la termodinamica statistica. [21] Ad esempio, in un prossimo capitolo dedicato al lavoro di Poincaré del 1912 verrà illustrato come Nernst (l'inventore del terzo principio della termodinamica) concepisca una situazione dinamica non completamente caotica in cui (sembrerebbe incredibile) la legge di Planck risulta essere compatibile con il principio di equipartizione.[22]

[17] De Roeck e Huveneers, Maiocchi Bambusi Carati, Giorgilli Paleari Penati, Maiocchi

[18] L. Berchialla, L. Galgani, A. Giorgilli, *Localization of energy in FPU chains*, Discrete Continuous Dynam. systems A 11, 855-866 (2004); G. Benettin, A. Ponno, *Time scales to equipartition ion the FPU problem*, J. Stat. Phys. **144**, 793 (2011).

[19] A. Carati, L. Galgani, F. Gangemi, R. Gangemi, *Relaxation times and ergodic properties in a realistic ionic–crystal model, and the modern form of the FPU problem*, Physica A (2019).

[20] A. Carati, L. Galgani. A. Maiocchi. F. Gangemi, R. Gangemi, *Classical infrared spectra of ionic crystals and their relevance for statistical mechanics*, Physica A**506**, 1 (2018).

[21] Su questo problema molto delicato potrebbe avere una notevole influenza il tipo di modello considerato, che dovrebbe tenere conto delle azioni elettromagnetiche, addirittura ritardate, e della nota forza di reazione di radiazione (si veda la seconda parte).

[22] Tutto questo è legato al fatto che Nernst riesce a concepire un modo quantitativo per distinguere tra energia posseduta da un sistema e energia che il sistema può scambiare. Solo la seconda

Sembra dunque che si presenti una nuova riformulazione del " paradosso FPU", e che, dopo più di sessant'anni, il problema FPU risulti ancora aperto. Mentre il problema FPU in senso stretto consiste nello stabililre se il sistema evolve ad uno stato di equipartizione dell'energia partendo da uno "stato iniziale alla Fermi" (solo pochi modi eccitati), il problema FPU in senso lato consiste nello stabilire se dopo il rilassamento all'equipartizione la dinamica garantisca l'applicabilità della statistica classica di equilibrio.

Nota biografica su Fermi. Enrico Fermi nacque a Roma il 17 Settembre 1901 e morì a Chicago il 29 Novembre 1954, poco dopo avere compiuto il 53–esimo anno.*" Nel 1938 le leggi razziali del fascismo, che colpivano la moglie Laura Capon e i figli, e che ripugnavano al suo senso morale, lo indussero ad emigrare negli Stati Uniti d'America"*. Così si esprime Emilio Segrè nella bellissima nota biografica scritta all'inizio del primo volume delle Note e Memorie (Collected Papers), dove giustamente dà il seguente commento riassuntivo (pag. XLV): *"Con lui si spense l'ultimo fisico universale nella tradizione dei grandi del secolo XIX, quando era possibile a una persona sola raggiungere i culmini più alti della teoria e dell'esperimento e dominare tutti i campi della fisica "*. Tra i fatti curiosi della vita di Fermi citati da Segrè riportiamo il seguente (significativo): *"A dieci anni il fatto che il cerchio fosse rappresentato dall'equazione $x^2 + y^2 = r^2$ gli aveva dato moltissimo da pensare e creato serie difficoltà"*.

costituirebbe l'energia interna termodinamica.

Capitolo 2

Boltzmann e la termodinamica statistica: la crisi del principio di equipartizione

Scopo di questo capitolo è di illustrare in che modo i metodi statistici intesi a fondare la termodinamica statistica, già anticipati dalle teorie cinetiche di Clausius, sono stati poi fondati da Boltzmann. Dal punto di vista fisico, il risultato fondamentale da lui ottenuto è il teorema di equipartizione (o, forse meglio. di equiripartizione) dell'energia: tipicamente, all'equilibrio gli oscillatori armonici hanno tutti la medesima energia, indipendentemente dalla loro frequenza. D'altra parte questa proprietà fallisce qualitativamente nella fenomenolgia del corpo nero e in quella del calore specifico dei solidi, che sono appunto i due problemi che condussero Planck ed Einstein alla primitiva introduzione della quantizzazione, nel 1900 e nel 1907. Tuttavia metteremo in luce come il procedimento di Boltzmann, che apparentemente ha lo stesso carattere puramente combinatorio che si presenta nel familiare trattamento del gioco dei dadi o dei lanci di monete, in realtà richieda invece una profonda giustificazione dinamica. Questo fu ripetutamente ed intensamente discusso da Boltzmann stesso, che introdusse a tale fine quello che egli stesso chiamò *il problema ergodico*. In tal modo egli mostrò come i suoi metodi combinatori siano significativi solo per il caso di gas (e non per le molecole, con i loro gradi di libertà interni), e addirittura nel caso in cui essi siano molto diluiti e ad alta temperatura. Risulta poi che la giustificazioe dinamica dei metodi statistici, anche quelli successivamente formulati da Gibbs, è ancora a tutt'oggi un problema completamente aperto.

2.1 Intermezzo: La rilevanza del principio di equipartizione per il programma riduzionista nelle scienze della natura

Nel capitolo precedente abbiamo mostrato come non sia chiaro a tutt'oggi se, per un determinato sistema fisico, siano applicabili, a bassa temperatura, i metodi della meccanica statistica classica. In particolare no è noto se la temperatura possa essere identificata con un multiplo dell'energia cinetica, come sembra richiedere il principio di equipartizione. Non abbiamo però spiegato perché tale principio fosse così rilevante già nella cultura della fine dell'800.

Tutto nasce dalla scoperta di Clausius, per cui è possibile (almeno nel caso dei gas) interpretare in maniera meccanicistica la temperatura, come energia cinetica proporzionale alla e nergia cinetica media dei baricentri delle molecole che compongono un corpo. Si tratta di un passo fondamentale, sentito come risolutivo, nell'implementazione del programma riduzionista, del programma cioè che si propone di ridurre tutti i fenomeni naturali al puro moto meccanico di enti elementari. Ora, non solo la temperatura è la prima grandezza non meccanica che si incontra studiando i fenomeni naturali, ma è anche la più importante dal punto di vista filosofico. Infatti, era già stato capito, per merito di Ostwald, che la chimica era "spiegata" dalla termodinamica, nel senso che le reazioni chimiche preocedono nel verso in cui l'energia libera diminuisce.[1] Dunque, se la biochimica spiega la biologia, anche quest'ultima, in ultima analisi, viene ricondotta alla termodinamica. Allora la scoperta di Clausius, riconducendo la termodinamica alla meccanica, può ben dirsi essere la chiave di volta di un programma meccanicistico generale che abbraccia tutta la scienza.

Ora, il risultato di Clausius contiene già in sé un primitivo principio di equipartizione, perché l'energia dei baricentri delle varie molecole viene ad essere in media la stessa, pari appunto alla temperatura del sistema, indipendentemente dal tipo di molecola (la sua massa, il numero degli atomi che la compongono, ecc.). Non dice però nulla sul valore dell'energia media totale delle molecole,[2] quantità anch'essa fondamentale per determinare l'energia interna U del sistema, e quindi le altre funzioni di stato termodinamiche. L'ingegnosa via intrapresa da Boltzmann per determinare queste quantità culminò con il famoso legame tra entropia e logaritmo della probabilità[3] e con il principio di equipartizione generale.

In un certo senso, questo fu l'acme del programma meccanicistico, che subito dopo piombò in una crisi profondissima, con la presa di coscienza, da parte

[1] Anzi, gli energetisti (ovvero gli antiatomisti) avevano addirittura proposto con Duhem di sostituire il principio fondamentale della dinamica, ovvero il principio di minima azione, con il principio di minimizzazione dell'energia libera. La meccanica sarebbe dunque, nell'ambito della termodinamica. il caso ideale di processi puramente adiabatici.

[2] Ricordiamo il teorema di König, che afferma che l'energia cinetica di un sistema è la somma dell'energia cinetica del baricentro e dell'energia cinetica del sistema rispetto al baricentro. Inoltre, vi é anche l'energia di interazione tra i costituenti della molecola.

[3] È questa la formula scritta come epitaffio sulla sua tomba.

della comunità dei fisici, che esso era in palese contraddizione con molti fatti sperimentali. Fatto questo che cercheremo di illustrare nel seguito del presente capitolo. Ne nacque un lungo travaglio che portò alla introduzione della qualtizzazione da parte di Planck ed Einstein e quindi, tra il 1925 e il 1930, alla nascita di una nuova meccanica (la MQ) con paradigmi completamente differenti da quelli della vecchia meccanica newtoniana. La nozione stessa di meccanicismo sembra essere in qualche modo svanita in tale contesto. Non a caso Heisenberg stesso si è sempre proclamato idealista, per quanto riguarda il suo atteggiamento filosofico. Tuttavia, ironicamente, per quanto riguarda la cultura ufficiale, che pur si dichiara scientista, sembra prevalere un generico atteggiamento meccanicistico, e nulla sembra essere cambiato dai giorni di Clausius, Maxwell e Darwin.

Ora, l'atteggiamento alla Einstein, che cerchiamo di "implementare" facendo riferimento a orbite di particelle, sembrerebbe addirittura avere un carattere iperrealistico. Si deve tuttavia tener presente che la concezione che perseguiamo deve necessariamente fare riferimento non solo alla teoria di Wheeler r Feynman, la cui trattazione dinamica richiede la conoscenza delle orbite per tutti i tempi (passato presente e futuro), ma anche alla versione globalistica data da Dirac all'elettrodinamica classica, con le sue condizioni al contorno, sicché la teoria che consideriamo risulta aver perso la connotazione deterministica laplaciana, che solitamente viene attribuita ad ogni teoria classica. Queste sono le ragioni per cui la concezione che proponiamo, pur avendo carattere realistico con il fare riferimento ad orbite di particelle, risulta tuttavia compatibile con diversi aspetti caratteristici della MQ.

Tralasciando questo problema generale, che riprenderemo nella terza parte delle presenti note, nel seguito di questo capitolo illustriamo come il programma atomistico sia stato implementato da Clausius e Boltzmann, con le interpretazioni della temperatura e dell'entropia e con il teorema di equipartizione, e illustreremo i fenomeni principali che lo misero in crisi. Nei capitoli successivi mostreremo poi come questi fenomeni forzarono Planck ed Einstein ad introdurre la quantizzazione.

2.2 Le teorie cinetiche e la scoperta di Clausius: l'interpretazione meccanicistica della pressione e della temperatura (per i gas)

Che la materia sia costituita di atomi e molecole appare oggi quasi ovvio, e si fa anzi fatica a comprendere che le cose non stessero in questo modo alla fine dell'ottocento, quando erano numerosi gli scienziati antiatomisti, ovvero continuisti (o energetisti, come allora si diceva).

Atomisti e antiatomisti nella seconda metà dell'ottocento. Da una parte si trovavano gli antiatomisti, tra cui Lord Kelvin e Planck stesso (oltre a Mach, Ostwald – premio Nobel per la chimica – e molti altri), dall'altra gli atomisti, con Clausius, Maxwell, Boltzmann e Lord Rayleigh.

Per quanto riguarda il caso paradigmatico di Planck, si usa parlare di una sua "conversione" alla concezione atomistica e ai procedimenti statistici. Questa ebbe luogo proprio quando egli si rese conto (14 dicembre 1900) che la sua legge (che diede appunto origine alla MQ), che egli aveva dapprima ottenuto per via empirica (19 ottobre 1900), poteva invece essere giustificata facendo uso di procedimenti ispirati a quelli statistici di Boltzmann. Ma prima era decisamente contrario.[4,5] Si capisce così come mai sia stato Planck stesso ad introdurre la cosiddetta costante di Boltzmann k_B definita da

$$k_B = R/N_A$$

dove R è la costante dei gas (quantità macroscopica) ed N_A il numero di Avogadro, che costituisce l'emblema stesso della concezione atomistica. Ad esempio, fino al 1917 Einstein continua a scrivere R/N_A (veramente, R/N), come prima faceva Boltzmann stesso.

L'emergere della teoria atomistica

Che debbano esistere delle teorie dei continui è ben naturale, perché le esperienze ordinarie ci portano ad esprimerci in termini di campi, ovvero quantità (scalari, vettoriali o tensoriali) funzioni del posto, cioè del vettore posizione $\mathbf{x}$. Ad esempio abbiamo la densità di materia $\rho(\mathbf{x})$, connessa alla massa $M(V)$ contenuta in un volume V dalla relazione $M(V) = \int_V \rho(\mathbf{x})\mathrm{d}_3\mathbf{x}$ e analogamente le densità di carica e di corrente. il campo di velocità $\mathbf{v}(\mathbf{x})$ di un fluido, il campo elettrico $\mathbf{E}(\mathbf{x})$, e così via. La prima efficace teoria dei continui fu introdotta fin dal 1750, con l'equazione della corda vibrante di d'Alembert, seguita dalla trattazione di Lagrange del 1759 in cui l'equazione di d'Alembert venne ritrovata con una operazione di "passaggio al continuo" a partire da una trattazione con elementi discreti. Vi furono poi le analoghe teorie della elasticità in materiali tridimensionali e le teorie dei fluidi perfetti di Eulero e dei fluidi viscosi di Navier–Stokes. Poi la teoria del campo elettromagnetico di Maxwell, e infine la teoria della gravitazione di Einstein.

D'altra parte, per non risalire ad Epicuro e a Lucrezio, la concezione corpuscolare in tempi moderni risale a Galileo (si veda la prima giornata delle Dimostrazioni matematiche intorno a due nuove scienze) e Newton, il quale ad esempio oppose la sua teoria corpuscolare della luce a quella ondulatoria di Huygens. Una svolta cruciale si verificò poi nell'ambito della chimica, con la legge di Dalton delle proporzioni multiple. Il fatto che nelle esperienze coinvolgenti

[4] Alcune citazioni in proposito sono riportate in P. Campogalliani, *La ragione sommersa*, Franco Angeli (Milano, 2007), specialmente pag. 113. Si veda inoltre il libro di Kuhn, ed anche S. Brush, in *The kind of motion we call heat*.

[5] Inoltre, il fatto che la concezione atomistica fosse tutt'altro che ovvia fino ai lavori di Planck ed Einstein, è documentato ad esempio anche da titolo che Nernst – forse il più grande chimico fisico del volgere del secolo, cui è dovuto, tra le altre cose, anche il terzo principio della termodinamica – scelse per il suo celebre libro di chimica teorica, ovvero: "*Chimica teorica dal punto di vista della teoria di Avogadro e della termodinamica*', del 1893. È disponibile la traduzione inglese W. Nernst, *Theoretical Chemistry from the standpoint of the Avogadro rule and thermodynamics*, MacMillan (London, 1895).

reazioni tra quantità macroscopiche (litri o chili, quantità della vita ordinaria) di sostanze diverse si riscontri che le varie sostanze reagiscono in rapporti di materia ben definiti, venne "spiegato" da Dalton ammettendo che la materia fosse composta da corpuscoli (atomi) che si combinano tra loro a formare molecole. Infine, la relazione tra mondo macroscopico e mondo microscopico (quello degli atomi e delle molecole) venne fissata nel 1811 da Avogadro, quando egli concepì come si possa determinare il numero di molecole contenute in una definita quantità macroscopica di materia.[6]

Nacque in tal modo, nel 1811, il *numero di Avogadro*, ovvero il numero di molecole (eventualmente monoatomiche) contenute in una *mole* di sostanza. La nozione di mole è alquanto sottile, anche se l'idea centrale è abbastanza semplice. La difficoltà che si incontrava consisteva nella differenza che si riscontrava tra numero atomico e peso atomico, ovvero nel fatto che i numeri che si incontrano concretamente nella legge delle proporzioni multiple per le reazioni chimiche sono prossimi, ma non esattamente uguali, a numeri interi. Dopo il 1932, anno della scoperta del neutrone, sappiamo che questa differenza è dovuta alla presenza di isotopi, ovvero al fatto che nel nucleo di un atomo (con numero atomico individuato dal numero di protoni – uguale al numero di elettroni) sono presenti anche dei neutroni, e che l'abbondanza relativa dei vari isotopi determina il peso atomico di una data specie atomica. Ma riuscire a cavarsela in una simile situazione senza sapere dell'esistenza dei neutroni (e addirittura del nucleo, scoperto da Rutherford nel 1911) doveva risultare abbastanza difficile.[7]

Per il *numero di Avogadro*, che denoteremo con N_A, il valore oggi "consiglia-

[6]Un buon riassunto si trova nelle prime pagine del libro M. Born, *Atomic Physics*, Dover (New York, 1969), che costituisce un gioiello come introduzione alla fisica atomica.

[7]Sostanzialmente, si può pensare al numero di Avogadro come al numero di atomi contenuti in un grammo di idrogeno in forma atomica. Questa è in effetti la primitiva definizione introdotta da Dalton, e anche quella cui si riferiva sempre ad esempio, almeno fino al 1910, Einstein stesso. In effetti, le cose sono un po' più complicate, ed è giusto essere più precisi, avendo però l'accortezza di non perdere con tale precisione la intuizione fondamentale che sta dietro alla definizione. Il problema è che l'idrogeno non è comodo come sostanza di riferimento, perché ha poca disposizione alle combinazioni chimiche. Inoltre, in condizioni ordinarie l'idrogeno si trova in forma molecolare, come molecola biatomica H_2. Dunque, dopo la proposta di Dalton di fare riferimento all'idrogeno in forma atomica, in un secondo tempo Berzelius propose di prendere come riferimento l'ossigeno ponendo il suo peso atomico convenzionalmente pari a 100. In seguito si venne ad un compromesso con la scelta di Dalton, tenendo l'ossigeno come riferimento ma attribuendogli peso atomico esattamente 16, sicché il peso atomico dell'idrogeno (pur non essendo intero, a causa alla presenza dei suoi isotopi) rimaneva in ogni caso molto prossimo ad 1. Infine, si pervenne alla convenzione di fare riferimento all'isotopo del Carbonio–12, definendo il numero di Avogadro come *"il numero di atomi contenuti in* 12 *grammi dell'isotopo del carbonio detto Carbonio–12"*. Dunque per definizione una mole di Carbonio–12 pesa 12 grammi. Analogamente per le moli delle altre sostanze.

to" è [8,9]

$$N_A = 6.02 \cdot 10^{23} \ . \tag{2.2.1}$$

Analogamente, per la costante dei gas il valore "consigliato" è

$$R = 8.314 \times 10^7 \ (cgs) \equiv 8.314 \ (mks) \ . \tag{2.2.2}$$

sicché per la costante di Boltzmann $k_B = R/N_A$ si ha[10]

$$k_B = 1.380 \times 10^{-16} \, erg/K = 1.380 \times 10^{-23} J/K \ .$$

La legge dei gas perfetti

Dunque, dagli studi in ambito chimico l'ipotesi atomistica appare ben fondata. Tuttavia, concretamente, in termodinamica abbiamo a che fare con una teoria dei continui. Ad esempio, tutti conosciamo la legge dei gas perfetti

$$pV = \mathcal{N} RT \ , \tag{2.2.3}$$

che lega pressione p, volume V, temperatura (assoluta) T, numero di moli $\mathcal{N}$, mentre R è la costante dei gas. Anzi, sappiamo anche che furono proprio le proprietà termodinamiche dei gas che condussero all'introduzione della temperatura assoluta T a partire da quella empirica, confrontando l'incremento del prodotto pV con quello della temperatura empirica, e giungendo in tal modo a trovare il celebre numero 273, che conduce ad "inventare" la nozione di "zero assoluto".[11] È proprio il prodotto pV per un gas ideale (riferito ad una mole e diviso per la costante R) che definisce la temperatura.

Ora, se devo interpretare tale legge in termini molecolari, essa comporta che, avendo fissato volume V, temperatura T e numero di moli $\mathcal{N}$, la pressione debba essere la medesima per tutti i gas (nel limite di gas ideali), indipendentemente dalla struttura (in particolare la massa) delle molecole costituenti. Conta solo il numero di molecole. Questo è il cuore della concezione di Avogadro. La dimostrazione di questa proprietà (che ora ricostruiremo), e l'esperienza della meraviglia che essa suscita, è un utilissimo esercizio per entrare nello spirito delle teorie cinetiche. Si pensi alla meraviglia che si prova nel comprendere come, a

[8] Ricordiamo che il numero di Avogadro venne determinato in maniera notevolmente precisa per la prima volta da Planck, nel lavoro del 14 Dicembre 1900 – precisamente nell'ultimo paragrafo, dove lo denota con $1/\omega$ – , proprio sulla base della sua legge per il corpo nero, e poi da Einstein nei suoi lavori successivi.

[9] È importante cercare di raffigurarsi l'enormità di questo numero, che fa da ponte tra mondo microscopico e mondo macroscopico. In un manuale per studenti americani questo viene descritto nel modo seguente. Se ci raffiguriamo una molecola come una pallina da ping pong, allora un numero di palline dell'ordine di N_A ricoprirebbe la superficie degli Stati Uniti da oceano ad oceano, con uno strato dello spessore di qualche metro.

[10] Una raccolta degli articoli rilevanti della teoria cinetica può essere reperita nel già citato libro S. Brush, *The kind of motion we call heat*.

[11] W. Nernst, *Theoretical Chemistry*, pag. 31–32.

fissati volume e temperatura, la pressione raddoppia se si verifica la dissociazione di una sostanza biatomica (perché raddoppia il numero di costituenti).[12]

Il teorema di Clausius nel modello delle molecole trattate come punti materiali

Prima di considerare il caso generale di un gas perfetto costituito di molecole, cominciamo a considerare il modello prototipo di un gas che sia privo di gradi di libertà interni, ovvero sia costituito di atomi trattati come punti materiali (che hanno dunque come uniche proprietà caratteristiche la posizione e la massa[13]) Ricordiamo che un gas si dice *perfetto* o ideale se sono nulle le forze mutue tra le molecole che lo costituiscono: le uniche forze sono quelle esercitate sui punti dalle pareti. In questo caso il teorema ha la forma data sotto.[14]

Teorema 1 (di Clausius per un gas perfetto costituito di punti materiali) *Per un gas perfetto costituito di punti materiali di massa m_i si ha*

$$pV = \frac{2}{3}\overline{K}\,, \tag{2.2.4}$$

dove K è l'energia cinetica totale

$$K = \sum_i \frac{1}{2} m_i v_i^2\,. \tag{2.2.5}$$

Qui la barra denota una opportuna media sui dati iniziali, mentre la pressione viene identificata come una media temporale della forza esercitata dal gas *(ovvero dai punti che lo costituiscono)* sulla parete.

Diamo qui uno schizzo della dimostrazione, indicando quale è l'idea centrale; in particolare, procediamo "alla garibaldina" per quanto riguarda il procedimento con cui si eseguono le due medie che riguardano la pressione e l'energia cinetica.[15]

[12] Questo è appunto uno dei fenomeni principali per cui si impose "l'ipotesi" di Avogadro. Si tratta delle difficoltà che si presentava nelle cosiddette *densità di vapore anomale*, che si osservano nei procedimenti di determinazione dei pesi molecolari di molecole complesse. Ad esempio, tipicamente per il cloruro di ammonio si trova che esso appare avere una densità uguale a circa la metà di quella calcolata secondo la formula NH_4Cl. Di fronte a questa difficoltà, quasi contemporaneamente, Cannizzaro (1857), Kopp (1858) e Kekule (1858) fecero osservare che le basse densità di vapore debbono essere spiegate come corrispondenti a un minore o maggiore grado di dissociazione. Si veda W. Nernst *Theoretical chemistry*, pag. 30 e pag. 301.

[13] Ad esempio, se si considerano gli atomi come delle sfere rigide (una buona approssimazione ad esempio per i gas nobili), ciò vuol dire che stiamo trascurando le rotazioni.

[14] Si veda L. Boltzmann, *Lectures on Gas Theory. Dover ed.*. Introduzione, sec.2, pag. 104 seg.

[15] Su questo punto ritorneremo ripetutamente nel seguito. Molto interessante a questo proposito è la lunga discussione tra diversi autori, in particolare tra Einstein e Poincaré, che seguì alla relazione che Einstein tenne alla prima conferenza Solvay (si veda la riproduzione e traduzione nella raccolta italiana di opere scelte di Einstein).

Per una migliore formalizzazione, che fa uso del teorema del viriale (di Clausius), si vedano ad esempio le nostre lezioni di Meccanica Analitica 2.

Idea della dimostrazione. Consideriamo un sistema di N punti materiali di massa m che rimbalzano elasticamente tra pareti opposte di un cubo di lato L.

1. Problema delle medie sui dati iniziali. Un primo punto riguarda la scelta dei dati iniziali (posizione e velocità di tutti gli atomi). È chiaro che si porebbero considerare dati iniziali in cui tutti gli atomi hanno velocità dirette in una stessa direzione, ad esempio la direzione mormale a due pareti opposte, ma è evidente che questi sono casi eccezionali. Nei *casi generici*, o, come si dice, *per dati iniziali tipici*, si avrà una densità uniforme, e inoltre le velocità saranno orientate in maniera isotropa. Problemi di questo tipo sono quelli affrontati da Maxwell e Boltzmann, sui quali ritorneremo in seguito. Qui saremo interessati alla seguente proprietà, che ammetteremo come plausibile. Se prendiamo assi ortogonali paralleli agli spigoli del cubo, e denotiamo con $\mathbf{v}_i = (v_i, u_i, w_i)$ la velocità della i–esima particella, e con K_x l'energia cinetica lungo l'asse x e analogamente per K_y e K_z, dove

$$K_x = \sum_{i=1}^{N} \frac{1}{2} m v_i^2 \quad K_y = \dots , \quad K_z = \dots ,$$

allora, per dati iniziali tipici queste tre quantità saranno uguali, o meglio, i loro valori medi rispetto a dati iniziali distribuiti in maniera tipica saranno uguali, ovvero si avrà (denotando la media con una sopralineatura)

$$\overline{K_x} = \overline{K_y} = \overline{K_z} = \frac{1}{3}\overline{K} \tag{2.2.6}$$

dove

$$K = K_x + K_y + K_z$$

è l'energia cinetica totale. Si noti che le funzioni del tipo dell'energia cinetica totale si dicono *"funzioni di tipo somma"*, perché definite come somme, estese a tutte le particelle, della medesima quantità relativa a ogni singola particella. Sono funzioni di questo tipo quelle che intervengono nella identificazione di quantità termodinamiche in termini di quantità meccaniche.[16] [17]

2. Definizione meccanica della pressione come media temporale. La pressione può essere identificata con un opportuna valor medio della forza esercitata *dagli atomi sulla parete* (anzi su una sua area unitaria), quando la parete sia vincolata a restare fissa. Si pensi al caso significativo di una forza *a breve range*, che si annulla a distanze piccolissime dalla parete. Possiamo pensare al caso ideale di una forza di "tipo deltiforme", detta anche forza impulsiva, che è diversa da zero (in effetti infinita) solo sulla parete.[18]

[16] Si veda Khinchin. ....

[17] Se i dati iniziali su cui si media non sono distribuiti in modo appropriato, ovvero, come si dice, se lo stato iniziale non è uno stato di equilibrio, il tempo necessario perché si raggiunga l'equilibrio puó essere incredibilmente lungo. Un esempio che riguarda proprio le collisioni di punti materiali con le pareti di un cubo, come quello qui considerato, fu studiato dal grande Poincaré in un articolo del 1906.

[18] Ovvero una forza F con la propreità

$$F(t) = f\,\delta(t - t^*) , \quad f = \text{costante} .$$

se t^* è il tempo al quale avviene l'urto con la parete.

Il punto significativo è comunque che la forza (per unità di area) esercitata dal gas sulla parete è una quantità estremamente fluttuante nel tempo, essendo dovuta alle collisioni con le numerosissime particelle costituenti il gas. Quindi la pressione deve essere pensata come la media temporale della somma delle forze esercitate sulla parete dalle particelle che la urtano.[19] Consideriamo la pressione esercitata su una delle pareti ortogonali all'asse delle x, e denotiamo con $F_i^{ext}(s)$ la componente x della forza esercitata al tempo corrente s *dalla parete* sulla particella i–esima. Allora definiamo la pressione, valutata *fino al* tempo di osservazione t, mediante la relazione

$$L^2 p(t) = -\frac{1}{t}\int_0^t \sum_i F_i^{ext}(s)\,\mathrm{d}s \ ,$$

dove si è tenuto conto della legge di azione e reazione.

Cominciamo a considerare il contributo $p^{(i)}$ dovuto agli urti con la particella i–esima,

$$L^2 p^{(i)}(t) = -\frac{1}{t}\int_0^t F_i^{ext}(s)\,\mathrm{d}s \ , \tag{2.2.7}$$

che possiamo anche scrivere nella forma

$$L^2 p^{(i)}(t) = -\frac{1}{t} I(t)$$

dove abbiamo introdotto il cosiddetto impulso $I(t)$ della forza fino al tempo t, definito da

$$I_i(t) = \int_0^t F_i^{ext}(s)\mathrm{d}s \ .$$

Il calcolo dell'impulso segue immediatamente dall'equazione di Newton (o meglio, dalla sua componemte lungo l'asse x). Infatti, facendo l'ipotesi che l'urto sia elastico, ovvero che esso non alteri l'energia cinetica della particella, l'ipotesi che la forza abbia carattere impulsivo significa che l'unico effetto dell'urto è di invertire istantaneamemte il verso della velocità. Allora, integrando l'equazione di Newton

$$\frac{\mathrm{d}}{\mathrm{d}t}(mv_i) = F_i^{ext}$$

in un intervallo di tempo $(0, t)$ in cui avvenga uno e un solo urto con la parete considerata, si ha

$$I_i(t) = mv_i(t) - mv_i(0) = -2mv_i(0) \ .$$

Dunque, se al tempo iniziale $t_0 = 0$ la particella parte dalla parete opposta a quella considerata, con una assegnata velocità v_i (che prendiamo di segno positivo, avendo orientato l'asse x in maniera corrispondente), l'impulso resta nullo fino al tempo di collisione con la parete considerata (che avviene al tempo L/v_i), e poi prende il valore è $-2mv_i$.

È chiaro allora quale è l'andamento di $I_i(t)$ dell'impulso, in funzione del tempo t. Esso è una funzione a scalino, inizialmente nulla, che compie un salto di $-2mv_i$

[19] La prescrizione di dovere compiere una media temporale, viene *dimostrata* (almeno nella sostanza), nella trattazione che Boltzmann dà nel suo libro sulla teoria dei gas.

immediatamente dopo il primo urto con la parete considerata, al tempo $t_1 = L/v_i$, e resta costante fino al tempo $t_2 = t_1 + \Delta t$, dove

$$\Delta t = \frac{2L}{v_i}$$

Poi il processo si ripete, e si hanno salti tutti identici, aventi valore $-2mv_i$, a tempi

$$t_n = t_1 + n\Delta t \ , \quad n = 2, 3, \dots \ . \tag{2.2.8}$$

Allora la quantità $L^2 p^{(i)}(t) = -I_i(t)/t$ che ci interessa (si veda la (2.2.7)), viene subito calcolata, almeno per i tempi discreti t_n, nei quali essa assume il medesimo valore

$$\frac{2mv_i}{\Delta t} = \frac{mv_i^2}{L} \ ,$$

(se si prende come tempo iniziale quello a cui avviene il primo salto). Ma per tempi lunghi il calcolo diviene esatto anche per tempi t nel continuo (e anche prendendo t_0 arbitrario), e si ha

$$L^2 p^{(i)}(t_n) = \frac{n\,2mv_i}{t_1 + n\,\Delta t} = \frac{2mv_i}{(t_1/n) + \Delta t} \simeq \frac{2mv_i}{\Delta t} = \frac{mv_i^2}{L} \ ,$$

ovvero

$$V p^{(i)} = mv_i^2 \ .$$

Sommando sulle particelle si ottiene allora[20]

$$pV = 2K_x \ .$$

Infine, prendendo la media sui dati iniziali e usando l'ipotesi di isotropia (2.2.6), si ha la legge di Pascal (pressione indipendente dalla direzione della parete), e la legge di Clausius (2.2.4), $pV = (2/3)\overline{K}$.

Confrontando poi la relazione cinetica (2.2.4) con la legge dei gas perfetti $pV = \mathcal{N}RT$, si vede che il teorema di Clausius fornisce una identificazione della quantità microscopica $\overline{K}$ (energia cinetica totale, mediata) in termini della quantità macroscopica T: **Identificazione di Clausius**

$$\overline{K} = \frac{3}{2}\mathcal{N}RT \equiv \frac{3}{2}Nk_BT \ . \tag{2.2.9}$$

In particolare, questo comporta che l'energia cinetica media di un gas (nel modello di atomi puntiformi) sia indipendente dalla natura degli atomi: atomi più pesanti hanno velocità più piccole. Questo sarà confermato dai metodi di Maxwell–Boltzmann, che permetteranno inoltre di dedurre la legge della distribuzione delle velocità degli atomi.

[20] Si avrebbe qui un problema, perché il tempo Δt dipende dalla velocità della particella considerata, e diverge quando la velocità tende a zero. Ma all'equilibrio la frazione di particelle attorno a una data velocità tende a zero con la velocità.

Ma prima di discutere il teorema di equipartizione dell'energia, cosa che faremo nel prossimo Capitolo, veniamo al problema di come si modifica il teorema di Clausius quando si prendano in considerazione i "gradi di libertà interni", cioè si considerino atomi non puntiformi, oppure molecole. Apparirà allora che sull'energia che compete ai gradi di libertà interni il teorema di Clausius non dice nulla.

Nel caso delle molecole monoatomiche si deve tener conto del fatto che gli atomi hanno una struttura, ad esempio, nel modello più semplice concepibile, una struttura di sfera rigida. Dunque, oltre alle coordinate del baricentro (o centro di massa) si avranno tre coordinate angolari, cui corrispondono tre ulteriori termini di energia cinetica (in aggiunta ai tre del baricentro). Le cose sono poi alquanto più complesse quando si vengono a considerare le molecole. Ad esempio nel caso delle molecole biatomiche si hanno tipicamente, in aggiunta ai tre termini di energia cinetica del baricentro, tre ulteriori termini di energia cinetica e uno di energia potenziale.[21] In ogni caso, il punto rilevante è che, sull'energia che compete a tutti questi termini aggiuntivi, la trattazione che generalizza quella di Clausius appena discussa non dice assolutamente nulla. Una risposta verrà poi data (entro opportune ipotesi) dal teorema di equipartizione. Questa risposta è soddisfacente solo parzialmente, ma in generale fallisce, e qui appunto nascerà la meccanica quantistica.

Il teorema di Clausius per le molecole

Per fissare le idee, consideriamo il caso di N molecole identiche, ciascuna costituita di n atomi. Essendo interessati alla "spiegazione" dell'equazione di stato dei gas perfetti che coinvolge (oltre al numero di moli $\mathcal{N}$, al volume V e alla temperatura T) la pressione p, concentriamoci ancora sulle interazioni che le singole molecole hanno con una parete, trascurando di discutere le mutue interazioni tra le molecole (sul ruolo delle quali torneremo in un prossimo paragrafo).

Denotiamo con $\mathbf{x}_{is}$ il vettore posizione dell's–esimo atomo della molecola i–esima. Allora il moto della molecola sotto l'azione delle forze esterne della parete e delle forze mutue tra i diversi atomi della molecola considerata ha la forma del sistema di n equazioni

$$m_{is}\ddot{\mathbf{x}}_{is} = \sum_{s'\neq s} \mathbf{F}_{is,is'} + \mathbf{F}_{is}^{ext} \quad , \quad i = 1,\ldots,N;\ s = 1,\ldots,n \ ,$$

dove $\mathbf{F}_{is,is'}$ sono le forze "interne" agenti sull'atomo s della molecola i–esima considerata (dovute agli altri atomi $s' \neq s$ della stessa molecola), e $\mathbf{F}_{is}^{ext}$ le forze "esterne" sui singoli atomi della molecola dovute alla parete.

A questo punto si tiene conto del risultato fondamentale della meccanica dei sistemi di punti, ovvero del teorema del baricentro. Questo si ottiene sommando

[21] Se i due atomi sono a distanza fissata, si ha come un manubrio (*dumbell*) con i due atomi agli estremi, e quindi, fissato il baricentro, si hanno come coordinate libere due angoli che fissano l'orientazione del manubrio (come avviene per un atomo singolo, descritto come una sfera rigida, quando si trascuri la "rotazione propria"), In aggiunta però i due atomi possono oscillare, e il manubrio può essere trattato come una molla abbastanza rigida, e quindi si ha un oscillatore armonico monodimensionale, con i familiari termini di energia cinetica e di energia potenziale.

termine a termine tutte le equazioni di Newton, tenendo conto del principio di azione e reazione e della definizione del vettore posizione del baricentro. Allora per il moto della molecola i–esima tale teorema afferma che

$$\frac{\mathrm{d}}{\mathrm{d}t}(M_i \mathbf{v}_i^{cm}) = \sum_s \mathbf{F}_{is}^{ext}$$

dove M_i è la massa totale della molecola e $\mathbf{v}_i^{cm}$ la velocità del suo baricentro. Dunque il baricentro si muove come se fosse un punto materiale soggetto al risultante (la somma vettoriale) delle forze esterne (qui, le forze esercitate dalla parete sui singoli atomi della molecola considerata).

In conseguenza, per quanto riguarda l'interazione di ogni molecola con la parete, ci si è ridotti al caso discusso sopra, di un sistema costituito di singoli punti materiali (i baricentri, appunto, trattati come atomi puntiformi). Si conclude allora che ogni molecola contribuisce al prodotto pV soltanto con i tre termini di energia cinetica del baricentro. Ovvero, si ottiene ancora l'equazione di stato dei gas perfetti (2.2.3) con una identificazione tra temperatura ed energia cinetica analoga alla (2.2.9), dove però ora l'energia cinetica è solo quella dei baricentri.

Riassumendo, abbiamo dunque il teorema generale di Clausius e la corrispondente identificazione:

Teorema di Clausius e identificazione della temperatura, per un gas costituito di molecole. Per un gas perfetto costituito di molecole si ha

$$pV = \frac{2}{3}\overline{K^{cm}}\,, \tag{2.2.10}$$

dove

$$K^{cm} \stackrel{\text{def}}{=} \sum_i K_i^{cm}\,.$$

è l'energia cinetica totale dei baricentri. Si ha allora la identificazione

$$\overline{K^{cm}} = \frac{3}{2}\mathcal{N}RT \equiv \frac{3}{2}Nk_BT\,. \tag{2.2.11}$$

Nel procedimento seguito sopra, la identificazione della temperatura è stata ottenuta confrontando il risultato del calcolo della pressione con la legge fenomenologica dei gas perfetti. Facciamo osservare tuttavia che tale risultato può essere dedotto anche senza fare riferimento alla legge fenomenologica, mostrando direttante che nel caso del gas perfetto la quantità $(2/3)K^{cm}$ è un denominatore integrante della forma differenziale del calore $\delta Q = \mathrm{d}U + p\mathrm{d}V$.

Dunque, mediante il teorema di Clausius abbiamo una "spiegazione microscopica" della legge dei gas perfetti. In particolare essa comporta che l'energia totale dei baricentri (media) ha il medesimo valore per tutti gas, indipendentemente dalla struttura dei costituenti (molecole monoatomiche o poliatomiche) e dalle loro costanti atomiche caratteristiche (masse, frequenze delle oscillazione

interne, ...). Questo verrà confermato dal teorema di equipartizione di Boltzmann. Ma sorgerà allora un nuovo problema, perché tale teorema darà anche informazione sulle energie medie dei gradi di libertà interni, in disaccordo con l'esperienza.

I problemi che si presentano quando esistono gradi di libertà interni è ovvio. Se, ad esempio, a fissato volume si incrementa la pressione, e quindi la temperatura, dal punto di vista microscopico ciò vuol dire che, in base al teorema di Clausius, si incrementa l'energia (cinetica) totale dei baricentri, ma non abbiamo alcuna informazione su cosa succede delle energie dei gradi di libertà interni. Una eventuale "termalizzazione" che porti ad un incremento corrispondente delle energie interne è possibile (come richiesto dal treorema di equipartizione) , ma dovrebbe essere determinata da scambi di energia dovuti ad esempio alle collisioni tra le molecole. Se poi questo possa avvenire, completamente o forse solo in parte, è un problema aperto legato alla dinamica e alla modellizzazione delle forze mutue. Problema tutt'altro che semplice, la cui rilevanza fu già fortemente sottolineata da Boltzmann stesso, nel suo celebre contributo alla rivista Nature del 1895.

2.3 Il procedimento statistico di Boltzmann, la distribuzione di Maxwell–Boltzmann, e le interpretazioni probabilistiche dell'entropia e del secondo principio

Abbiamo accennato a come la trattazione cinetica di Clausius faccia riferimento a procedimenti statistici, riguardanti sia la probabilità dei dati iniziali delle molecole, sia medie coinvolgenti la dinamica (attraverso medie temporali di variabili dinamiche, come le forze esercitate dagli atomi su una parete). Si presenta allora il problema se sia possibile ed in che modo trattare le variabili dinamiche come fossero variabili casuali. È questo un aspetto del cosiddetto problema ergodico (introdotto da Boltzmann stesso), che costituisce il cuore della meccanica statistica.

Dopo Clausius, un importantissimo contributo venne dato da Maxwell, che giustificò, attraverso un procedimento astuto, la forma che ci si deve aspettare per la distribuzione di probabilità delle velocità delle molecole in un gas perfetto. Il contributo fondamentale venne però dato da Boltzmann, che riuscì non solo a generalizzare la distribuzione di Maxwell, ma soprattutto fu in grado di giustificarla sulla base di un profondo ripensamento dei fondamatenti della meccanica statistica. In effetti, egli viene considerato il creatore dei *metodi delle grandi deviazioni*[22], che costituiscono il fulcro di tutte le trattazioni meccanico–statistiche moderne.

[22]Per una prima informazione si può vedere anche Wikipedia, *Large deviations theory*. Un classico manuale è quello di Ellis.

Ora le teorie cinetiche vennero sviluppate per i gas, che sono costituiti da sottosistemi (atomi o molecole) che possono essere considerati come indipendenti, e dunque trattati con metodi statistici analoghi a quelli classici familiari rigurdanti lanci di monete e di dadi. In effetti Boltzmann, pur partendo dal caso particolare del gas, considerò anche casi del tutto generali.Tuttavia, questi suoi studi sono pochissimo noti e non conclusivi, e comunque di lettura molto difficile.[23].

Dopo di lui venne[24] il contributo dato da Gibbs nel suo fondamentale libro del 1902, moltissimo apprezzato anche da Poincaré (si veda ad esempio *La science et l'hypothèse*, del 1905). Tuttavia, rispetto a quello di Boltzmann l'approccio di di Gibbs fu alquanto diverso, anche se certamente ispirato ai tentativi di trattazione generale di Boltzmann. L'approccio di Gibbs dà di più e di meno. In qualche modo il procedimento di Gibbs è più limpido, perché ad esempio permette di separare con chiarezza gli aspetti dinamici da quelli statistici, più di quanto avvenga nella trattazione di Boltzmann. Per questo, giustamente, al procedimento di Gibbs si ispirano praticamente tutti i manuali. Tuttavia, negli studi di Gibbs poca attenzione è data al ruolo della dinamica, e ai procedimenti connessi al principio delle grandi deviazioni. Infine, attorno al 1957 vi fu il contributo di Green–Kubo, che permise di fondere in maniera limpida gli aspetti statistici e quelli dinamici, attraverso l'introduzione delle funzioni di correlazione delle variabili dinamiche di interesse. I contributi di Gibbs e di Green–Kubo verranno illustrati, anche se solo parzialmente, nell'ultimo capitolo di questa parte. Qui cominciamo con il contributo che Boltzmann diede alla meccanica statistica di equilibrio nello spirito della teoria dei gas, e tralasciamo completamente gli altri suoi contributi alla meccanica statistica di equilibrio e nonequilibrio per sistemi di tipo generale.

2.3.1 La trattazione probabilistica di Boltzmann per i gas diluiti. Gli stati macroscopici come successioni $\{n_k\}$ di numeri di occupazione nello spazio μ. Analogia con la statistica del lancio di dadi

Ben sappiamo che i metodi statistici si applicano a situazioni coinvolgenti un grande numero di eventi individuali, che nei casi più semplici possono essere considerati indipendenti, e anche identicamente distribuiti (cioè con le medesime distribuzione di probabilità per gli eventi singoli). Questi metodi vennero applicati da Boltzmann al caso dei gas (costituiti da atomi o molecole) in analogia con le trattazioni tradizionali per il lancio di N monete identiche, che può equivalentemente pensarsi come una successione di N lanci di una medesima moneta. In tal caso, le quantità di interesse si esprimono come somme di quantità relative ai lanci individuali, ovvero, nel linguaggio alla Khinchin,[25] sono *grandez-*

[23]Si veda la rassegna G. Gallavotti ....

[24]Quasi in mezzo tra Boltzmann e Gibbs stanno i primi lavori di Einstein del 1903, 1904, sui quali ritorneremo più avanti.

[25]A.I. Khinchin, *Mathematical foundations of statistical mechanics*, Dover (New york, 1949).

ze di tipo somma. Così, nel caso del lancio di N monete, in relazione all'i–esimo lancio si considera la funzione χ_T^i che assume il valore 1 se è uscita testa, e il valore 0 se non è uscita testa, e analogamente la funzione χ_C^i relativa all'uscita o meno di croce (naturalmente, si ha anche $\chi_C^i = 1 - \chi_T^i$, perché si hanno solo due possibili uscite). Allora, nell'evento globale costituito da una successione di N lanci, il numero n_T di teste è dato dalla somma $n_T = \sum_i \chi_T^i$ che coinvolge tutti i lanci (tutte le monete). Analogamente, nel caso del gioco dei dadi si hanno a ogni lancio sei possibili uscite, e quindi per ogni lancio individuale i si avranno sei funzioni χ_k^i, $k = 1, 2, \ldots, 6$. La teoria delle probabilità si interessa di eventi globali, ciascuno dei quali è definito da una ben definita successione di N lanci. Denotiamo con ω un tale evento globale (gli N risultati corrispondenti a una successione di N lanci). La conoscenza dello stato globale ω ci dà una informazione completa del sistema, perché ci dice esattamente lo stato (testa T o croce C) di ogni singola moneta (o di ogni singolo lancio). Di solito tuttavia si ricercano informazioni molto meno dettagliate, come tipicamente il numero n_T di teste uscite in una successione di N lanci (cioè in uno stato globale ω), indipendentemente dall'ordine di uscita, e dunque si considera la somma $n_T = \sum_i \chi_T^i$. La stima di somme di questo tipo, somme di funzioni relative ciascuna a un singolo evento elementare (qui, a ogni singolo lancio), costituisce il problema fondamentale del calcolo delle probabilità. In tale contesto, ad esempio, la legge dei grandi numeri afferma sostanzialmente che per la stragrande maggioranza degli eventi globali ω si ha (per monete oneste, o "*unbiased*") $\frac{1}{N}\sum_i \chi_T^i \simeq 1/2$, ovvero che il numero di teste è uguale al numero di croci. Dapprima venne studiato il caso più semplice, in cui i singoli eventi "costituenti" un evento globale ω siano indipendenti e addirittura "identicamente distribuiti", ovvero "iid" (*independent, identically distributed*). In seguito i risultati vennero estesi al caso più generale di eventi singoli non identicamente distribuiti o addirittura non indipendenti.

Anche la meccanica statistica tratta sistemi fisici costituiti da un enorme numero N di sottosistemi (con N dell'ordine del numero di Avogadro N_A). Nel caso di un gas, poi, i sottosistemi (le molecole) possono essere pensate sostanzialmente indipendenti, perché interagenti solo all'atto delle collisioni descritte da forze a breve range. Se infine consideratiamo il caso più semplice di molecole tutte uguali, abbiamo chiaramente l'analogo di un sistema di monete identiche o dadi identici. L'analogo dello stato totale (che avevamo denotato con ω) è ora il punto rappresentativo del sistema totale nello spazio delle fasi totale del sistema. Questo spazio delle fasi totale è di solito denotato (seguendo Boltzmann) come spazio Γ (leggi gamma – maiuscolo), mentre lo spazio delle fasi corrispondente a un singolo sottosistema è chiamato spazio μ (leggi mu oppure mi).

Vediamo dunque come procede Boltzmann. Ogni sottosistema (molecola), che individuiamo con un indice $i = 1, \ldots, N$, abbia un numero μ – è questo proprio il simbolo usato da Boltzmann – di gradi di libertà, e sia riferito a coordinate canoniche $z_i = (\mathbf{q}_i, \mathbf{p}_i)$ (dove $\mathbf{q}_i, \mathbf{p}_i \in \mathbb{R}^\mu$), ed abbia hamiltoniana $H(\mathbf{q}_i, \mathbf{p}_i)$.

L'hamiltoniana totale allora avrà la forma

$$H^{tot} = \sum_{i=1}^{N} H(\mathbf{q}_i, \mathbf{p}_i) + V(\mathbf{q}_1, \ldots, \mathbf{q}_N) \,.$$

dove V è una definita energia potenziale che descrive le interazioni mutue, a corto range, tra le molecole. Questa energia potenziale ha un ruolo fondamentale, perché in sua assenza ogni sottosistema si muoverebbe come se gli altri non esistessero, e in particolare l'energia di ogni sottosistema sarebbe una costante del moto. Dunque le energie dei sottosistemi variano nel tempo, e si può presumere che dopo un certo tempo (che di solito si trova essere molto breve, dell'ordine del tempo di poche collisioni molecolari) le energie dei singoli sottosistemi possano essere trattate come se fossero *statisticamente* indipendenti. Una ipotesi più o meno di questo tipo potrebbe essere chiamata, per usare un celebre termine di Boltzmann, la *Stosszahlansatz*[26] o "ipotesi del caos molecolare". Si osservi anche che, almeno per gas sufficientemente diluiti, in punti tipici dello spazio delle fasi totale Γ avviene che il valore numerico dell'energia d'interazione delle molecole è del tutto trascurabile rispetto alla parte restante dell'hamiltoniana totale (ovvero la somma delle energie dei singoli sottosistemi).

A questo punto Boltzmann si è sostanzialmente ricondotto al livello delle trattazioni classiche tradizionali del problema delle successioni di lanci di monete o di dadi ("schema di Bernoulli", nella terminologia moderna). Si ha un numero N (un numero enorme, dell'ordine del numero di Avogadro N_A) di sistemi individuali (i sottosistemi, le molecole), le cui energie, chiamiamole E_i, possono essere trattate come variabili casuali (*random variables*) che, possono assumere valori tipicamente nell'intervallo $[0, \infty)$. In effetti, per ogni sistema individuale l'informazione più rilevante (anzi l'informazione completa) è la conoscenza del punto dello spazio μ (con 2μ coordinate) in cui quel sistema si trova (questo è l'analogo del conoscere lo stato della i–esima moneta: testa T o croce C). Qui Boltzmann introduce nella descrizione una semplificazione, perché discretizza il problema, introducendo in ognuno dei singoli spazi μ una suddivisione in cellette. Si pensi ad esempio al caso di un oscillatore armonico, in cui ogni sottosistema ha due sole coordinate, q_i e p_i, e si pensi di introdurre in tale spazio cartesiano una quadrettatura.

Anticipazione: Planck vs Boltzmann. Naturalmente, Boltzmann pensa all'introduzione di questa discretizzazione (quadrettatura) come a un espediente provvisorio, con l'intento di passare poi al limite in cui la quadrettatura scompaia e si riottenga il continuo. Anticipando le cose, facciamo subito presente che quando Planck prenderà in considerazione i metodi di Boltzmann, la sua scoperta sarà che il metodo di Boltzmann produce il risultato 'giusto" se, evitando di "passare al continuo", si mantiene la cella finita, esattamente con area uguale ad $\hbar$ per ogni grado di libertà, dove $\hbar$ é la "costante ridotta di Planck, $h/2\pi$.[27]

[26] Assunzione, o ipotesi, sul numero di urti.

[27] Si constata immediatamente che, qualunque sia la coordinata configurazionale q, con le sue

L'analogia con la descrizione statistica del problema del lancio della moneta è ormai sostanzialmente completa. Ci resta anzitutto da definire per ogni sottosistema i una funzione che sia l'analogo di quello che sono le funzioni χ_T^i, χ_C^i nel caso delle monete, le quali ci davano una informazione completa sul risultato del lancio della singola moneta i, ovvero quale dei due possibili stati (testa T, o croce C) si è presentato. Qui l'informazione completa sarebbe sapere in quale punto dello spazio μ si trova il sottosistema i. Per via della discretizzazione introdotta, l'informazione completa sul singolo sottosistema è allora sapere in quale cella esso si trova. A ciò provvede la funzione χ_k^i che assume il valore 1 se il sottosistema i si trova nella cella k e il valore 0 se si trova fuori.

Ci resta infine un ultimo passaggio. Si tratta dell'analogo della informazione che dice che, in corrispondenza di un completo lancio di monete (ovvero in corrispondenza di quello che abbiamo chiamato un evento totale, e denotato con ω), sono uscite un certo numero n_T di teste e un numero n_C di croci (naturalmente, con il vincolo $n_T + n_C = N$). Ovviamente, l'informazione fornita dalla conoscenza dell'evento globale ω ci dice molto di più, perché ci dice anche esattamente quali monete (ad esempio. la quinta. le dodicesima, la ventesima, ...) sono nello stato testa. Nel nostro caso, l'analogo del numero di teste o del numero di croci è il numero di sottosistemi, n_k che si trova nella generica cella k, che viene di solito chiamato con il nome "*numero di occupazione*" della cella k. Evidentemente ancora abbiamo a che fare con funzioni "di tipo somma", perché si ha

$$n_k = \sum_i \chi_k^i \,,$$

dove ogni addendo dipende da uno solo dei sottosistemi (quello i–esimo). Abbiamo già osservato che la descrizione data dalla successione dei numeri di occupazione in

$$\{n_k\} \equiv (n_1, n_2, n_3, \ldots) \tag{2.3.1}$$

è una descrizione altamente ridotta, perché ci dice quanti sottosistemi sono in ognuna delle celle, ma non quali. In questo senso la descrizione del sistema totale tramite la successione $\{n_k\}$ dei numero di occupazione è una descrizione **macroscopica**, e non **microscopica**. Essa ci permette di avere un informazione

---

proprie dimensioni (ad esempio, una lunghezza o un angolo, adimensionale), se p è il corrispondente momento coniugato, allora il prodotto qp (o anche $\mathrm{d}q\mathrm{d}p$, che rappresenta un elemento di area nello spazio delle fasi) ha le dimensioni di una azione, come h. Basta a tal fine ricordare la definizione di momento p coniugato a q in termini della lagrangiana (che è dimensionalmente un'energia) sicché si ha

$$pq = \frac{\partial L}{\partial \dot{q}} q \,.$$

Dunque le dimensioni di q si eliminano tra numeratore e denominatore, e si resta con una energia per un tempo. Ne segue ad esempio che, se q è un angolo, allora il momento coniugato (che è un momento angolare) ha le dimensioni di una azione. Non è un caso che lo stato fondamentale dell'atomo di idrogeno sia quello per cui, facendo riferimento al modello classico, il momento angolare dell'elettrone ha valore $\hbar$.

significativa anche sul sistema globale, ma solo per quantità che abbiamo chiamato (seguendo Khinchin) "di tipo somma", ovvero quantità che sono somme di funzioni identiche relative ad ogni singolo sottosistema. Infatti, consideriamo una qualsiasi variabile dinamica F di un singolo sottosistema, definita quindi nello spazio μ, e sia F_k un tipico valore che essa assume nella cella k (ad esempio nel suo "centro"). Allora, se è nota la successione $\{n_k\}$, il corrispondente valore per il sistema totale sarà dato da $\sum_k n_k F_k$. Queste sono le quantità di cui si occupa Boltzmann, eventualmente passando alla corrispondente descrizione "continua".

Naturalmente, la descrizione ridotta in termini di successione $\{n_k\}$ di numeri di occupazione può essere espressa in maniera completamente equivalente anche attraverso la successione di numeri

$$p_k = \frac{n_k}{N} \qquad \left(0 \le p_k \le 1\,, \quad \sum_k p_k = 1\right). \tag{2.3.2}$$

che evidentemente hanno il significato di **"frequenze empiriche di occupazione delle celle"**, analoghe alle frequenze di teste o di croci nel lancio di N monete.

2.3.2 La probabilità a priori nello spazio Γ e il problema ergodico

Abbiamo dunque mostrato come Boltzmann si concentri sugli stati macroscopici, che è il nome che abbiamo dato alle successioni $\{n_k\}$ di numeri di occupazione delle celle nello spazio μ. Che il nome di stato macroscopico sia ragionevole si capisce bene. Infatti, se trascuriamo le velocità e ci concentriamo sulle posizioni, nel caso di un gas è ben concepibile di suddividere in cellette il volume fisico V in cui si trova ogni molecola del gas e misurare la densità del gas in ogni cella. In tal modo, la successione $\{n_k\}$ definisce la densità spaziale del sistema, come funzione del posto. Essa ci dà però solo una informazione macroscopica, perché non ci dice quali particelle sono nella cella (o vulumetto) k–esima.

A questo punto Boltzmann introduce in maniera ragionevolissima una stima della probabilità di ogni stato macroscopico $\{n_k\}$, che egli denota con $W(n_1, n_2, \ldots)$ (la ragione della notazione W è che in tedesco probabilità si dice *Wahrscheinlickeit*, letteralmente verosimiglianza). Questo gli permetterà poi di calcolare quale è lo stato macroscopico che ha massima probabilità (con i vincoli che siano fissati l'energia totale E e il numero N di particelle) e trova che tale stato è il cosiddetto stato di Maxwell–Boltzmann. Inoltre, risulta non solo che tale stato ha probabilità massima, ma addirittura che ogni altro stato ha probabilità potremmo dire evanescente, nulla, nel senso che essa ha una taglia che decresce esponenzialmente al crescere di N, da valutarsi con N dell'ordine del numero di Avogadro.

Il punto di partenza consiste nell'assumere come postulato fondamentale che la probabilità (relativa)[28] che il sistema si trovi in una regione dello spazio delle

[28] Si parla di probabilità *relativa*, perché lo spazio delle fasi ha evidentemente misura di Lebesgue infinita, sicché tale misura può essere usata solo per dedefinire il rapporto delle probabilità di due regioni aventi misura finita.

fasi totale Γ sia proporzionale alla misura di Lebesgue di quella regione. In particolare, la probabilità $W(n_1, n_2, n_3, \ldots)$ di una stato macroscopico, sarà allora proporzionale alla misura di Lebesgue del volume dell'insieme dei punti dello spazio Γ cui corrisponde la assegnata successione $\{n_k\}$. Questa ipotesi viene talvolta chiamata *ipotesi della equiprobabilità a priori*, nel senso che, facendo una scelta a caso, sia più probabile finire in una regione "grande" (cioè di volume grande) che in una regione "piccola".

Osservazione. Il fatto di usare la misura di Lebesgue può essere motivato in base alla proprietà che essa è l'unica misura invariante per traslazioni spaziali. Infatti, poiché il passaggio da un osservatore inerziale ad un altro corrisponde ad una traslazione nello spazio delle fasi, quella proprietà pone la misura di Lebesgue in una posizione privilegiata dal punto di vista della relatività galileiana. Inoltre, si dimostra che per sistemi hamiltoniani la misura di Lebesgue è invariante anche per traslazione temporali. Si tratta del cosiddetto *teorema di Liouville*, di cui discuteremo nel capitolo 6.

Premessa: il problema ergodico, come problema della giustificazione dinamica dell'ipotesi di equiprobabilità, o di ogni altra analoga ipotesi che si assuma

Nel corso della sua vita, Boltzmann affrontò il problema della probabilità degli stati macroscopici, dapprima per analogia con il problema classico del lancio delle monete o dei dadi, in cui si ammette implitamente che sia assegnata a priori una probabilità di avere ad esempio testa o croce. Ma sempre piú si occupò poi della giustificazione dinamica dell'analogo problema nel caso della meccanica statistica, fondando quindi il problema che dopo di lui viene detto problema ergodico, che ora cerchiamo di spiegare.

Per quanto riguarda i sistemi dinamici hamiltonani, sappiamo che l'energia é una costanre del moto, ovvero che ogni dato iniziale nello spazio delle fasi Γ dà luogo ad un'orbita che giace su una superficie Σ_E di fissate energia E, ovvero, come si dice, che ogni superficie Σ_E è una *superficie invariante* (se contiene un punto, allora contiene tutta l'orbita che esso genera). Ora, se esiste un'altra costante del moto, indipendente dall'energia (cioè dalla hamiltoniana H) esistono famiglie di superfici invarianti che sono trasversali alle superfici Σ_E, e quindi esistono ostacoli dinamici a che le orbite possano estendersi per tutta la superficie Σ_E. Sostanzialmente, i sistemi ergodici son quelli per cui non esiste alcuna ostruzione dinamica a visitare tutte le regioni della superficie Σ_E, e che, anzi, le visitano "tutte". In maniera un poco più precisa, un sistema viene detto ergodico in senso topologico se, su ogni superficie Σ_E, ogni punto iniziale (a meno di un insieme di misura nulla) dà luogo a un'orbita densa su Σ_E. Si ha poi ergodicità in senso metrico (cioè rispetto alla misura) se l'orbita non solo visita ogni regione di Σ_E, ma anche la visita per un tempo (detto solitamente tempo di soggiorno) uguale alla misura dele regione. Questo problema ergodico, sollevato in maniera esplicita per la prima volta da Boltzmann stesso, fu poi ripreso in ambito fisico, "modernamaente", da FPU nel 1954. Esso è implicitamente al cuore del primo ca-

pitolo del volume 5 (Meccanica Statistica) del trattato di fisica teorica del "divino Landau" (ovvero di Landau Lifshitz).

La probabilità degli stati macroscopici nella statistica di Boltzmann

Dunque, Boltzmann implicitamente ammette di trattare con un sistema ergodico, e quindi assume il principio di equiprobabilità nello spazio Γ, riservandosi poi di introdurre "a posteriori", nel corso delle sue argomemtazioni, il vincolo che sia fissata l'energia E del sistema considerato. In tal modo, avendo assunto che *la probabilità di trovarsi in una regione dello spazio* Γ *sia propozianale al suo volume*, egli può procedere a determinare la probabilità di uno stato macroscopico con procedimenti tipici del del calcolo combinatorio.

Infatti egli dìmostra che, *a meno di una costante moltiplicativa*,[29] si ha

$$W(n_1, n_2, \ldots) = \frac{N!}{n_1!\, n_2!\, n_3! \ldots} \,. \tag{2.3.3}$$

che è il familiare *coefficiente multinomiale* del calcolo combinatorio.

La dimostrazione della formula di Boltzmann (2.3.3) è quasi immediata. L'argomento centrale, che ora ilustriamo, prescinde dalla finitezza delle celle, e riguarda la corrispondenza tra punto $z \in \Gamma$ da una parte, e N–uple di punti nello spazio μ dall'altra, coinvolgendo argomenti di tipo combintorio. Infatti, ogni punto z di Γ produce una ben definita successione $\{n_k\}$, mentre la corrispondenza inversa à tutt'altro che univoca, e questo fatto, come ora mostriamo, conduce alla spiegazione del fattore multinomiale dato dalla (2.3.3), che Boltzmann chiama con lo strano nome di *numero di complessioni*.

Infatti un punto $z \in \Gamma$ è definito da una N–upla *ordinata* di punti nello spazio μ

$$z \equiv (z_1, z_2, \ldots, z_N)$$

perché ogni punto z_i si riferisce a un preciso sottosistema, quello i–esimo. Dunque un punto z di Γ produce N punti nello spazio μ, che a loro volta producono una ben definita successione $\{n_k\}$ di numeri di occupazione (stato macroscopico). Invece, se si assegna una successione $\{n_k\}$ di numeri di occupazione, restano assegnati (a meno dell'indeterminazione dovuta alla finitezza del volume delle celle elementari) N punti nello spazio μ, ma questi punti, se ci limitiamo a contarli, sono sprovvisti di indice. È quindi evidente che (ancora trascurando l'indeterminatezza dovuta alla finitezza del volume delle cellette) esistono un numero ben definito di successioni $(z_1, z_2, \ldots, z_N)$, ovvero di punti $z \in \Gamma$, che producono l'assegnato stato macroscopico. Evidentemente se nell'assegnato stato macroscopico i numeri di occupazione n_k fossero tutti 0 o 1, allora il numero di punti diversi in Γ sarebbe $N!$, che è il numeratore della formula di Boltzmann per il numero di complessioni (infatti i diversi punti si ottengono scambiando i

[29] Avendo trascurato questa costante, la W rappresenta piuttosto una *probabillità relativa*, nel senso che permette di calcolare il rapporto delle probabilità di due stati.

nomi dei sottosistemi, e quindi compiendo $N!$ permutazioni). Per comprendere la ragione del denominatore, seguendo il ragionamento originale di Boltzmann esposto nel suo lavoro del 1877, bisogna tenere conto di un altro fatto, ovvero che se ad esempio 3 punti dello spazio μ sono sovrapposti (coincidono), allora, tra gli $N!$ punti dello spazio Γ ottenuti per permutazione, che già abbiamo individuato, ve ne sono 3! che coincidono (sempre a meno dell'indeterminazione dovuta alla finitezza della celletta).[30] Occorre però tenere presente che questi diversi punti in Γ corrispondenti ad un medesimo stato $(n_1, n_2 \ldots)$ sono in generale molto distanti tra di loro, per cui l'ipotesi ergodica svolge qui un ruolo essenziale.

Esempio. Consideriamo il caso in cui ci dimentichiamo del momento e si abbia una sola coordinata spaziale x, sicché lo spazio μ è monodimensionale, e inoltre si ha $N = 2$, sicché lo spazio Γ è il piano. con coordinate (x_1, x_2). In tal caso, se nello spazio μ si hanno due punti con coordinate ad esempio 3 e 10, allora a questo stato corrispondono due punti nello spazio Γ, uno di coordinate $(3, 10)$ e un altro (simmetrico rispetto alla bisettrice) con coordinate $(10, 3)$. Ma se i due punti nello spazio μ coincidono, ad esempio con $x_1 = x_2 = 3$, allora a questo stato corrisponde nello spazio Γ un unico punto, di coordinate $(3, 3)$, disposto sulla bisettrice.

Se poi si tiene conto della finitezza delle celle (ad esempio tutte uguali, di volume δV_μ, che comporta una corrispondente suddivisione dello spazio Γ in celle di volume $(\delta V_\mu)^N$), si comprende come il procedimento che fa uso del calcolo combinatorio non sia altro che un modo per misurare la probabilità di uno stato macroscopico $(n_1, n_2, \ldots)$ come proporzionale al volume del corrispondente volume nello spazio Γ. Come avevamo osservato, questo è un fatto che rimanda a una giustificazione di carattere egrgodico.

Altro procedimento per il fattore muultinomiale. La formula per il coefficiente multinomale può essere dedotta anche nel modo seguente (si vedano i due libri di Fermi, *Introduzione alla fisica atomica*, e *Molecole e cristalli*) . Dato il numero di occupazione n_1, della prima celletta, il numero di modi in cui posso scegliere n_1 sottosistemi tra gli N è il numero di combinazioni di N oggetti a n_i a n_1, ovvero

$$\frac{N!}{n_1!(N-n_1)!} \, .$$

Messa a posto la prima celletta, restano $N - n_1$ sottosistemi, e allora quando assegno n_2 ho

$$\frac{(N-n_1)!}{n_2!\,(N-n_1-n_2)!}$$

scelte possibili per riempire la celletta 2, e così via, sicché il numero totale, che si ottiene evidentemente per moltiplicazione di quei numeri parziali, fornisce il risultato

[30] Considerazioni analoghe si compiono per determinare ad esempio i coefficienti dello sviluppo di $(a + b + c)^N$. Ogni termine contiene una successione (ordinata !) di N fattori, del tipo $aabacbc \ldots c$. La situazione è la stessa di quando si ha un punto "globale" ω nel problema del lancio di dadi a tre facce. Si hanno N lanci, con tre possibili uscite, a oppure b oppure c. Ad ogni punto ω corrispondono tre numeri di occupazione n_1, n_2, n_3 (le potenze a cui elevare i fattori a, b, c). Invece, assegnate le tre potenze (la successione di numeri di occupazione), si capisce subito che il corrispondente fattore è proprio il relativo coefficiente polinomiale.

(ad ogni moltiplicazione si elimina un fattore a denominatore con un fattore uguale a numeratore).

2.3.3 Lo stato di Maxwell–Boltzmann come stato di massima probabilità, nell'ambito della statistica di Boltzmann

Boltzmann può ora chiedersi quale sia lo stato macroscopico (ovvero la successione $\{n_k\} \equiv (n_1, n_2, \ldots)$) più probabile.

Gli stati ad energia finita come stati di grande deviazione

Qui si incontra un primo fatto significativo, ovvero che, con la definizione data per la probabilità degli stati macroscopici, lo stato più probabile ha energia media *infinita*.[31] Questo è ovvio, perché l'energia è una quantità illimitata superiormente, e abbiamo assunto uniforme la probabilità a priori nello spazio Γ, data dalla misura di Lebesgue. Quindi è nulla la probabilità che il valor medio dell'energia del sistema sia finita. Boltzmann assume allora il punto di vista tipico di quella che modernamente, proprio a seguito dei suoi lavori, viene chiamata la *teoria delle grandi deviazioni*. Ovvero, prende atto del fatto che possiamo considerare un sistema ben concreto (costituito di N particelle) che possiede una ben concreta, nota, energia totale finita E, nonostante che tale situazione abbia a priori probabilità nulla (si tratta appunto di una "grande deviazione" da quello che a priori sarebbe il valor medio, che qui è infinito). Quindi egli ricerca quale sia lo stato macroscopico che ha probabilità massima, quando si imponga dal di fuori (o *a posteriori*, come si dice) la condizione che l'energia sia fissata a un valore finito E (si tratta dunque di una *probabilità condizionata*, nel linguaggio comune della teoria delle probabilità).

Calcolo del punto di stazionarietà

Dobbiamo dunque massimizzare la funzione $W(n_1, n_2, n_3, \ldots)$ con i due vincoli $\sum n_k = N$, $\sum \epsilon_k n_k = E$, dove ϵ_k è un valore tipico dell'energia della celletta k–esima, ed E l'energia totale del sistema. Boltzmann procede pensando al caso di numeri n_k grandi che possano essere trattati come numeri reali (anziché interi), sicché valga l'approssimazione di Stirling $n! \simeq (n/e)^n$, ovvero

$$\log n! \simeq n \log n - n \ .$$

Massimizzare W è equivalente a massimizzare $\log W$, come ora faremo.

Questo è ovvio, perché il logaritmo è una funzione monotóna. Ma nel nostro caso vi sono anche due ragioni concrete per questa scelta. La prima è che per calcolare il punto

[31] Naturalmente, per energia di uno stato macroscopico intendiamo l'energia totale del sistema, data da

$$E = \sum_k \epsilon_k n_k \ ,$$

dove ϵ_k è un valore tipico della cella k, ad esempio il valore nel centro della cella.

di stazionarietà si devono eseguire della derivate, e la derivata di una somma à molto più semplice di quella di un prodotto. Inoltre si ha a che fare con dei numeri fattoriali, che nel nostro caso sono tanto grandi da essere, potremmo dire, "disumani". I loro logaritmi sono invece più "umani, come anche la loro approssimazione sopra ricordata.

Cominciamo a determinare una forma significatìva che assume $\log W$, che ci interesserà in seguito. Dalla definizione si ha

$$\log W = N \log N - N - \sum n_k \log n_k + \sum n_k \ .$$

Ma, utilizzando $\sum n_k = N$ per cancellare il termine N, e scrivendo $N \log N = \sum_k n_k \log N$, si ha immediatamente $\log W = -\sum n_k \log \frac{n_k}{N}$, ovvero, in termini delle frequenze di occupazione $p_k = n_k/N$,

$$\log W = -N \sum_k p_k \log p_k \ . \tag{2.3.4}$$

Questa è una prima forma di quella che vedremo essere la fondamentale relazione di Boltzmann tra entropia S e probabilità W, ovvero [32]

$$S = \log W$$

Osservazione. Un abuso di notazione. La relazione esatta che coinvolge l'entropia termodinamica S sarà in effetti$S = k_B \log W$, come richiesto anche per ragioni dimensionali. Per ora la relazione scritta sopra appare solo come una comoda notaione per $\log W$, e quindi dovremmo scegliere per $\log W$ un altro simbolo, come per esempio $\tilde{S}$. Per non copmlicare le notazioni teniamo comunque la scelta fatta sopra, e al momento opportuno inseriremo il fattore dimensionale k_B.

Procediamo ora alla massimizzazione di $(\log W)/N$. In termini delle frequenze di occupazione, il vincolo sul numero di particelle e sull'energia prendono allora la forma $\sum p_k = 1$, $\sum p_k \epsilon_k = E/N$. In effetti, quello rilevante è il secondo vincolo (sull'energia), perché, quando sarà stato utilizzato il secondo vincolo, il primo verrà soddisfato in maniera banale mediante una *normalizzazione*, cioé dividendo il risultato per un opportuno fattore. Usando il metodo dei moltiplicatore di Lagrange per tener conto del vincolo sull'energia dobbiamo dunque massimizzare la funzione[33]

$$-\sum_k p_k \log p_k - \beta \sum p_k \epsilon_k$$

[32]Quando si passa al continuo, l'analoga della funzione $\sum_k p_k \log p_k$ (si noti, senza il segno meno), espressa dunque come un integrale, coincide con la celebre "funzione H di Boltzmann".

[33]Evidentemente il segno $-$ davanti al moltiplicatore β è arbitrario. Esso è stato scelto perché in tal modo il moltiplicatore di Lagrange risulterà positivo (nel caso molto significativo in cui le energie ϵ_k sono positive, come per una particelle libera o per un oscillatore armonico). Ma ci sono casi più generali in cui β può essere negativo – con la scelta qui fatta per il segno. D'altra parte si trova che β é la temperatura inversa, e allora si dovrà ammettere che in tali casi si ha a che fare con temperature negative. Si veda il libro di Landau Lifshitz sulla Meccanica Statistica, sec. 71.

espressa in termini delle variabili p_k, dove β è appunto il moltiplicatore di Lagrange. Ricordando che la derivata di $x \log x$ è $\log x + 1$, e uguagliando a zero ognuna delle derivate rispetto alle variabili p_k, si ha allora

$$\log p_k = -1 - \beta \epsilon_k \ ,$$

ovvero (ricordandosi della libertà del fattore moltiplicativo),

$$p_k(\beta) = C e^{-\beta \epsilon_k} \ .$$

La condizione $\sum p_k = 1$ fornisce poi $C = 1/\mathscr{Z}(\beta)$, dove abbiamo introdotto la *funzione di partizione* (detta anche, con evidente significato, *sum over states* dal tedesco *Zustandsumme*), definita da

$$\mathscr{Z}(\beta) = \sum_k e^{-\beta \epsilon_k} \ . \tag{2.3.5}$$

Naturalmente, si ammette che la serie definente la funzione di partizione $\mathscr{Z}(\beta)$ converga, il che richiede anzitutto che si prenda $\beta > 0$ (nel caso molto significativo in cui le energie siano positive). Vedremo poi che β viene solitamente interpretato come l'inverso della temperatura, o precisamente come $1/k_B T$).[34]

Abbiamo dunque trovato la distribuzione di Maxwell–Boltzmann

$$p_k(\beta) = \frac{e^{-\beta \epsilon_k}}{\mathscr{Z}(\beta)} \quad \text{ovvero} \quad n_k(\beta) = N \frac{e^{-\beta \epsilon_k}}{\mathscr{Z}(\beta)} \tag{2.3.6}$$

con il suo celebre esponenziale. Inoltre, per quanto riguarda il valore di β, esso viene in linea di principio determinato *implicitamente* tramite la condizione di fissata energia $\sum \epsilon_k n_k(\beta) = E$, ovvero

$$E = N \frac{\sum \epsilon_k e^{-\beta \epsilon_k}}{\mathscr{Z}(\beta)} \ . \tag{2.3.7}$$

La proprietà di massimo. Decadimento esponenziale (al crescere di N) della probabilità relativa degli altri stati

Resta ancora da mostrare che siamo in presenza proprio di un massimo. Vogliamo inoltre mostrare che, asintoticamente in N, gli stati macroscopici diversi da quello di MB hanno probabilità del tutto trascurabile. Ad esempio, nell'agile libretto di Pauli, questo risultato si ottiene compiendo, per la probabilità di uno stato macroscopico, uno sviluppo al secondo ordine negli scarti δp_k dal punto di stazionarietà, e mostrando che essi sono distribuiti in maniera gaussiana. Invece,

[34] Si noti che non si ha convergenza nel caso dell'atomo di Idrogeno. Questo problema è stato discusso nel caso quantistico da Fermi nel 1923, nel lavoro *Sulla probabilità degli stati quantici*, lavoro n. 7a, pag. 118 delle Note e Memorie (Collected Papers), Vol.I., corrispondente al lavoro E. Fermi, Z. Physik, **26**, 54–56 (1924), non ristampato nelle opere di Fermi.

una dimostrazione ancora più generale, ovvero non limitata al secondo ordine negli scarti, si ottiene in maniera addirittura più semplice con il procedimento seguente, che è caratteristico dei metodi delle grandi deviazioni, i cui inizi sono appunto fatti risalire a Boltzmann.

Facciamo ancora riferimento alla espressione (2.3.4) per $\log W$, e riscriviamo la probabilità W di un generico stato macroscopico nella forma significativa

$$W(p_1, p_2, \ldots) = \exp\big[Ns(p_1, p_2, \ldots)\big] ,$$

dove abbiamo introdotto la funzione s definita da

$$s(p_1, p_2, \ldots) \stackrel{\text{def}}{=} -\sum_k p_k \log p_k \ .$$

Vedremo che, a meno del fattore k_B, la funzione s ha il significato fisico di entropia specifica per sottosistema. Fin da ora, con abuso di linguaggio, chiameremo già entropia la funzione $S(p_1, p_2, \ldots) = Ns(p_1, p_2, \ldots)$. Il punto essenziale è invece che all'esponente viene messo in evidenza il fattore N, mentre la quantità specifica s risulta indipendente da N, almeno asintoticamente in N.

Evidentemente s ha valori positivi, e si controlla facilmente che, come funzione delle variabili p_k, essa è convessa (cioè concava verso il basso, ovvero il suo grafico è quello di una montagnetta, e non di una scodella), in quanto somma di funzioni evidentemente convesse (si disegni la funzione $-p \log p$ della variabile reale p nell'intervallo $(0, 1]$). Ora, dato che la funzione s è convessa e presenta un punto di stazionarietà, si conclude che tale punto può essere solamente un massimo. Dunque abbiamo mostrato che lo stato di MB è lo stato macroscopico di massima probabilità.

Veniamo infine al confronto con la probabilità degli altri stati. Questo si ottiene scrivendo l'espressione della probabilità di uno stato generico relativamente alla probabilità dello stato di massimo (ovvero lo stato di MB). Questa probabilità relativa è data dal rapporto

$$\frac{W(p_1, p_2, \ldots)}{W_{max}} = \exp\Big(N\big[s(p_1, p_2, \ldots) - s_{max}\big]\Big) . \tag{2.3.8}$$

dove abbiamo denotato con con W_{max} la probabilità dello stato di massimo e con s_{max} la corrispondente entropia specifica. Allora è evidente che, se ci si pone in uno stato $(p_1, p_2, \ldots)$ diverso dal massimo, si ha

$$s - s_{max} < 0 ,$$

e dunque al crescere di N la probabilità relativa di quel punto decresce esponenzialmente con N, ovvero il massimo diventa piccato esponenzialmente in N (o anche, come potremmo dire, si ha asintoticamente un massimo di tipo deltiforme).

Vedremo qui sotto che un elemento fondamentale della meccanica statistica classica consiste nell'interpretare β come "temperatura inversa" (più precisamente come $\beta = 1/k_B T$ dove T è la temperatura assoluta) e S (o piuttosto $k_B S$) come entropia. In effetti queste sono proprietà tutt'altro che ovvie, e anzi a tutt'oggi non completamente chiarite.

2.3.4 La termodinamica statistica negli stati macroscopici di equilibrio, e la celebre relazione di Boltzmann $S = k_B \log W$

Occupiamoci ora di definire le funzioni termodinamiche nello stato di MB (Maxwell–Boltzmann). In effetti daremo prima una trattazione più generale, e poi verremo al caso di MB.

Cominciamo dunque con la prima identificazione tra quantità meccaniche e quantità termodinamiche (o *analogie termodinamiche*, come diceva Boltzmann e dirà poi Gibbs).[35] Si tratta della identificazione tra energia meccanica E ed energia termodinamica U, che viene realizzata in maniera ovvia, come

$$E \equiv U \ . \tag{2.3.9}$$

Chiaramente, la quantità U/N esprime la *energia specifica per sottosistema*, ovvero l'energia totale divisa per il numero di sottosistemi.[36] [37]

Il punto delicato consiste invece nell'identificare in termini meccanici la temperatura assoluta T e l'entropia S, Si tratta delle due quantità termodinamioche tipiche, che in genereale non possono essere pensate come valori medi di grandezze meccaniche, ovvero di funzioni definite nella spazio delle fasi totale Γ. È vero che, dal teorema di Clausius precedentemente dimostrato nel caso del gas perfetto la temperatura appare come valor medio di una grandezza meccanica,un multiplo dell'energia cinetica specifica, $2K/3N$ (o, meglio, la corrispondente espressione in termini energia cinetica del baricentro). Ma, come vedremo, questo è garantito solo nel caso di gas molto diluiti e a temperature altissime.

La identificazione di temperatura ed entropia si comprende addirittura meglio nel caso generale in cui la probabilità W degli stati macroscopici sia lasciata

[35] Per la termodinamica statistica si può vedere il bel libretto di Schrödinger e una appendice dell' *Atomic Physics* di Born. Per quanto riguarda i lavori originali di Boltzmann, in particolare dove introduce la relazione dell'entropia con il logaritmo del corrispondente volume nello spazio Γ, si veda Boltzmann, *Gesammelte Werke*, n. 39 pag. 121 e n. 42 pag. 193, citato da Sommerfeld (NB Le pagine sopra riportate sono quelle indicate da Sommerfeld, che probabilmente aveva a disposizione una edizione delle opere diversa da quella disponibile oggi). Si veda A. Sommerfeld, *Thermodynamics and Statistical Mechanics*, pag. 213. La relazione tra entropia e probabilità è data nelle opere di Boltzmann a pag, 215 del secondo volume. Veramente, sembrerebbe trovarsi nella sezione V del lavoro 42 (1877), pag. 215 fino alla fine, nell'edizione delle opere disponibile oggi.

[36] Quando si parla di "energia specifica" si intende solitamente l'energia totale per mole, ovvero divisa per il numero di moli. Dal punto di vista termodinamico è come se avessimo introdotto una nuova unità di misura della quantitá di materia, in modo che sia 1 mole $= N_A$ particelle.

[37] Talvolta ci si riferisce alla quantità U/N come all'energia media di *ogni singolo* sottosistema, coerentemente con l'analoga interpretazione per le frequenze p_k come probabilità empiriche, citata in precedenza.

ancora impregiudicata, senza richiedere che essa coincida con l'espressione data dalla (2.3.3): ricordiamo che quest'ultima definisce la statistica di Boltzmann, e risulta essere appropriata solo nel caso di un sistema ergodico, per il quale non si hanno vincoli dinamici sulla superficie di energia considerata. Lasciamo dunque la forma della funzione $W(p_1, p_2 \ldots)$ indeterminata. Solo in seguito torneremo al caso particolare di Boltzmann.

Scriveremo dunque in generale

$$W(n_1, n_2 \ldots) = e^{S(n_1, n_2 \ldots)} \,, \quad \text{ovvero} \quad S \stackrel{\text{def}}{=} \log W \,, \tag{2.3.10}$$

con una funzione $S(n_1, n_2 \ldots)$ imprecisata, richiedendo solo che essa sia positiva e convessa. Per comodità la chiameremo già entropia. Dunque in virtù della ipotesi di convessità esisterà un unico stato di massima entropia, che chiameremo *stato di equilibrio* rispetto alla assegnata forma dell'entropia, e ci occupiamo di ritrovarne la termodinamica in termini statistici.

Lo stato $(n_1, n_2 \ldots)$ è dunque anzitutto uno stato di stazionarietà dell'entropia, per un assegnanto valore di energia $E = \sum \epsilon_k n_k$. Dunque esso soddisfa le condizioni $\frac{\partial S}{\partial n_k} - \beta \epsilon_k = 0$, per tutti i k, ovvero

$$\frac{\partial S}{\partial n_k} = \beta \epsilon_k \,, \quad k = 1, 2, \ldots \,, \tag{2.3.11}$$

dove β è ancora il moltiplicatore di Lagrange. Dunque il differenziale della funzione S prende l'espressione

$$\mathrm{d}S = \beta \sum \epsilon_k \mathrm{d}n_k \,. \tag{2.3.12}$$

D'altra parte mostreremo subito sotto che la forma differenziale $\sum \epsilon_k \mathrm{d}n_k$ è nient'altro che l'espressione meccanico–statistica di quello che in termodinamica è il calore δQ trasferito al sistema durante un processo in cui si incrementano di $\mathrm{d}n_k$ i numeri di occupazione. Ovvero, si ha

$$\sum \epsilon_k \mathrm{d}n_k = \delta Q \,. \tag{2.3.13}$$

Dunque la relazione (2.3.12) si scrive nella forma

$$\mathrm{d}S = \beta \delta Q \,, \tag{2.3.14}$$

che costituisce un risultato fondamentale per la termodinamica statistica nell'approccio alla Boltzmann. Denotiamo con questo nome l'approccio in cui gli stati macroscopici sono individuati da successioni $(n_1, n_2, \ldots)$ di numeri di occupazione nello spazio μ, la cui probabilità $W(n_1, n_2, \ldots)$ abbia la proprietà che $S \stackrel{\text{def}}{=} \log W$ sia un funzionale[38] positivo e convesso. In particolare in questo ambito si ha anche la statistica di Maxwell–Boltzmann, identificata dallla scelta del

[38] NOTA PER GLLI AUTORI. Spiegare funzionale vs funzione.

funzionale $W(n_1, n_2, \ldots)$ dato dalla (2.3.3), che è la statisitca rilevante per il caso di dinamica ergodica.

La rilevanza della identificazione (2.3.12) è dovuta al fatto che la forma differenziale δQ del calore ha un ruolo fondamentale in termodinamica fenomenologica. Infatti il secondo pricipio (nella sua prima parte, relativa alle trasformazioni reversibili[39]) afferma che tale forma ammette un fattore integrante (ovvero un fattore che rende la forma differenziale δQ una forma esatta): il fattore integrante è $1/T$ (l'inverso della temperatura assoluta), e la funzione in tal modo definita è l'entropia termodinamica S_{th}, ovvero si ha

$$\frac{1}{T}\,\delta Q = \mathrm{d}S_{th}\ . \tag{2.3.15}$$

Il confronto delle relazioni (2.3.12) e (2.3.15) mostra che, a meno di un fattore costante, si debba identificare β con $1/T$ e contemporaneamente S con l'entropia termodinamica S_{th}. Precisamente, la nota espressione termodinamica dell'entropia del gas perfetto mostra che si devono avere le identificazioni

$$\beta = \frac{1}{k_B T}\ , \qquad \log W = \frac{S_{th}}{k_B} \tag{2.3.16}$$

dove k_B è la costante di Boltzmann. Denotando con S l'entropia termodinamica S_{th}, la seconda di tali identificazioni si scrive nella forma

$$S = k_B \log W\ , \tag{2.3.17}$$

che costituisce la celebre relazione di Boltzmann, incisa sulla lapide della sua tomba. [40]

Intermezzo: La forma differenziale del calore nella meccanica statistica "alla Boltzmann"

Ricordiamo come, nella meccanica statistica "alla Boltzmann", si giustifica l'espressione (2.3.14) per il calore δQ fornito al sistema studiato, in una trasformazione reversibile. Si tratta di una osservazione banalissima, di cui tuttavia abbiamo trovato conferma solo nell'agile libretto di Schrödinger.

A tal fine ritorniamo alla definizione dell'energia interna

$$U = \sum \epsilon_k n_k\ ,$$

[39] La seconda parte invece afferma che nelle trasformazioni irreversibili l'entropia cresce.

[40] Questa formula di Boltzmann venne a costituire un cavallo di battaglia per Einstein. Questi ne diede una rideduzione interessante nel suo secondo articolo di termodinamica statistica (quello del 1904) in cui, denotando con $\Omega(E)$ il volume della regione dello spazio Γ racchiusa entro la superficie di energia E, egli trovò

$$\Omega(E) = C\ e^{S(E)/k_B}\ .$$

e consideriamo un processo elementare (ovvero, con incrementi infinitesimi) in cui siano incrementati sia i numeri di occupazione n_k , sia le energie ϵ_k delle celle, cioè si abbia

$$\mathrm{d}U = \sum \epsilon_k \mathrm{d}n_k + \sum n_k \mathrm{d}\epsilon_k \ .$$

Il secondo termine deve essere identificato come il lavoro compiuto *sul* sistema, ottenuto modificando un parametro esterno (tipicamente il volume V in cui è cintenuto il sistema). Si deve pensare che le energie delle celle dipendano ad esempio dal volume, $\epsilon_k = \epsilon_k(V)$, sicché una variazione $\mathrm{d}V$ del volume in cui è contenuto il sistema produca il lavoro esterno $-p\mathrm{d}V$ (si veda Khinchin o Schrödinger). Se, come di solito in termodinamica, si fa riferimento al lavoro compiuto *dal* sistema studiato, denotandolo con δL, avremo allora la identificazione

$$\delta L = -\sum n_k \mathrm{d}\epsilon_k \ .$$

Allora ritroveremo il primo principio se identificheremo il calore δQ trasferito *al* sistema come

$$\delta Q = \sum \epsilon_k \mathrm{d}n_k \ ,$$

come d'altra parte è ben spontaneo fare. Infatti, l'espressione di δQ descrive un incremento di energia in cui vengono alterati soltanto i numeri di occupazione, ad esempio aumentando le popolazioni delle celle con alta energia e diminuendo le popolazioni di quelle con bassa energia. È pertanto giustificata l'espressione elementare del calore δQ che abbiamo usato nelle pagine precedenti.

Il caso particolare della statistica di Maxwell–Boltzmann

Torniamo ora alla statistica di MB, in cui la probabilità di uno stato macroscopico $W(n_1, n_2, \ldots)$ è proporzionale al corrispondente volume dello spazio Γ, e dunque si fa riferimento al particolare funzionale $S(p_1, p_2, \ldots) = -N \sum p_k \log p_k$ e al corrispondente punto di massimo dato dalla distribuzione di MB, ovvero $p_k(\beta) = \exp(-\beta\epsilon_k)/\mathscr{Z}(\beta)$. Abbiamo visto che, nel caso di Maxwell–Boltzmann l'energia termodinamica U viene espressa in termini di β dalla formula (2.3.7). Osserviamo ora che essa si scrive anche nella forma significativa[41]

$$U = -N \frac{\partial}{\partial \beta} \log \mathscr{Z}(\beta) \equiv -N \partial_\beta \log \mathscr{Z}(\beta) \ , \tag{2.3.18}$$

Inoltre, conoscendo la identificazione dell'entropia, possiamo mostrare che si ha anche l'identificazione delle energia libera F, mediante la formula

$$-\beta F = N \log \mathscr{Z} \ , \quad \text{ovvero} \quad \mathscr{Z} = e^{-\beta F/N} \ . \tag{2.3.19}$$

[41] Basta ricordare la formula per la derivata del logaritmo, e scambiare la derivata con la somma su k. Ammettendo che tale scambio sia possibile. Ciò non è consentito, ad esempio, nel caso dell'atomo di idrogeno, in cui addirittura si ha che la serie definente Z diverge.

Questo si verifica immediatamente valutando la funzione $S = -N \sum p_k \log p_k$ nel punto di massimo $p_k = \exp(-\beta\epsilon_k)/\mathscr{Z}$. Infatti, usando $U = N \sum p_k \epsilon_k$ e $\sum p_k = 1$, si ha

$$S(p_1, p_2, \ldots) = N \sum p_k \big(\beta\epsilon_k + \log \mathscr{Z}\big) = \beta U + N \log \mathscr{Z}$$

e quindi

$$N \log \mathscr{Z} = S - \beta U \ ,$$

che fornisce subito la (2.3.19) in virtù della nota definizione $F = U - TS$, ovviamente equivalente a $\beta F = \beta U - S/k_B$.[42] [43] [44]

2.3.5 Note complementari

Boltzmann e i fondamenti della meccanica statistica di non equilibrio

La visione di Boltzmann è allora la seguente. Un sistema potrebbe a priori trovarsi in un punto qualsiasi dello spazio delle fasi Γ (o di una definita superficie di energia costante, se è conosciuta l'energia). Ma, come abbiamo visto, lo stato di MB ha probabilità enormemente grande. E, poiché tale probabilità è definita in termini di volume nello spazio delle fasi, ciò significa che la regione corrispondente allo stato di MB (è questo il cosiddetto "mare di Boltzmann") occupa sostanzialmente tutta la superficie dell'energia. Se invece il dato iniziale corrispondesse a uno stato macroscopico diverso, la dinamica lo porterebbe prima o

[42] La (2.3.19) è compatibile con la (2.3.18), perché in termodinamica è noto ch si ha

$$U = \partial_\beta(\beta F) \, .$$

Infatti, per le familiari proprietà della trasformata di Legendre, da $F = U - TS$ e da $\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V$ (che esprime insieme il primo e il secondo principio) si ottiene $\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V$ e quindi $S = -\partial_T F$, e infine $F = U + T\partial_T F$. D'altra parte esplicitiamo ora l'equazione di Helmholtz rispetto ad U, ed osserviamo, come subito si verifica, che vale $T\partial_T = -\beta\partial_\beta$. Dunque si ha

$$U = F - T\partial_T F = F + \beta\partial_\beta F = \partial_\beta(\beta F) \, .$$

[43] Ricordiamo che la relazione

$$F = U + T\partial_T F \ ,$$

costituisce la cosiddetta *equazione di Helmholtz*, che era considerata da Nernst *l'equazione fondamentale della termodinamica*. La ragione è che essa si presenta come una equazione differenziale del primo ordine per l'energia libera F, e quindi permette di calcolare F quando sia nota l'energia interna U (e sia assegnato il valore iniziale di F, che può essere preso nullo per $T = 0$, alla luce del terzo principio di Nenrst.) Dunque da dati termici (ovvero dalla funzione U, che è conosciuta attraverso misurazioni del calore specifico) viene determinata F che, in base al secondo principio, coincide con il lavoro massimo ottenibile in una trasformazione isoterma, e tale lavoro è la quantità di interesse massimo per le macchine termodinamiche.

[44] NOTA PER GLI AUTORI FARE I TRE ESERCIZI DEL CAPITOLO (UNA PAGINA) DI SCHRODINGER PER LA TERMODINAMIA DI GAS PERFETTO , ODCILLATORE ARMONICO E OSCILLATORE DI FERMI

poi nello stato di MB.[45] Qui allora Boltzmann compie un salto, riguardante l'entropia degli stati macroscopici che non sono stati di equilibrio generici, perché concepisce di definire una entropia, e più in generale anche una termodinamica, anche per gli stati macroscopici diversi da quello di MB, di massima entropia. L'entropia di tali stati sarebbe nient'altro che il corrispondente valore di S. In questo senso egli è il fondatore della meccanica statistica di non equilibrio.

La trattazione di questi problemi, nei lavori originali di Boltzmann pubblicati su riviste, è di lettura non semplice (anche Maxwell se ne lamentava, in maniera esplicita). Una rassegna (ancora di lettura non agevolissima) è stata data recentemente da Gallavotti.[46] Tuttavia, una trattazione più compatta e decisamente più abbordabile è stata data da Boltzmann stesso nel suo libro sulla teoria dei gas. Particolarmente rilevante ai nostri fini è la discussione data nel primo capitolo. Nel paragrafo 6 (pag. 55 e seguenti della traduzione inglese pubblicata da Dover), vengono date le formule probabilistiche che noi abbiamo riportato sopra. Poi, nel paragrafo 8 (pag. 68 e seguenti), dedicato al calore specifico di una gas, viene discusso il significato matematico della sua celebre funzione H (la nostra s, a parte il segno e il fattore k_B), e viene mostrato che, nello stato di equilibrio di MB, nel caso del gas perfetto il logaritmo della probabilità dello stato coincide (a meno di una costante moltiplicativa) con l'entropia termodinamica del gas, scritta in funzione di densità ρ e temperatura T (pag. 74). Poi, sulla base del fatto che lo stato di equilibrio è lo stato di massima probabilità, egli viene a congetturare che gli stati di non equilibrio evolvano spontaneamente in modo da portarsi in stati di probabilità crescente, fino a giungere eventualmente allo stato di equilibrio, come stato di massima probabilità. Infine conclude dicendo (pag. 75):

"*In one respect we have even generalized the entropy principle here, in that we have been able to define the entropy in a gas that is not in a stationary state*".

Il mare di Boltzmann e il problema di Fermi Pasta ed Ulam

Abbiamo mostrato come, a fissata energia totale E (e a fissato numero di sottosistemi N, grande), l'insieme degli stati microscopici cui corrispondono stati macroscopici (successioni $\{n_k\}$ di numeri di occupazione) diversi da quelli di MB abbia misura estremamente piccola, sostanzialmente nulla. La stragrande maggioranza dei punti corrisponde allo stato di MB. Si usa dire che tale regione, che prende sostanzialmente tutta la superficie di energia E nello spazio Γ, costitui-

[45]Si potrebbe credere che il tempo richiesto debba essere piccolo, come in effetti avviene per i gas diluiti, ad alte temperature. Ma Boltzmann inisteva continuamente che ciò non è vero in generale, portando tipicamente l'esempio della sfera perfettamente liscia. Nel 1954 il lavoro di Fermi Pasta ed Ulam del 1954 ha mostrato che il tempo impiegato potrebbe essere lungo nel caso di oscillatori accoppiati.Inoltre, nel primo capitolo del loro libro di Meccanica Statisitca, Landau e Lifshitz riportano esempi di reazioni chimiche in cui i tempi sono lunghissimi, come avviene anche per i vetri.

[46]G. Gallavotti, *Ergodicity: a historical perspective. Equilibrium and Nonequilibrium*, Eur. Phys. J. .. .

sce il *mare di Boltzmann*. Prendendo "a caso" un punto su quella superficie di energia si cade sostanzialmente sempre nel mare di Boltzmann.

Tuttavia è possibile assegnare "a mano" un dato iniziale fuori dal mare. Si tratta di una situazione di grande deviazione, come è avvenuto quando abbiamo assegnato a mano un valore dell'energia E finito, mentre secondo la probabilità a priori assegnata nello spazio Γ l'energia media sarebbe infinita. Era comune opinione[47] che se a mano si assegna un dato iniziale fuori dal mare di Boltzmann, allora rapidamente la dinamica ricondurrà l'evoluto temporale dentro il mare di Boltzmann (ma abbiamo ricordato che non era questa l'opinione di Boltzmann stesso, facendo tipicamente riferimento al suo esempio prediletto della sfera liscia). Costituì allora una grande scoperta quella fatta da FPU nel 1954. Come illustrato nel Capitolo 1, per una catena di particelle con interazioni non lineari a primi vicini, essi assegnarono condizioni iniziali con energia solo su pochi modi normali di bassa frequenza. Questo è un dato iniziale completamente fuori dal mare di Boltzmann, perché come vedremo nella prossima Sezione, per tutti i punti del mare l'energia è in media ugualmente distribuita tra tutti i modi normali (appunto secondo il principo di equipartizione dell'energia). Ci si sarebbe attesi che l'evoluzione dinamica conducesse rapidamente il sistema nel mare, e quindi all'equipartizione. Si trova invece (Izrailev e Chirikov) che ciò avviene rapidamente a energia specifica E/N abbastanza alta, mentre per energia specifica bassa le energie medie (in media temporale) si distribuiscono solo su un pacchetto di modi di bassa frequenza. Solo dopo un tempo alquanto lungo il sistema rilassa poi ad uno stato di apparente equilibrio, con una approssimata equipartizione dell'energia dei modi. Tuttavia è ancora un problema aperto stabilire se in un tale stato si possa applicare la statistica di MB, o invece si abbia una diversa termodinamica statistica. Il problema è che la diatribuzione di MB riguarda la probabilià delle energie dei singoli sottosistemi, e non dice assolutamente sulla sulla probabiltà congiunta delle energie di due o piú sottosistemi..

È chiaro che qui il problema ergodico è particolarmente acuto. Infatti man mano che si abbassa l'energia specifica il sistema si approssima sempre più a un sistema di oscillatori armonici disaccoppiati, che presenta N costanti del moto (le energie dei singoli modi normali). e quindi nello spazio delle fasi si presentano delle barriere sempre più insormontabili. Ma nessuno ancora è stato in grado di descrivere in maniera chiara cosa ciò comporta, ai fini della meccanica statistica.

Nota critica sulla interpretazione dei numeri di occupazione

Se ogni particella occupasse la cella k–esima con probabilitá p_k assegnata a priori, in maniera indipendente dalle altre particelle, si avrebbe $n_k \simeq N p_k$. Questo porta a pensare che viceversa si possa definire $p_k = n_k/N$ come la probabilità che un singolo sottosistema si trovi nello stato microscopico k–esimo (quello che diventerà un "livello di energia", nel caso in cui le celle individuino delle

[47] NOTA PER GLI AUTORI. Controllare il libro di Ford e Uhlenbeck.

"strisce" di energia). Naturalmente questo ragionamento non è corretto, perché il processo di occupazione non é indipendente, avendo noi fissato l'energia totale.

La domanda è però lecita, e domandiamoci dunque se sia possibile calcolare quanto vale la probabilità che una **fissata** molecola stia nella cella k–esima. Questo è un calcolo classico nella teoria delle probabilità, che viene chiamato *calcolo di una probabilità marginale*, e richiede però di conoscere la probabilità nello spazio delle fasi Γ del sistema totale. In particolare, se si assume l'equiprobabilità sulla superficie di energia fissata, il calcolo si trova nel libro di Khinchin, ed il risultato è proprio quello che ci si aspetta, ovvero

$$p_k = \frac{e^{-\beta\epsilon_k}}{\mathcal{Z}(\beta)} \, .$$

D'altra parte, come abbiamo discusso in questo capitolo e in quello su FPU, l'equiprobabilità a priori è garantita se la dinamica del sistema è ergodica. Se il sistema non è ergodico, niente si può dire su questo problema.

Deduzione della distribuzione di MB per via termodinamica

Nel lavoro di Einstein del 1917 di cui parleremo più avanti, in cui egli fa uso della distribuzione di Maxwell–Boltmann (formula (5) del lavoro), si trova la frase seguente: *"La formula (5) può essere ricavata dal principio di Boltzmann, oppure per via puramente termodinamica."*[48] Questo è proprio vero, ed è un fatto pochissimo conosciuto. Il procedimento è il medesimo che conduce alla deduzione della cosiddetta *legge di azione di massa* (la nota legge degli equilibri chimici), in cui si mostra che il rapporto tra il prodotto delle concentrazioni dei prodotti di una reazione e il prodotto delle concentrazioni dei reagenti è una costante (a fissata temperatura). La dimostrazione della legge degli equilibri chimici è riportata ad esempio nel libro di termodinamica di Fermi, e fa uso della cosiddetta *Van't Hoff reaction box*.[49] L'analogia con la distribuzione di Maxwell–Boltzmann è la seguente. Consideriamo ad esempio il caso semplice in cui abbiamo un atomo che può esistere in diversi livelli energetici E_n. In tal modo gli atomi che si trovano in un certo livello sono considerati come una ben definita specie chimica, e si hanno tante specie chimiche quanti sono i livelli. Allora il passaggio da un livello a un altro viene considerato come la reazione che conduce da una specie chimica ad un'altra, e si applica la legge delle reazioni chimiche (o legge di azione di massa).

Si noti che il medesimo procedimento cui accenna Einstein nel suo articolo del 1917 venne esplicitamente seguito da Fermi nel suo articolo del 1923, dal titolo *Sulla probabilità degli stati quantici*, riportato come contributo n. 17a, a pag. 118 delle *Note e Memorie*, Vol. I. In tale articolo egli mostra come si possa superare

[48] Per una analogo procedimento dovuto a Kramers, al fine di giustificare la distribuzione di Bose–Einstein, si veda Whittaker, *A history of the theories of ether and electricity*.

[49] E. Fermi, *Thermodynamics*, Dover (New York, 1956), pag. 101.

la difficoltà relativa al fatto che la funzione di partizione dell'atomo di idrogeno è espressa da una serie divergente.

Due Osservazion sulla formula di Boltzmann (2.3.3) per la probabilitá di uno stato macroscopico

La costante additiva dell'entropia e i lavori di Sackur e Tetrode. Quando si definisce la probabilità di uno stato macrorscopico $\{n_k\}$ mediante il volume della corrispondente regione nello spazio delle fasi Γ, resta indefinito un fattore moltiplicativo nella formula fondamentale (2.3.3) della statistica di Boltzmann. Questo fattore produce nell'entropia una costante additiva, che a prima vista potrebbe apparire irrilevante.

Invece nel 1911 fu messo in luce da Sackur e Tetrode, che questa costante additiva dell'entropia che ha un significato fondamentale per le reazioni chimiche, perché viene ad entrare nella legge di "azione di massa" che determina gli equilibri chimici. A tal punto che, dai valori empirici noti che entrano in tali formule, quei due autori furono in grado di trovare una stima empirica per la costante di Planck, del tutto indipendente dalle determinazioni empiriche che si ottengono dalle leggi fenomenologiche del corpo nero. Inoltre, avendo compreso la rilevanza del volumetto elementare, essi ne diedero una interpretazione che anticipa la relazione di indeterminazione di Heisenberg. Dato che importa solo la misura di Lebesgue del volumetto, e non la sua forma, questo significa (ad esempio nel caso di un solo grado di libertà) che il volumetto può avere base δq arbitraria e altezza δp arbitraria, purché per l'area si abbia il vincolo $\delta q \cdot \delta p = \hbar$. Infine facciamo notare che i lavori di Sackur e Tetrode ebbero notevole importanza per il lavoro del 1926 in cui Fermi introdusse quella che ora si chiama la statistica di Fermi–Dirac.

2. Il fattore di normalizzazione di Boltzmann, e la formula usata da Planck nel suo lavoro del 14 dicembre 1900. Abbiamo già osservato che la probabilità W di uno stato macroscopico evidentemente non è normalizzata, perché lo spazio delle fasi totale Γ ha misura di Lebesgue infinita. Invece, in generale è finita la misura di una superficie Σ_E, indotta dall misura di Lebesgue nello spazio Γ.[50] In tale caso il fattore di normalizzazione relativo alla (2.3.3) determinato dalla condizione di energia fissata fu calcolato da Boltzmann stesso, e stranamente esso ha un aspetto strettamente parente di quello che presenta la formula usata la Planck il 14 dicembre 1900 nella deduzione teorica che egli diede della sua formula.

[50] Ad esempio, per una circonferenza la misura di Lebesgue nel piano assegna misura nulla, mentre la "misura indotta" é $2\pi r$.

2.4 La distribuzione di Maxwell–Boltzmann nel limite del continuo, e il teorema di equipartizione dell'energia

Passando al continuo (è proprio questo il passaggio che verrà poi eliminato da Planck nella sua seconda comunicazione, e poi specialmente da Einstein nel suo lavoro sui calori specifici), in luogo dei numeri di occupazione n_k si ottiene una densità nello spazio μ (numero di sottosistemi per unità di volume nello spazio μ). Questa viene denotata da Boltzmann con la lettera f. La densità di massima probabilità è allora data da

$$f(q,p) = N\frac{e^{-\beta H(q,p)}}{\mathscr{Z}(\beta)} \tag{2.4.1}$$

con

$$\mathscr{Z}(\beta) = \int e^{-\beta H(q,p)}\,\mathrm{d}q\,\mathrm{d}p\ . \tag{2.4.2}$$

In analogia con il caso discreto si ha allora

$$\frac{U}{N} = \int H(q,p) f^*(q,p)\mathrm{d}q\mathrm{d}p = \frac{\int H(q,p)e^{-\beta E(q,p)}\,\mathrm{d}q\,\mathrm{d}p}{\int e^{-\beta E(q,p)}\,\mathrm{d}q\,\mathrm{d}p}$$

ovvero

$$\frac{U}{N} = -\frac{\partial}{\partial\beta}\log\mathscr{Z}\ . \tag{2.4.3}$$

Il teorema di equipartizione, o di equiripartizione, dell'energia

Il risultato più rilevante della teoria cinetica è il teorema di equipartizione. Lo discutiamo qui nella sua forma più semplice possibile. Succede che per una grandissima categoria di sistemi (pensati come costituiti di un enorme numero di sottosistemi identici) l'hamiltoniana di un singolo sottosistema è la somma di un certo numero l di termini quadratici nelle posizioni o nei momenti, con coefficienti costanti. Ad esempio, nel caso della particella libera e in quello dell'oscillatore armonico monodimensionale (con equazione $m\ddot{x} = -kx$), che sono quelli che maggiormente ci interesseranno, si ha rispettivamente[51]

$$H = \frac{1}{2m}\left(p_x^2 + p_v^2 + p_z^2\right) \quad , \quad H = \frac{1}{2m}p^2 + \frac{1}{2}kx^2\ .$$

[51] Per l'oscillatore armonico, si può procedere anche nel modo seguente. Si introduce la frequenza ω mediante $\omega^2 = \varkappa/m$. Poi nell'hamiltonianasi effettua la fattorizzazione di ω in maniera da potere eseguire una trasformazione canonica $(x,p) \to (Q,P) = (x\sqrt{m\omega}, p/\sqrt{m\omega})$, perchè allora x e p risultano rispettivamente moltiplicati e divisi per il medesimo fattore, $\sqrt{m\omega}$, quindi la trasformazione é canonica, e l'hamltoniana diviene

$$H = \frac{\omega}{2}\left(\frac{p^2}{m\omega} + m\omega x^2\right) = \frac{\omega}{2}(P^2 + Q^2)\ .$$

Nel caso della molecola monoatomica non puntiforme bisognerebbe tenere conto dei gradi di libertà rotazionali che entrano nell'energia cinetica e hanno una struttura più complicata, come anche i gradi di libertà rotazionali della molecola biatomica (invece, la "molla" che descrive le oscillazioni della distanza tra gli atomi può essere descritta come un oscillatore armonico).

Dunque l'energia meccanica E del sistema totale risulta decomposta nella somma di l contributi $E = E^{(1)} + E^{(2)} + \ldots + E^{(l)}$, e possiamo pensare che in maniera analoga si decomponga anche l'energia termodinamica U,

$$U = U^{(1)} + U^{(2)} + \ldots + U^{(l)} \tag{2.4.4}$$

Si ha allora la

Proposizione 1 (Teorema di equipartizione dell'energia. Prima parte.) *Si consideri un sistema costituito di N sottosistemi identici. L'hamiltoniana di ogni sottosistema sia la somma di l termini quadratici nelle posizioni o nei momenti, e sia*

$$U = \sum_{j=1}^{l} U^{(j)}$$

la corrispondente decomposizione dell'energia termodinamica U. Se il sistema si trova nello stato di Maxwell–Boltzmann (nel continuo) allora tutti gli l termini $U^{(j)}$ in cui si decompone l'energia totale contribuiscono per la medesima quantità (equipartizione !), data da

$$U^{(j)} = N \frac{1}{2\beta}\,, \quad j = 1, \ldots l\,. \tag{2.4.5}$$

Dimostrazione

La dimostrazione è elementare. La difficoltà consiste solo nel trovare notazioni generali ma semplici. Sceglieremo di trattare un caso particolare, considerando dunque un sistema costituito di N oscillatori armonici identici, ciascuno con hamiltoniana che scriviamo nella forma

$$H(x, p) = a p^2 + b x^2\,.$$

All'energia totale E o equivalentemente all'energia totale termodinamica U contribuiscono due termini, quello corrispondente al termine di energia cinetica $a p^2$ in H e quello corrispondente al termine di energia potenziale $b q^2$ in H. Li denoteremo con $U^{(kin)}$, $U^{(pot)}$.

Basta ora calcolare l'energia totale termodinamica secondo la formula generale (2.4.3). Si osservi che la funzione di partizione, che coinvolge l'hamiltoniana H di un sottosistema, si fattorizza, con un fattore in corrispondenza di ogni termine quadratico dell'hamiltoniana:

$$\mathcal{Z} = \mathcal{Z}^{(kin)} \cdot \mathcal{Z}^{(pot)}\,,$$

$$\mathcal{Z}^{(kin)} = \int e^{-\beta a p^2} \mathrm{d}p \ , \quad \mathcal{Z}^{(pot)} = \int e^{-\beta b q^2} \mathrm{d}q \ ,$$

e dunque si ha $U = U^{(kin)} + U^{(pot)}$ con

$$U^{(kin)} = -N \frac{\partial}{\partial \beta} \log \mathcal{Z}^{(kin)} \ , \quad U^{(pot)} = -N \frac{\partial}{\partial \beta} \log \mathcal{Z}^{(pot)} \ ,$$

Ma ad esempio si ha[52]

$$\mathcal{Z}^{(kin)} = \int e^{-\beta a p^2} \mathrm{d}p = c \, \frac{1}{\sqrt{\beta}}$$

dove c è una costante[53] indipendente da β. Qui sta tutto il "trucco", perché dobbiamo prendere il logaritmo, ed eseguire la derivata rispetto a β, sicché le costanti moltiplicative indipendenti da β sono irrilevanti. Si noti che nell'eseguire la derivata del logaritmo sparisce anche la dipendenza dalla costante a che entra in c, e caratterizza le diverse specie chimiche. Si trova in tal modo $U^{(kin)} = N/(2\beta)$. Evidentemente il calcolo per il termine potenziale è assolutamente identico.

Non è difficile dimostrare che il medesimo risultato vale anche nel caso dei termini quadratici che descrivono le parti rotazionali dell'energia cinetica[54]

Complementi: il teorema di equipartizione con il procedimento di Einstein. Seguiremo qui, per dimostrare il teorema di equipartizione, il procedimento che sembra essere quello preferito da Einstein, come si vede ad esempio leggendo il suo lavoro sui calori specifici.

Abbiamo evidentemente

$$\frac{U^{(kin)}}{N} = \frac{\int\int a p^2 e^{-\beta\left(a p^2 + b x^2\right)} \, \mathrm{d}x \mathrm{d}p}{\int\int e^{-\beta\left(a p^2 + b x^2\right)} \, \mathrm{d}x \mathrm{d}p} = \frac{\int a p^2 e^{-\beta a p^2} \, \mathrm{d}p}{\int e^{-\beta a p^2} \, \mathrm{d}p}$$

dove i due integrali sono entrambi estesi a tutto l'asse reale; scrivendoli come integrali ripetuti, un integrale si elimina tra numeratore e denominatore. Dunque, estraendo il *fattore dimensionale* $1/\beta$, si ha

$$\frac{U^{(kin)}}{N} = \frac{1}{\beta} \, \frac{\int t^2 e^{-t^2} \, \mathrm{d}t}{\int e^{-t^2} \, \mathrm{d}t} \ .$$

[52] Si compie il cambiamento di variabili $\sqrt{\beta a}\, p = t$. Si usa dire che in tal modo "si adimensionalizza la variabile di integrazione.

[53] $c = (\int e^{-t^2} \mathrm{d}t)/\sqrt{a}$.

[54] Il metodo più semplice è indicato nel libro G. H Wannier, *Statistical Physics*, Dover (New York, 1966), pag. 76. Si osserva che in ogni caso il momento p entra in maniera quadratica, e il fattore a, invece di essere costante, dipende da una variabile q, un angolo. Ancora si usa la legge di Maxwell–Boltzmann, e si eseguono gli integrali come integrali "ripetuti", ciascuno rispetto ad una delle variabili. Allora eseguendo l'integrale su p il fattore $a(q)$ "esce" dall'integrale, e si semplifica con l'analogo fattore al denominatore, esattamente come nel caso di a costante.

Già questo costituisce il risultato principale, perché il fattore a, che distingue una specie atomica da un'altra, è stato eliminato. Inoltre, scambiando il ruolo di p e q, allo stesso modo si ottiene anche

$$\frac{U^{(pot)}}{N} = \frac{1}{\beta} \frac{\int t^2 e^{-t^2}\, \mathrm{d}t}{\int e^{-t^2}\, \mathrm{d}t} ,$$

e quindi abbiamo già ottenuto il teorema di equipartizione (i due termini portano il medesimo contributo all'energia totale).

Comunque si determina subito il fattore adimensionale che resta da calcolare, e si trova

$$\frac{\int t^2 e^{-t^2}\, \mathrm{d}t}{\int e^{-t^2} \mathrm{d}t} = \frac{1}{2} .$$

come si vede eseguendo al denominatore una integrazione per parti (nella quale il termine finito si annulla) e si trova che il denominatore è il doppio del numeratore.[55]

Il problema dell'identificazione della temperatura. Il "principio" di equipartizione

Ora, il teorema di equipartizione induce ad introdurre la seconda identificazione fondamentale (dopo la identificazione $E \equiv U$) tra quantità termodinamiche e quantità meccaniche, ovvero $\beta = 1/k_B T$. A tal punto che tale identificazione viene comunemente intesa come parte del teorema di equipartizione. Ci pare più significativo procedere invece in maniera diversa, mettemdo di evidenza che questa seconda identificazione costituisce piuttosto un assioma, la cui validità può essere discutibile:

Principio di identificazione della temperatura. Si ha la seconda identificazione fondamentale

$$\beta = \frac{1}{k_B T} . \tag{2.4.6}$$

Si giunge in tal modo al

Corollario 1 (Principio di equipartizione) . *Si ha*

$$U^{(j)} = N \frac{k_B T}{2} , \quad j = 1, \dots l . \tag{2.4.7}$$

Il principio di equipartizione conduce immediatamente al secondo risultato di interesse fondmentale per la fisica.

Corollario (Calori specifici nello stato di MB). *Si consideri un sistema costituito di N sottosistemi identici, e l'hamiltoniana di ogni sottosistema sia la somma di l termini quadratici nelle posizioni o nei momenti (a coefficienti costanti). Se il*

[55] Si ha

$$\int e^{-t^2} \mathrm{d}t = -\int t\, \mathrm{d}e^{-t^2} = 2 \int t^2 e^{-t^2}\, \mathrm{d}t .$$

sistema si trova nello stato di Maxwell–Boltzmann (nel continuo), allora l'energia termodinamica $U(T)$ in funzione della temperatura è data da

$$U(T) = \frac{l}{2} N k_B T \ . \tag{2.4.8}$$

Pertanto la capacità termica $\partial_T U$ e il calore specifico C_V (capacità termica di una mole) sono dati da

$$\frac{\partial U}{\partial T}(T) = \frac{l}{2} N k_B \ , \quad \text{ovvero} \quad C_V(T) = \frac{l}{2} R \ , \tag{2.4.9}$$

dove $R = N_A k_B$ è la costante dei gas.

Vedremo tuttavia che i calori specifici predetti dal principio di equipartizione corrispondono ai dati sperimentali solo parzialmente, e anzi sono qualitativamente inadeguati per basse temperature e/o alte frequenze. Questo è appunto il fatto che darà origine alla crisi che condurrà all'introduzione della MQ.

Nota critica sul principio di identificazione della temperatura e sul principio di equipartizoine

L'assioma di identificazione della temperatura ha una sua ragionevolezza. Infatti, nel caso del gas il contributo dei baricentri all'hamiltoniana H di una singola molecola è dato da 3 termini quadratici, sicché in base al teorema di equipartizione il corrispondente contributo $U^{(cm)}$ all'energia termodinamica U è dato da

$$U^{(cm)} = N \frac{3}{2\beta} \ .$$

D'altra parte, in base al teorema di Clausius si ha

$$U^{(cm)} = \frac{3}{2} N k_B T \ .$$

e quindi l'identificazione (2.4.6).

Tuttavia, la validità dell'identificazione nel caso generale (tipicamente nel caso di un solido, o di un sistema di oscillatori armonici debolmente accoppiati) è tutt'altro che ovvia, ed anzi, a nostro parere, completamente ingiustificata. Il pumto cruciale, che abbiamo messo in luce precdentemente, è la validità stessa della statistica di Boltzmann, che appare giustificata solo per sistemi per cui valga l'ipotesi ergodica. Ora, tale ipotesi sembra esere giustificata nel caso dei gas dituiti ad alte temperature, ma evidentemente fallisce completamente in casi generici, tipicamente nel caso dei solidi. Infatti, mentre è spontaneo considerare come indipendenti i moti dei singoli atomi di un gas, ciò ovviamente non può dirsi per i costituenti di un solido: ad esempio, in un solido ionico gli ioni non si scambiano di posiione con ioni lontani, e neanche con quelli vicini, tranne che nel caso in cui si verifichi la fusione, il che avviene a temperature alte. E

comunque, quando avvengono tali scambi non si ha più a che fare con un solido, ma con un liquido. Ma anche quando si trascurasse questo aspetto, nei modelli di tipo FPU si vede bene che anche i modi normali non si comportano come indipendenti, per tempi estremamente lunghi. Questo fatto è stato dimostrato recentissimamente con le più sofisticate tecniche moderne di teoria dei sistemi dinamici, da Alberto Maiocchi[56]. Inoltre, è stato messo in luce, attraverso calcoli numerici degli spettri infrarossi dei cristalli ionici, che a basse energie la temperatura non deve essere identificata come un mutiplo dell'energia cinetica, ovvero non è valido il familiare principio di identificazione della temperatura. [57]

Difficoltà del principio di equipartizione per le molecole

Per le molecole monoatomiche, nell'approssimazione in cui si trascura il contributo all'energia termodinamica (o energia interna) dovuto alle rotazioni, il calore specifico teorico è $(3/2)R$, indipendente dal numero atomico. Questo valore è ben verificato per molecole monoatomiche, ma il fatto strano è che esso è ben verificato anche in casi in cui sembrerebbe più adeguato un modello del tipo della sfera rigida, sicché non si comprende (in ambito classico) perché sia lecito trascurare il contributo dell'energia rotatoria. Abbiamo già ricordato che questo fatto stava molto a cuore a Boltzmann.

In effetti, le verifiche sperimentali furono condotte non direttamente sul valore del calore specifico, perché questo coinvolge la costante dei gas R, il cui valore era conosciuto entro una notevole incertezza. Si preferì allora fare riferimento a una quantità nella quale scompariva R. Questa quantità è il rapporto γ tra calore specifico a pressione costante e quello a volume costante. Infatti, poiché dalla prima legge della termodinamica si ha (per una mole di sostanza)

$$C_p = C_V + R \ ,$$

ricordando che, come appena mostrato, secondo il principio di equipartizione si ha

$$C_V = \frac{R}{2} l \ ,$$

dove l è il numero di termini quadratici che costituiscono l'energia totale di un sottosistema, risulta infine che il principio di equipartizione comporta

$$\gamma \stackrel{\text{def}}{=} \frac{C_p}{C_V} = 1 + \frac{2}{l} \ . \tag{2.4.10}$$

Si noti tra l'altro che evidentemente il risultato è indipendente anche dal numero di moli della sostanza su cui si compie la misura (che non è agevole determinare

[56] A Maiocchi, Comm. Math. Phys., in corso di stampa (2019).

[57] A. Carati, L. Galgani, F. Gangemi. R. Gangemi, *Relaxation times and ergodic properties in a realistic model of an ionic crystal, and the modern form of the FPU problem*, Physica A (2019).

Gas	Values of $\gamma - 1$	
	According to the M.-B. doctrine	by Observation
Air	$\frac{2}{7} = .2857$	.406
H_2	" "	.40
O_2	" "	.41
Cl_2	" "	.32
CO	" "	.39
NO	" "	.39
NO_2	$\frac{1}{6} = .1667$	.30
N_2O	" "	.331
NH_3	$\frac{1}{9} = .1111$	.311

Tabella 2.1: La tavola dei valori di $C_p/C_v - 1$ previsti secondo il principio di equipartizione e quelli effettivamente misurati, come riportata nel celebre lavoro di Kelvin, *Nineteenth-Century Clouds over the Dynamical Theory of Heat and Light*, Phil. Mag. **6**, 1 (1901).

sperimentalamente). Dunque misurando γ si determina l come

$$l = \frac{2}{\gamma - 1} \,. \tag{2.4.11}$$

Secondo il principio di equipartizione, ci si aspetterebbe quindi di trovare valori di $2/(\gamma - 1)$ vicini ad interi.

Invece in generale i risultati sono molto discordanti. Questo si vede dalla tabella 2.1, che è presa dal lavoro originale di Kelvin, in cui i dati sperimentali sono confrontati con le corrispondenti previsioni teoriche. In generale, come commenta Kelvin stesso, i contributi rotatori per le molecole monoatomiche non si manifestano, quelli rotatori per le molecole biatomiche si manifestano, ma non sempre, e in genere non in maniera adeguata. Lo stesso avviene per il termine di energia potenziale nelle molecole biatomiche, ed in generale anche per le molecole poliatomiche.

In conclusione, la situazione può descriversi dicendo che il numero dei termini quadratici efficaci di energia cinetica o potenziale dipende dalla temperatura e in generale non coincide con un intero. Questo numero in generale diminuisce al decrescere della temperatura: alcuni gradi di libertà appaiono come *dormienti*,[58], e non sono efficaci.[59]

[58] Il termine *dormancy* è proprio quello che viene comunemente usato. Si veda G.H. Wannier, *Statisticall physics*, Dover (New York, 1966), pag. 78–80, 114–115, 215, 224, 258.

[59] NOTA PER GLI AUTORI Riportare la bellissima descrizione del libro di Nernst *The new heat theorem*, e il lavoro di Kelvin.

Difficoltà del principio di equipartizione per i solidi

Una situazione analoga si presenta per il calore specifico dei solidi. In effetti, sembrerebbe che l'approccio di Boltzmann non si applichi direttamente al caso dei solidi o dei liquidi, perchè in questo caso gli atomi o le molecole interagiscono fortemente e cadrebbe l'ipotesi in cui è stata dedotta la distribuzione di Maxwell–Boltzmann (cioè l'ipotesi che ci si trovi in presenza di un numero enorme di sottosistemi debolmente interagenti). Tuttavia, si presenta un fatto notevole cui abbiamo già accennato nel primo capitolo in connessione con il problema FPU, e che mostreremo più avanti. Ovvero che nel caso dei solidi, effettueando un cambiamento di variabili ai modi normali del sistema, formalmente ci si riduce dal punto di vista matematico, al caso di sottosistemi debolmente interagenti. In ogni caso, ci si era convinti abbastanza presto che il principio di equipartizione potesse applicarsi anche in questo caso, sopratutto in virtù del fatto che esso forniva una semplice spiegazione della legge di Dulong e Petit. Infatti se si ammette che il contributo termico sia dovuto ad atomi o ioni che oscillano attorno a posizioni di equilibrio, indipendenti l'uno dall'altro, allora, secondo il principio di equipartizione, si avrebbe un calore specifico indipendente dalla temperatura, e proporzionale al numero atomico della sostanza considerata. Questa proprietà fu in effetti osservata sperimentalmente da Dulong e Petit addirittura nella prima metà dell'ottocento (molto prima che si dimostrasse il teorema di equipartizione), e veniva utilizzato perfino per stimare il numero atomico delle sostanze. Ma vi erano delle nubi o ombre, perché questa proprietà non era verificata, a temperatura ambiente, per tutte le sostanze, e inoltre, generalmente, il calore specifico sembrava diminuire al decrescere della temperatura. Anche qui, come per i gas, si avevano dei gradi di libertà *dormienti*.

Ci proponiamo di riportare altrove citazioni da Boltzmann, da Kelvin e da Rayleigh riguardanti queste difficoltà. Per cominciare, tuttavia, preferiamo citare quasi alla lettera quello che dice Einstein nel suo lavoro sui calori specifici del 1907. Vedremo che il problema centrale si presenta in due forme apparentemente diverse, ma in realtà complementari.

Citiamo da Einstein (pag. 194 dell'edizione italiana). Anzitutto egli ricorda il successo del teorema di equipartizione nello spiegare la legge di Dulong e Petit, dicendo: "*Possiamo rappresentarci nel modo più semplice il moto termico nei solidi immaginando che i singoli atomi contenuti nel corpo compiano oscillazioni sinusoidali attorno alle posizioni di equilibrio. In tali ipotesi, applicando la teoria cinetico–molecolare* [ovvero, il principio di equipartizione], *e tenendo conto del fatto che a ogni atomo competono tre gradi di libertà, si deduce che il calore specifico, riferito al grammo equivalente* [cioè ad una mole] *è* $3Rn$*, ovvero, in calorie,* $5.94n$*, dove* n *è il numero di atomi in una molecola.*[60] *Come è noto, questa relazione è soddisfatta con ottima approssimazione per la maggior parte degli elementi e per molti composti allo stato solido (legge di Dulong e Petit, regola di F. Neumann e Kopp).*"

[60] Einstein sta considerando il caso più generale, di molecole non monoatomiche.

Ma poi continua: " *A un esame più attento si incontrano due difficoltà che sembrano limitare fortemente l'applicabilità della teoria molecolare. Anzitutto vi sono elementi (carbonio, boro e silicio) che allo stato solido posseggono, a temperatura ambiente, un calore atomico specifico notevolmente inferiore a* $5.94\,n$. *Inoltre tutti i composti solidi che contengono ossigeno, idrogeno o almeno uno degli elementi sopra menzionati, hanno un calore specifico per mole più piccolo di* $5.94\,n$. *In secondo luogo Drude, riconducendo le frequenze proprie infrarosse a oscillazioni degli atomi (ioni atomici), e le frequenze proprie ultraviolette a oscillazioni di elettroni, ha dimostrato che i fenomeni ottici (la dispersione) portano ad attribuire a ciascuna molecola più masse elementari mobili tra loro indipendenti. Da qui nasce per la teoria cinetico–molecolare del calore una seconda difficoltà significativa, perché il calore specifico dovrebbe superare di parecchio il valore* $5.94n$, *dato che il numero di masse puntiformi mobili per molecola è maggiore del numero dei suoi atomi.*" [61]

Fermiamoci qui. La prima difficoltà è semplice da capirsi, ed è di tipo prevalentemente quantitativo (ma non del tutto): i valori dei calori specifici sono diversi da quelli previsti. Dopo Planck ed Einstein (come vedremo più avanti), oggi sappiamo che questa difficoltà si supera affermando che secondo la meccanica quantistica l'energia media di un oscillatore dipende dalla temperatura secondo una formula che contiene anche la frequenza (la legge di Planck, per l'appunto), e questo sostanzialmente "spiega" i fatti.

Il secondo punto è più sottile, ma presumibilmente è ancora una diversa forma del primo punto. Einstein ricorda che Drude ha fatto presente che nei solidi ci sono non solo gli ioni (che oscillano secondo frequenze infrarosse), ma anche gli elettroni, e che questi, secondo il principio di equipartizione, dovrebbero avere anch'essi la loro parte (il loro *share*) di energia, contribuendo almeno quanto gli ioni. Questa difficoltà addirittura si esaspera nel caso dei metalli, in cui gli elettroni di conduzione devono essere descritti secondo Drude, come un gas di elettroni liberi. Qui sembra che persino il Teorema di Clausius cada in difetto, perché anche la pura energia cinetica di traslazione degli elettroni non entra in nessuna espressione termodinamica. Questa difficoltà, come è ben noto e accenneremo più avanti, venne in seguito superata, nell'ambito della meccanica quantistica, utilizzando per gli elettroni la statistica di Fermi–Dirac.

Si incontrano poi altri problemi analoghi, sostanzialmente in tutti i campi della fisica. In generale, possiamo dire, il problema è di stabilire quale sia il numero di gradi di libertà di un sistema.[62] Per un atomo si devono contare i gradi di libertà rotazionali? Dai dati sperimentali sembra di no. Per una molecola biatomica, i due termini che entrano nell'energia della "molla", relativo alla distanza tra i due atomi, devono essere presi in conto? Anche su questi problemi la meccanica quantistica sembra dare una risposta corretta, come vedremo.[63]

[61] NOTA PER GLI AUTORI. Questi argomenti sono illustrati in modo ancora migliore nel contributo di Einstein alla Conferenza Solvay.

[62] La referenza più significativa è forse quella del lavoro di Boltzmann su *Nature* del 1895.

[63] NOTA PER GLI AUTORI. Resta tuttavia un certo problema di principio, riguardo il numero dei gradi di libertà da assegnare a un sistema. Ad esempio, perché nel calcolo del numero di

Il principio di equipartizione nella teoria dei campi, in particolare nel problema del corpo nero

NOTA DIDATTICA. Questa parte deve essere riscritta. Riportiamo tuttavia una versione preliminare che richiama il passaggio ai modi normali e riporta il contributo di Rayleigh nel suo lavoro del giugno 1900.

Come abbiamo detto più sopra, l'approccio di Boltzmann non si applicherebbe ai solidi, in quanto nei solidi i costituenti (atomi o molecole) interagiscono fortemente tra loro, per cui l'energia totale non si può scrivere come somma delle energie ϵ_k calcolate nello spazio μ, e dunque l'espressione stessa per i numeri di occupazione n_k perde senso. Questo problema può essere aggirato (come abbiamo ricordato nel capitolo sul problema FPU) se si considerano sistemi per cui le equazioni del moto sono lineari, cioè della forma

$$M\ddot{q}_i = -\sum_k \hat{A}_{ik} q_k \ , \tag{2.4.12}$$

dove le matrici $\hat{M} \stackrel{\text{def}}{=} m_i \delta_{ik}$ ed $\hat{A}$ siano simmetriche e definite positive. Allora usando le variabili $\tilde{q}_i = \sum \hat{M}_{ik}^{1/2} q_k$ ci si riduce all'equazioni nella forma standard

$$\ddot{\tilde{q}}_i = -\sum \hat{B}_{ik} \tilde{q}_k \ ,$$

dove $\hat{B}$ è la matrice simmetrica (definita positiva) definita da $\hat{B} \stackrel{\text{def}}{=} \hat{M}^{-1/2} \hat{A} \hat{M}^{-1/2}$. Esiste allora una matrice unitaria $\hat{U}$ che diagonalizza $\hat{B}$, di modo che, definito il cambiamento di variabili $v_i = \sum \hat{U}_{ik} \tilde{q}_k$. Infine le equazioni si scrivono come

$$\ddot{v}_i = -\omega_i^2 v_i \ ,$$

essendo ω_i^2 gli autovalori della matrice $\hat{B}$.[64] Quindi abbiamo ridotto il sistema a somma di oscillatori armonici disaccoppiati, i cosiddetti modi normali di oscillazione. Quindi nel caso dei sistemi lineari lo spazio μ non è lo spazio delle fasi di un sottosistema concreto, qualcosa di fisicamente reale, ma lo spazio delle fasi

gradi di libertà non dobbiamo comprendere anche i nucleoni (protoni e neutroni) che costituiscono i nuclei? Sembra che ai nuclei si attribuisca una temperatura di circa 10^{14}K. Quindi non sono in equilibrio con il baricentro dell'atomo, che ha una temperatura ordinaria. Essi hanno una grandissima energia, ma non la scambiano con noi. Un fatto analogo avviene anche per le due popolazioni in equilibri separati che si incontrano nel caso delle molecole biatomiche (orto idrogeno e para idrogeno). Ricordare la possibile via d'uscita dalle difficoltà secondo Boltzmann e Rayleigh. Inoltre, rilevanza dei tempi di rilassamento, e congettura dei tempi lunghi. Paragrafo non è ancora stato scritto. Si veda A. Carati, L. Galgani, B. Pozzi, *The Problem of the Rate of Thermalization and the Relations between Classical and Quantum Mechanics* , in *Mathematical Models and Methods for Smart Materials*, edited by M. Fabrizio et al., Series of Advances in Mathematics n. 62, World Scientific (Singapore, 2002). Disponibile nelle *home page* degli autori di queste note.

[64]Essendo la matrice $\hat{B}$ definita positiva, i suoi autovalori sono positivi e quindi possono essere pensati come il quadrato di un numero reale.

dei singoli modi normali, che sono combinazioni lineari delle coordinate di tutte le particelle del sistema.

Se consideriamo un solido, possiamo supporre che le molecole che lo costituiscono compiano (piccole) oscillazioni attorno alla posizione di equilibrio. Allore se linearizziamo le equazioni attorno a tale posizione, otteniamo equazioni del tipo (2.4.12) (dove le q_i sono opportune coordinate lagrangiane). In definitiva, passando ai modi normali otterremo che le equazioni di moto si possano riscrivere in termini di modi normali nella forma

$$\ddot{v}_i = -\omega_i^2 v_i + \text{ termini non lineari }.$$

Dunque, se i termini non lineari sono pensati piccoli, le equazioni di moto si scrivono nella forma in cui si può applicare la teoria di Boltzmann, che, come già ampiamente discusso, prevede l'equipartizione nel caso degli oscillatori armonici. Come abbiamo riportato sopra nelle parole di Einstein, l'equipartizione porta alla legge di Dulong e Petit per i calori specifici dei solidi, che è sperimentalmente ben verificata solo per temperature sufficientemente alte, cadendo invece in difetto alle basse temperature.

Il ragionamento appena esposto fallisce invece a tutte le temperature se si considera il campo elettromagnetico. Ricordiamo, ma sarà meglio discusso nella seconda parte di queste note, che le equazioni che reggono il campo elettromagnetico sono lineari, e della forma (2.4.12), se ad esempio le quantità q_i rappresentano i coefficienti dello sviluppo di Fourier dei campi. La differenza con il caso dei solidi è dovuta al fatto che, anche se si considera l'energia per unità di volume (anziché quella totale), le q_i formano una successione attualmente infinita, e non un insieme finito come nel caso dei solidi. Passando ai modi normali, il teorema di equipartizione porta allora al fatto che l'energia di questo insieme infinito di oscillatori è attualmente infinita, e ciò contrasta ovviamente con i risultati sperimentali.

Questo è il tipo di approccio usato da Rayleigh nel suo articolo di due pagine e mezzo del giugno 1900,[65] per discutere appunto il problema del corpo nero, cioè il problema della radiazione termica. Egli inizia con il considerare i gradi di libertà del campo elettromagnetico, pensato come un sistema analogo alla corda vibrante, la cui energia può essere decomposta formalmente, come ricordato più sopra, in energie di oscillatori armonici (i modi normali di oscillazione), sicché ad ognuno di tali oscillatori (formali, o virtuali, come potremmo dire) dovrebbe essere assegnata, secondo il teorema di equipartizione, una energia media $k_B T$. Ovvero, come egli dice, "*According to the Boltzmann–Maxwell doctrine of the partition of energy, every mode of vibration should be alike favoured.*". Continua poi, calcolando il numero di modi normali che hanno frequenza di oscillazione compresa tra ν e $\nu + d\nu$, di modo che la densità di energia u_ν del campo elettromagnetico di frequenza ν (si veda il capitolo seguente), sarebbe data dalla cosiddetta

[65] Rayleigh, *Remarks upon the law of complete radiation*, Phil. Mag. **49**. 539 (1900), Opere Vol. 5, n. 260, pag. 259.

legge di Rayleigh–Jeans

$$u_\nu(T) = \frac{8\pi\nu^2}{c^3} k_B T \ . \tag{2.4.13}$$

Tale legge è palesemente assurda, perché non presenta decadimento alle alte frequenze ed addirittura produce una densità di energia (energia per unità di volume) infinita. Dunque, tenendo un atteggiamento pragmatico, Rayleigh prende atto che le alte frequenze decadono esponenzialmente (è la cosidetta legge di decadimento di Wien per le alte frequenze che discuteremo nel prossimo capitolo), e si limita ad ammettere che la legge di equipartizione sia adeguata solo nel caso asintotico delle basse frequenze (e/o equivalentemente delle alte temperature). Infatti, dopo la frase sopra riportata, egli immediatamente aggiunge "*And although for some reason not yet explained the doctrine fails in general, it seems possible that it may apply to the graver modes*", e pertanto, avendo eseguito il conteggio dei modi normali tra ν e $\nu + \mathrm{d}\nu$ (egli in effetti si riferisce alla decomposizione in lunghezze d'onda) propone le legge che, scritta in termini di frequenze, ha la forma

$$u_\nu(T) = c_1 \nu^2 \ k_B T \ e^{-c_2 \nu / T} \ . \tag{2.4.14}$$

Si noti che questa legge rappresenta i dati sperimentali (per il comportamento a basse frequenze e/o alte temperature) molto meglio di quanto non faccia la pura legge di Wien, la quale svolse un ruolo centrale nello sviluppo che condusse alla legge di planck.

Per quanto riguarda il conteggio del numero dei modi normali, prendiamoci il piacere di leggere le sue stesse parole. In effetti, Rayleigh determina tutto a meno di costanti moltiplicative, prendendo uguale ad uno la velocità di propagazione: daremo subito sotto la forma completa dell'espressione. Egli dunque così si esprime.

"*Let us consider in illustration the case of a stretched string vibrating transversally. According to the Boltzmann–Maxwell law the energy should be equally divided among all the modes, whose frequencies are as* $1, 2, 3, \ldots$. *Hence, if* k *be the reciprocal of* λ, *representing the frequency, the energy betweem the limits* k *and* $k + \mathrm{d}k$ *is (when* k *is large enough) represented by* $\mathrm{d}k$ *simply.*

When we pass from one dimension to three dimensions, and consider for example the vibrations of a cubical mass of air, we have (Theory of Sound, section 247) as the expression for k^2,

$$k^2 = p^2 + q^2 + r^2 \ ,$$

where p, q, r *are integers representing the number of subdivision in the three directions. If we regard* p, q, r *as the coordinates of points forming a cubic array,* k *is the distance of any point from the origin. Accordingly the number of points for which* k *lies between* k *and* $k + \mathrm{d}k$, *proportional to the volume of the corresponding spherical shell, may be represented by* $k^2 \mathrm{d}k$, *and this expresses the distribution of energy according to the Boltzmann–Maxwell law, as far as regards the wavelength or frequency.*"

Poi egli conclude dando, a meno di fattori moltiplicativi, la forma della cosiddetta legge di Rayleigh–Jeans (2.4.13) che seguirebbe dal principio di equipar-

tizione, e infine la legge (2.4.14) che egli propone, che corrisponde ad assumere l'equipartizione soltanto *for the graver modes* (per i modi più gravi, ovvero per i modi di bassa frequenza).

Per completezza ricordiamo che, come si controlla immediatamente, la formula per il conteggio dei modi tra ν e $\nu + \mathrm{d}\nu$ per unità di volume, per una campo scalare tridimensionale caratterizzato da una velocità c (ricavata con il metodo spiccio di Rayleigh)[66] risulta essere

$$dN_\nu = \frac{4\pi\nu^2}{c^3}\,\mathrm{d}\nu\ .$$

Nel caso del campo elettromagnetico, a causa del carattere trasversale del campo, che comporta due possibili polarizzazioni, si hanno poi due vibrazioni trasversali anziché una sola, e la formula diventa

$$dN_\nu = \frac{8\pi\nu^2}{c^3}\,\mathrm{d}\nu\ .$$

2.5 Boltzmann contro l'equipartizione

Si veda L. Boltzmann, Lectures on gas theiry, Dover 1964, pag, 330–3311, Nota 4, dove viene detto: Nei casi in cui si trova sperimentalmente un calore specifico inderiore a quello di Dulong e Petit (previsto dal teorema di equipartizione) "*one must assume that the motions related to certain parameters (gradi di libertà) come into equilibrium with the others* so slowly *that they do not contribute to the specific heats determined by experiment.*"

NOTA PER GLI AUTORI: Riportare qui i commenti di Boltzmann sui tempi lunghi per l'equipartizione. Citare Boltzmann dall'articolo con Barbara e da Nature. Citare anche l'articolo di Poincaré del 1906 in cui discute il gas dentro un cubo a pareti riflettenti. Jeans Sovay.

[66] Metodi più sofisticato furono trovati da Hermann Weyl e da David Hilbert.

Capitolo 3

Planck e il corpo nero: 19 ottobre e 14 dicembre 1900

Il corpo nero, che ha dato origine alla Meccanica Quantistica, occupava già un posto molto rilevante nella Fisica Teorica di fine '800. Era infatti il banco di prova per tre nuove discipline, elettromagnetismo, termodinamica, meccanica statistica, nate tutte nella seconda metà del secolo. Ma queste sembravano in qualche modo essere inconciliabili tra di loro. Ed infatti la soluzione proposta da Planck ai problemi legati al corpo nero, portò ad una vera rivoluzione nella concezione della Natura, tra le maggiori mai avvenute in ogni epoca.

Nei moderni manuali il problema del corpo nero viene spesso relegato al ruolo di un banale esercizio, brevemente discusso alla fine dì tutta una precedente trattazione. In particolare viene quasi a scomparire l'aspetto fondamentale dell'equilibrio termico tra radiazione e materia, e in tal modo viene quasi completamente sottovalutato, banalizzandolo, anche il ruolo del lavoro del 1907 di Einstein sui calori specifici.

Nel presente capitolo tentiamo di dare una succinta descrizione del cammino percorso da Planck per giungere alla sua sofferta quantizzazione. Cercheremo di mettere in luce come il problema dell'equlibrio radiazione–materia sia stato dapprima trasportato da Planck in un problema che sembrerebbe coinvolgere solo la meccanica, quella di un oscillatore materiale, pervenendo in maniera empirica (con un fit su dati sperimentali) alla sua legge. Vedremo come egli riprodusse poi tale formula in termini statistici, con un procedimento alla Boltzmann, e come tale procedimento appaia richiedere la quantizzazione dell'energia. In ogni caso, lo studio de problema del corpo nero ci insegna, come enfatizzato da Feynma nella sua Nobel Lecture. é che il problema della quantizzazione coinvolge in una meniera inestricabile radiazione e materia, e non può essere banalizzato ad apparire un problema di pura meccanica.

Più in generale, il problema del corpo nero costituisce un esenpio paradigmatico per comprendere cosa vuol dire "creare una teoria". Anzitutto, si ha a che fare con "bruti fatti sperimentali", che con la loro bruta forza si ergono davanti a

noi provocandoci a cercarne una comprensione. In una prima fase, si formulano ragionamenti che si concretizzano in formule, che appaiono descrivere abbastanza bene i fenomeni. Quando i ragionamenti diventano un po' più stringenti e apparentemente formalizzati, e inoltre investono un ambito abbastanza generale, allora tutti concordano sostanzialmente sul fatto che si abbia una teoria. Nel caso della teoria quantistica questo sviluppo si è determinato in una maniera incredibilmente affascinante, sia nella sua origine con la legge di Planck, sia nelle successive fasi di Bohr (1913) e di Heisengerg e Schrödinger del 1925–26. Infine si è passati addirittura ad affermare che la teoria così fissata sia l'ultima che si possa dare. Una affermazione questa. che, come sappiamo, non era molto gradita a Planck, ad Einstein e a Schrödinger.

Anzitutto illustreremo la fenomenologia del corpo nero, proseguendo poi con i lavori di tipo termodinamico di Planck che lo condussero alla deduzione della legge di Wien. Questa, in effetti, coglie un elemento essenzile della fenomenologia del corpo nero, stabilendo un decadimento esponenziale per l'energia sulle alte frequenze e/o alle basse temperature. Verremo poi alla crisi che si produsse quando nuove misure sperimentali rivelarono l'inadeguatezza della legge di Wien alle basse frequenze e/o alle alte temperature. Mostreremo infine come Planck fu condotto (17 ottobre 1900) a proporre una nuova legge fenomenologica (la legge di Planck, appunto) che egli riuscirà poi a giustificare (14 dicembre 1900) tramite ragionamenti "teorici" meccanico–statistici ispirati a quelli di Boltzmann (ma con una loro originalità), a patto di introdurre i quanti di energia.

3.1 Il corpo nero come sistema termodinamico

Richiami sul corpo nero. La legge di Kirchhoff

Una descrizione compatta e molto bella del problema del corpo nero è data da Fermi, nel suo libro *Fisica Atomica*, cap. IX, paragrafo 4 (Spettro del corpo nero), pag 276 e seguenti.

È noto che un corpo incandescente emette luce, e dunque radiazione elettromagnetica. Così avviene ad esempio per il Sole, la cui radiazione era stata studiata sperimentalmente in maniera spettroscopica (cioè come essa viene dispersa da un prisma), determinandone lo 'spettro", ovvero l'intensità in funzione della frequenza ν (o della corrispondente lunghezza d'onda λ). Come è noto, tale spettro si estende ben oltre lo spettro visibile, nell'infrarosso e nell'ultravioletto. Esperienza significative a questo proposito erano state condotte in Italia da Macedonio Melloni e da A.G. Bartoli.[1]

Ricordiamo che, come è ben noto, lo spettro di un corpo incandescente (intensità in funzione della frequenza, o della lunghezza d'onda) ha la forma grossomodo di una campana asimmetrica, con un decadimento maggiormente rapido dalla parte delle alte frequenze. Si veda la figura 3.1.

[1] A.G. Bartoli, *Sopra i movimenti prodotti dalla luce e dal calore*, Le Monnier (Firenze, 1876).

Ricordiamo anche che, come il più semplice modello di gas è il gas perfetto, così l'esempio ideale di corpo incandescente è il cosiddetto *corpo nero*, ovvero un corpo che assorbe tutta la radiazione che incide su di esso. Il motivo è legato ad una fondamentale legge stabilita da Kirchhoff, il quale aveva introdotto i classici procedimenti della termodinamica nello studio della radiazione elettromagnetica, in relazione al potere emissivo "e" e potere assorbente "a" di un corpo.

Se $e(\nu, T)\,\mathrm{d}\nu$ è la quantità di energia raggiante, appartenente all'intervallo di frequenza tra ν e $\nu+\mathrm{d}\nu$, che viene emessa, per unità di tempo, dall'unità di superficie del corpo,[2] alla temperatura T, allora e viene chiamato *potere emissivo* del corpo. Invece si dice *potere assorbente* a di un corpo il rapporto tra la quantità di energia assorbita e quella incidente sul corpo (Fermi, pag. 219–220).

Nota. Si osservi che l'uso del medesimo termine "potere" (emissivo o assorbente) per le due quantità puoò essere fuorviante, perché si tratta di due quantità fisiche di natura completamente diversa: infatti il potere emissivo è dimensionale, mentre il potere assorbente è un puro numero. La ragione di tale diversità sta nella diversità dei processi di emissione ed assorbimento: mentre il processo di emissione è di tipo "spontaneo", perché ogni corpo irraggia energia "per conto suo", naturalmente, invece il processo di assorbimento si produce solo quando il corpo viene investito da una radiazione. Dunque la sua capacità di assorbimento viene valutata misurando quale "frazione" esso assorbe, dell'energia che lo investe (per unità di superficie e di tempo).

In particolare, in base alle leggi della termodinamica, nel 1859 Kirchhoff aveva dedotto che il rapporto tra potere emissivo e potere assorbente relativi a una frequenza ν alla temperatura T è uguale per tutti i corpi, ed è dunque della forma

$$\frac{e}{a} = \mathscr{E}(\nu, T) \equiv \mathscr{E}_\nu(T)\,,$$

dove $\mathscr{E}$ è una *funzione universale della temperatura e della frequenza*.

Il cuore dell'argomento di Kirchhoff. Diamo qui un cenno dell'argomento di Kirchhoff, cercando di metterne in rilievo l'idea centrale, anche a costo della precisione. Consideriamo due corpi S_1 ed S_2 affacciati, in equilibrio alla medesima temperatura T. Tra le due facce vi sarà della radiazione, e sia R la radiazione che incide sulle superfici (per unitià di area e di tempo). Essendo in equilibrio, la quantità di energia emessa da ciascuno dei corpi è uguale a quella assorbita. In base alle definizioni appena date si avrà allora

$$e_1 = a_1 R\,, \quad e_2 = a_2 R\,,$$

e dunque

$$\frac{e_1}{a_1} = \frac{e_2}{a_2}\,.$$

Invece, se si avesse $e_1/a_1 \neq e_2/a_2$. allora uno dei due corpi si riscalderebbe a spese dell'altro, contro il secondo principio della termodinamica. Si ha poi un complemento,

[2] Il fatto che l'energia emessa sia proporzionale alla superficie del corpo, e non al suo volume, è una proprietà fondamentale, che non ha ancora ricevuto una "spiegazione" microscopica. È nostra opinione che la spiegazione possa essere trovata nell'ambito della "teoria di Wheeler e Feynman".

perché Kirchhoff dimostra anche che vale $R = (c/4)u$ dove u è l'energia della radiazione termica, per unità di volume. Infine, si osservi che tutto quanto detto sopra dovrebbe essere riferito a quantità (e, a, R, u) tutte relative ad una fissata frequenza ν.

Per definizione si dice **corpo nero** un corpo che assorbe completamente la radiazione incidente, ovvero che ha potere assorbente pari a uno, per cui cioè $a = 1$. Pertanto, *la funzione universale $\mathscr{E}_\nu(T)$ rappresenta il potere emissivo di un corpo nero*.

A proposito del corpo nero così si esprime Fermi (pag. 278):

"Osserviamo che la realizzazione più perfetta di una superficie nera è data da un piccolo foro che immetta in una grande cavità; infatti la luce che penetra nella cavità attraverso al foro deve subire numerose riflessioni sulle pareti interne della cavità prima di potere uscire dal foro di accesso e viene assorbita quasi completamente in queste riflessioni multiple. È questa la ragione per cui l'imboccatura di una grotta oppure la porta di una chiesa o di una grande sala appaiono nere a chi la vede dall'esterno."

Naturalmente, se all'interno della cavità è presente della radiazione, questa potrà essere emessa attraverso il foro, come un gas che fuoriesca da un recipiente bucato. *La radiazione elettromagnetica entro una cavità può dunque essere pensata come una qualche specie di sostanza avente certe sue proprietà termodinamiche specifiche che la caratterizzano*, allo stesso modo in cui certe altre proprietà specifiche caratterizzano macroscopicamente, dal punto di vista termodinamico, ad esempio la "sostanza" gas perfetto.

Funzioni di stato della termodinamica del corpo nero: l'energia interna e la pressione

Seguiamo qui il libro di Planck *Wärmelehrung*, la cui seconda edizione del 1912 (la prima è del 1906) è disponibile in traduzione inglese, e facilmente accessibile.[3] Una presentazione molto bella e compatta si trova anche nel bellissimo libro di Max Born, *Atomic Physics*, Dover (New York, 1969), Capitolo VII, pag. 204 e seguenti, e Appendici XXXIII–XXXIV.

La prima funzione di stato che si considera è l'energia interna, che qui denotiamo con $\mathscr{U}$, per la quale si ammette che valga

$$\mathscr{U}(T, V) = V u(T)\,, \tag{3.1.1}$$

dove la funzione $u(T)$ può essere chiamata *densità di energia (rispetto al volume)*, ovvero energia per unità di volume. Si noti bene che l'ipotesi (3.1.1) è tutt'altro che ovvia per una sostanza termodinamica. Infatti, ad esempio il gas perfetto è addirittura caratterizzato dal fatto che l'energia interna, pur essendo proporzionale alla massa del gas, è *indipendente* dal volume in cui il gas è contenuto.

Questa forma dell'energia interna è dimostrata da Planck nella sezione 22 del suo libro, e sarebbe dovuta unicamente al fatto che la luce si propaga con velocità finita. Comunque, la si potrebbe ritenere ragionevole come proprietà valida per ogni campo (e quindi

[3] Esistono però anche una terza edizione ed una quarta, reperibii solo in tedesco.

in particolare per il campo elettromagnetico), perché per i campi l'energia totale risulta essere l'integrale di una densità di energia locale (funzione del posto). Se si ha omogeneità, ovvero la densità non dipende dal posto (come dovrebbe avvenire all'equilibrio termico), si ha allora immediatamente la (3.1.1).

La seconda relazione che definisce la termodinamica del corpo nero è quella che prende il posto dell'equazione di stato del gas perfetto, $pV = \mathcal{N}RT$ (dove $\mathcal{N}$ è il numero di moli). Nel caso del corpo nero tale relazione viene rimpiazzata da

$$p = \frac{u(T)}{3} \ . \tag{3.1.2}$$

È questa infatti la relazione che era stata sostanzialmente stabilita teoricamente da Maxwell sulla base delle sue equazioni per il campo elettromagnetico.[4]

L'analogia con la legge del gas perfetto diventa ancora più evidente se si moltiplica la relazione (3.1.2) per il volume V, perché allora essa prende la forma

$$pV = \frac{1}{3}\mathcal{U} \ , \tag{3.1.3}$$

dove $\mathcal{U}$ è l'energia elettromagnetica (ma si veda la precisazione data subito sotto) contenuta nel volume V. Infatti, nella deduzione microscopica data da Clausius la legge del gas perfetto ha la forma

$$pV = \frac{2}{3}\overline{K^{cm}}$$

dove la barra denota media temporale, e K^{cm} denota l'energia cinetica totale dei baricentri delle molecole costituenti il gas. Analogamente, a destra della (3.1.3) compare una quantità microscopica, che è proprio l'energia del campo elettromagnetico.[5]

Vedremo tuttavia più avanti che la analogia del corpo nero con il gas perfetto vale, almeno in un certo senso, solo nel limite di alte frequenze e/o basse temperature, nel quale il corpo nero è descritto dalla legge di Wien anziché dalla legge di Planck. Questo fu dimostrato nel celebre articolo del 1905 di Einstein, dal quale si fu per la prima volta mostrato che un campo (qui, quello elettromagnetico) presenta proprietà particellari: si tratta dei cosiddetti fotoni.

Ricordiamo brevemente che il campo elettromagnetico "microscopico" (secondo la concezione di Lorentz e Planck) è descritto da campi vettoriali $\mathbf{E}(\mathbf{x}, t)$ (campo elettrico) e $\mathbf{B}(\mathbf{x}, t)$ (campo magnetico) che ubbidiscono alla leggi di Maxwell (scritte nelle unità di

[4]NOTA PER GLI AUTORI. Vedere come mai nel libro di Drude (pag. 490 dell'edizione inglese) si mostra – e in maniera molto semplice e bella – $p = u(T)$ senza il fattore 1/3. Forse si ha qui il problema citato da Boltzmann nel suo lavoro?

[5]Più precisamente, si tratta della parte radiativa, come discuteremo più avanti. Qui vi è una analogia con la legge del gas perfetto, perché in quel caso quella che entra non è tutta l'energia cinetica, ma solo quella dei baricentri delle molecole. A dire il vero, qui vi sarebbe un punto delicato, su cui ora sorvoliamo.

Gauss) seguenti

$$\operatorname{div}\mathbf{E} = 4\pi\rho(\mathbf{x},t)\,, \qquad \mathbf{rot}\,\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}$$

$$\operatorname{div}\mathbf{B} = 0\,, \qquad \mathbf{rot}\,\mathbf{B} = \frac{4\pi}{c}\mathbf{j}(\mathbf{x},t) + \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}\,,$$

dove si è indicata con c la velocità della luce nel vuoto, e con ρ e $\mathbf{j}$ rispettivamente la densità di carica e di corrente. La densità di energia "microscopica" è allora data da

$$\mathscr{E}(\mathbf{x},t) = \frac{E^2(\mathbf{x},t) + B^2(\mathbf{x},t)}{4\pi}$$

mentre la forza a livello "microscopico" può ottenersi a partire dal tensore degli sforzi

$$T_{i,k}(\mathbf{x},t) = \frac{1}{4\pi}\big(E_i(\mathbf{x},t)E_k(\mathbf{x},t) + B_k(\mathbf{x},t)B_i(\mathbf{x},t)\big) - \frac{1}{2}\,\mathscr{E}(\mathbf{x},t)\delta_{i,k}\,,$$

Viceversa le quantità macroscopiche, come l'intensità I di irraggiamento, la pressione p, o la stessa densità di energia $u(T)$, devono essere intese come opportune medie su tempi "lunghi" rispetto ai tempi della dinamica microscopica, e su volumi "grandi" rispetto alle dimensioni atomiche.[6]

Il fattore 1/3 che entra nell'equazione di stato (3.1.2) segue proprio facendo queste medie dei valori microscopici, come spiegato da Boltzmann nel 1884, e forse prima da Adolfo Bartoli, cui fa riferimento Boltzmann nel suo lavoro.

Relazione tra potere emissivo E_ν e densità spettrale di energia u_ν

A fissata temperatura, si considera non soltanto la densità di energia u ma anche la sua "componente" spettrale di frequenza ν, che denoteremo con u_ν. Riservandoci di definirla meglio più avanti (essendo il problema della sua definizione assai delicato), ricordiamo solo che intuitivamente $u_\nu \mathrm{d}\nu$ è l'energia (per unità di volume) contenuta nell'intervallo $(\nu, \nu + \mathrm{d}\nu)$ di frequenza. Si ha dunque

$$u(T) = \int_0^\infty u_\nu(T)\,\mathrm{d}\nu\,.$$

Tenendo conto che i campi si propagano alla velocità della luce c, Kirchhoff stabilì che tra il potere emissivo $\mathscr{E}_\nu(T)$ (la quantità effettivamente osservata) del corpo nero e la "componente" u_ν vale la semplice relazione di proporzionalità[7,8]

$$\mathscr{E}_\nu(T) = \frac{c}{4}u_\nu(T)\,.$$

[6]NOTA PER GLI AUTORI. Inserire lo spettro alla Wiener, che si trova nel file.tex con il percento.

[7]Si veda Planck, *Theory of heat radiation*, sect. 23, formula (24) dove c è denotata con q e $\mathscr{E}_\nu$ co K_ν, oppure formula (21) con un fattore 4π (oppure 8π ?)

[8]NOTA PER GLI AUTORI. Questa formula deve essere qualitativamente sbagliata, perché non distingue tra energia e energia scambiata, in analogia con quanto avviene nel problema del calore specifico trattato alla Kubo. Analoga critica si può fare alla legge di Planck per la relazione tra energia U_ν del risonatore materiale ed energia u_ν del campo.

La componente u_ν è la quantità concretamente studiata da Boltzmann, Wien, Planck, in relazione allo spettro del corpo nero.

Nota: curve isoterme e curve isocromatiche. L'intensità spettrale (o l'energia per unità di volume) è dunque una funzione delle due variabili ν, T. Se si considera T come parametro e si osserva "lo spettro", ovvero l'energia come funzione di ν, si dice che si sta considerando una *isoterma*, una curva che potremmo denotare con $u_T = u_T(\nu)$. Se invece si considera ν come un parametro, e si pensa l'energia come funzione della temperatura T, allora si dice che si sta considerando una *isocromatica* (cioè a curva a colore – ovvero frequenza – costante), che potremmo denotare con $u_\nu = u_\nu(T)$.

I dati sperimentali per le isoterme venivano più frequentemente riportati in funzione della lunghezza d'onda λ invece che della corrispondente frequenza ν (dove $\lambda\nu = c$), e allora si ha evidentemente una analoga decomposizione

$$u(T) = \int_0^\infty u_\lambda(T)\,\mathrm{d}\lambda\,.$$

A questo proposito si deve porre attenzione al fatto che, per ogni assegnata temperatura T, non si ha la relazione $u_\lambda = u_\nu$, bensí la relazione

$$u_\lambda = c\,\lambda^{-2} u_\nu \quad (\text{con} \quad \nu = c/\lambda)\,.$$

Questo si capisce immediatamente se si tiene presente che l'energia contenuta nell'intervallo $(\lambda, \lambda + \mathrm{d}\lambda)$ deve essere uguale a quella contenuta nel corrispondente intervallo $(\nu, \nu + \mathrm{d}\nu)$, ovvero si deve avere[9]

$$u_\lambda(T)\mathrm{d}\lambda = u_\nu(T)\mathrm{d}\nu\,,$$

e si usa

$$|\mathrm{d}\nu| = \frac{c}{\lambda^2}\,|\mathrm{d}\lambda|\,.$$

Un esempio di isoterme è riportato in Fig. 3.1, per temperature tra 700 e 1600 °C. Per una discussione relativa a questa figura, si veda il bel libro H. Kangro, *Early history of Planck's radiation law*, Taylor & Francis (London, 1976). Un esempio di isocromatiche è riportato più avanti, in Fig. 3.2. Quest'ultima figura ha una importanza fondamentale, perché è quella che convinse Planck che la legge di Wien vale solo come legge limite per grandi valori di ν/T, sicché egli venne forzato a trovare una diversa forma analitica per la legge del corpo nero.

La legge di Stefan–Boltzmann

Il primo grande risultato teorico dopo il teorema di Kirchhoff del 1859, fu ottenuto nel 1884 da Boltzmann,[10] che dimostrò la formula che era stata precedentemente stabilita sperimentalmente da Stefan. Secondo tale "*legge di Stefan–Boltzmann*"

[9] Formalmente, si scrive la condizione

$$\int_0^\infty u_\lambda(T)\mathrm{d}\lambda = \int_0^\infty u_\nu(T)\mathrm{d}\nu$$

e si esegue il cambiamento di variabile $\lambda = c/\nu$.

[10] L. Boltzmann, Wissenschaftliche Abhandunglen, III, Chelsea R.C. (New York, 1968), N. 72, Vedi anhe N. 71: *Ueber eine von Hrn. Bartoli entdeckte Beziehung der Waermestrahlung zum zweiten Hauptsatze*.

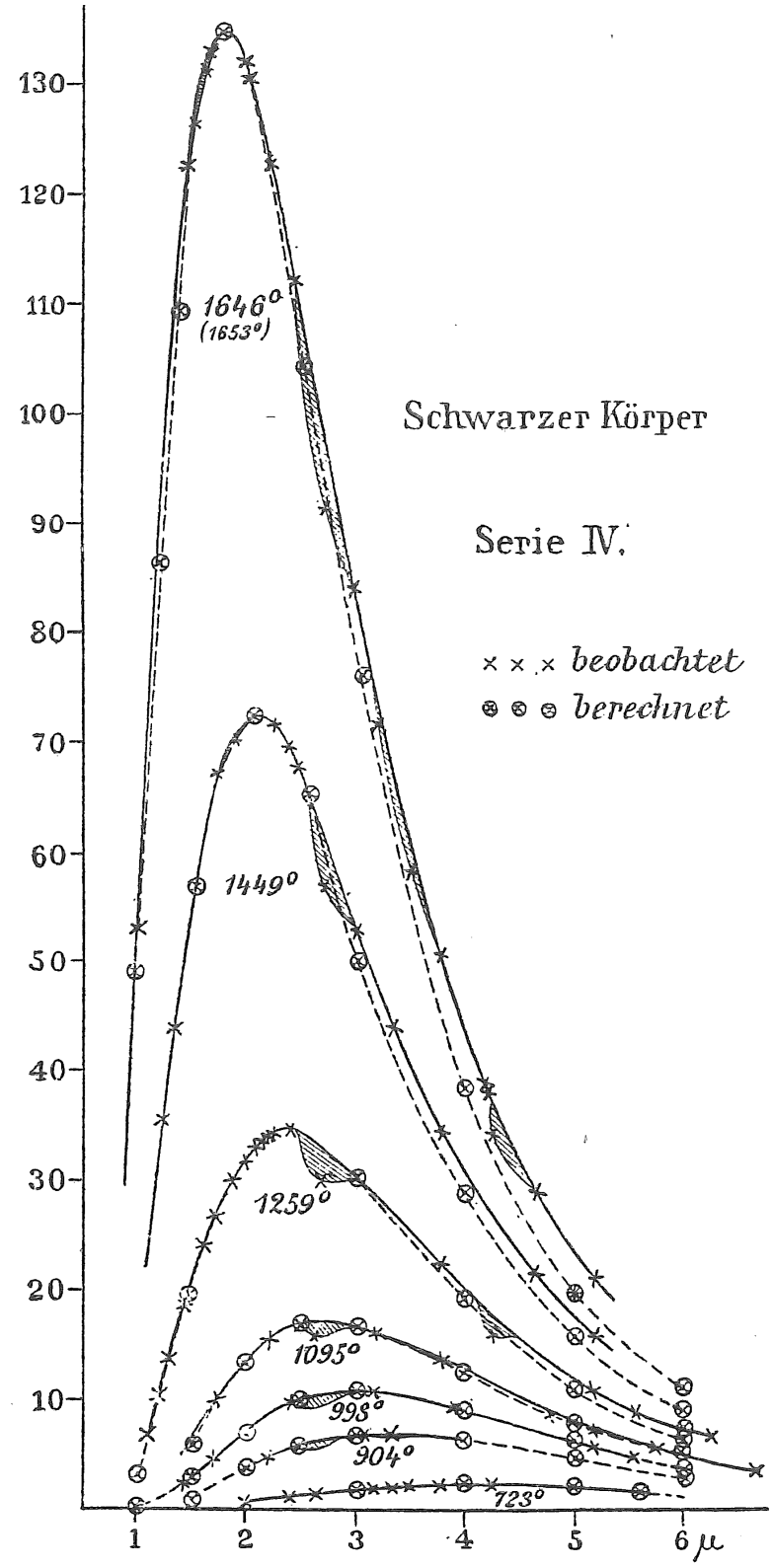

Figura 3.1: Spettro del corpo nero. Grafico di sette isoterme (intensità vs lunghezza d'onda) tra 700 e 1600 °C per lunghezze d'onda tra 1 e 6 μ. Risultati sperimentali di Lummer e Pringsheim nel Novembre 1899, c onfronto con la legge di Wien.

la potenza (energia per unità di tempo) emessa da un "corpo nero" alla temperatura T cresce come la quarta potenza di T. Equivalentemente, in virtù della relazione tra potere emissivo e densità di energia u, si ha

$$u(T) = \sigma T^4 \,, \tag{3.1.4}$$

dove σ è una certa costante, detta *costante di Stefan–Boltzmann*.

La dimostrazione è una immediata conseguenza dei primi due principi della termodinamica, applicata al "fluido elettromagnetico", il quale è caratterizzato dalle due funzioni di stato sopra illustrate, l'energia interna $\mathscr{U}$ (3.1.1), e la pressione p (3.1.2).

Il procedimento originale di Boltzmann, esposto in un articolo che fa seguito a una sua nota sui lavori di Bartoli, non è banale da seguirsi. Seguiamo qui la incredibilmente semplice esposizione di Planck[11]

Basta ricordare la relazione $\mathrm{d}U = \delta Q - p\mathrm{d}V$ che esprime il primo principio, e quella che connette calore ed entropia alla luce del secondo principio, $\mathrm{d}S = \delta Q/T$, dalle quali segue immediatamente

$$\mathrm{d}S = \frac{\mathrm{d}U + p\mathrm{d}V}{T} \,.$$

Sostituendo le relazioni $U = Vu$ e $p = u/3$, per la funzione $S(T, V)$ si ottiene immediatamente (usando $\mathrm{d}U = V\mathrm{d}u + u\mathrm{d}V$)

$$\mathrm{d}S = \frac{V}{T}\frac{\mathrm{d}u}{\mathrm{d}T}\mathrm{d}T + \frac{4}{3}\frac{u}{T}\mathrm{d}V \,.$$

Ma S è una funzione di stato, e dunque $\mathrm{d}S$ un differenziale esatto. Pertanto le due derivate seconde miste devono essere uguali, e si deve avere (come immediatamente si verifica)

$$\frac{\mathrm{d}u}{\mathrm{d}T} = \frac{4u}{T} \,.$$

Questa è una equazione differenziale del primo ordine nella funzione incognita $u(T)$, che ha evidentemente come soluzione proprio la (3.1.4) con una costante arbitraria σ, come si vede subito o come può comunque essere immediatamente verificato.

La legge di spostamento di Wien

Un grande passo venne poi compiuto da Wien nel 1896.[12] Si osservava, come si vede dalla figura (3.1), che al variare della temperatura le isoterme del corpo nero avevano tutte una forma alquanto simile, l'aspetto caratteristico essendo che al crescere della temperatura le curve u_ν vs ν (delle specie di campane con una pendenza decisamente più grande dalla parte delle alte frequenze) si alzavano (in accordo col fatto che l'area sottesa deve crescere come T^4 secondo la legge di Stefan–Boltzmann), ed inoltre il picco si spostava verso destra,[13] precisamente

[11] Wärmelehrung, sezioni 62–63.

[12] W. Wien, *Wied. Ann* **58**, 662 (1896).

[13] Un indicazione di questo tipo si ottiene facilmente scaldando uno spillo su un fornello a gas: inizialmente lo spillo è opaco, poi diventa incandescente il colore passa da un rosso cupo (diciamo mattone, circa 600 gradi), poi ciliegia (circa 700 gradi centigradi) poi rosso ciliegia chiaro (circa

secondo la legge $\lambda_{max} T = \text{cost}$ o equivalentemente

$$\nu_{max} = \text{cost } T \ .$$

Le curve parevano dunque potersi ricondurre tutte ad una unica curva mediante un opportuno riscalamento.

Questo riscalamento fu determinato da Wien con un argomento teorico che da una parte faceva uso della seconda legge della termodinamica (come già avveniva nei procedimenti di Kirchhoff, di Bartoli e di Boltzmann), e d'altra parte si rifaceva ad un argomento che oggi chiameremmo di teoria dei sistemi dinamici. Questo riguardava i cosiddetti *invarianti adiabatici*, particolarmente studiati da Helmholtz (1821–1894) e Boltzmann.

Gli invarianti adiabatici sono sostanzialmente variabili dinamiche che non soltanto sono praticamente costanti del moto fino a tempi sufficientemente lunghi, ma hanno inoltre la proprietà di mantenere il loro valore inalterato anche quando si varia lentamente un parametro. L'esempio più classico è quello delle piccole oscillazioni del pendolo di cui venga fatta variare lentamente la frequenza (si pensi ad esempio di allungare lentamente la corda a cui il pendolo è sospeso). In tal caso si mostra che, al variare della frequenza ν col tempo, varia anche l'energia E. Si dimostra però che fino a tempi lunghi resta sostanzialmente inalterato il valore del rapporto $I \equiv E/\nu$. L'invariante adiabatico I ha evidentemente le dimensioni di una azione, energia per tempo, come la costante di Planck ed il momento angolare.[14]

Nel caso del corpo nero, questi invarianti adiabatici riguardano l'energia contenuta nell'intervallo tra ν e $\nu + \mathrm{d}\nu$ quando si sposta lentamente il bordo della cavità per ottenere un lavoro $p\mathrm{d}V$. Infatti, in analogia con quanto avviene per la corda vibrante (la cui equazione di moto fu stabilita fin dal 1750), tutti sappiamo che, al variare della lunghezza della cavità indotto dallo spostamento di una parete, variano anche le frequenze caratteristiche della cavità (diminuendo al crescere del volume). Si ha dunque il problema di connettere l'energia contenuta nell'intervallo tra ν e $\nu + \mathrm{d}\nu$ con quella contenuta nel corrispondente intervallo tra ν' e $\nu' + \mathrm{d}\nu'$[15].

In ogni caso questi argomenti di Wien, esposti nel Capitolo III del libro di Planck, e nell'appendice XXXIII del libro *Atomic Physics* di Born, conducono

800 gradi). Con un fornello non si riesce ad ottenere una temperatura superiore, perchè l'energia che irraggia lo spillo è pari a quella fornita dalla fiamma. Se si ponesse lo spillo in un forno opportunamente costruito, ad esempio per la fusione dell'acciaio, si vedrebbe lo spillo diventare arancione (circa 1000 gradi), poi giallo circa (1200) gradi, poi bianco cominciando a fondere (circa 1400 gradi per l'acciaio).

[14]Questo punto fu particolarmente sottolineato da Paul Ehrenfest. Questi mise in luce come in generale la quantizzazione si dovesse ottenere per hamiltoniane che dipendono da variabili di azione, imponendo che ogni azione fosse un multiplo intero di $\hbar$. Nel lavoro del 1914 riportato nel libro di van der Waerden, la frase conclusiva di Ehrenfest è la seguente: *"In any case, it seems to me that the validity of Wien's displacement law shows that reversible adiabatic changes take a prominent place in the theory of quanta."*

[15]Secondo i presenti autori, la trattazione data da Wien lascia aperti problemi non banalissimi, che potrebbero forse essere chiariti in futuro. Di questa opinione è anche Giorgio Parisi, dell'Università di Roma La Sapienza.

alla *legge di spostamento di Wien*

$$u_v(T) = \nu^3 f\left(\frac{\nu}{T}\right) \equiv \nu^2 T g\left(\frac{\nu}{T}\right) , \tag{3.1.5}$$

dove f (oppure g) è ora l'unica significativa funzione universale caratteristica del corpo nero che resta da determinare, teoricamente e/o empiricamente.

Nota. Il più semplice modo per ottenere la legge di spostamento di Wien è quello illustrato da Einstein alla Conferenza Solvay del 1911 (pare, seguendo Jeans). Il risultato (che però è solo parzialmente una dimostrazione) segue immediatamente da una analisi dimensionale, allo stesso modo in cui si ottiene ad esempio la legge per il periodo del pendolo. Notiamo inoltre che, essendo il rapporto ν/T dimensionale, deve esistere una costante h con le dimensioni di una azione per cui si abbia $f = f(h\nu/k_B T)$. D'altra parte, poiché si pensa di fare una teoria microscopica dell'irraggiamento, nella teoria dovrà necessariamente comparire il valore e della carica microscopica ed il valore c della velocità della luce. In effetti, a partire da queste due costanti si ottiene una azione, data da e^2/c, il cui valore non è molto discosto da h, in quanto si ha $\hbar \simeq 137\, e^2/c$. Il numero adimensionale $\alpha \stackrel{\text{def}}{=} e^2/\hbar c \simeq 1/137$ viene chiamato *costante di struttura fina*, e la sua rilevanza venne messa in luce da Sommerfeld nello studio della struttura fina delle righe dell'atomo di idrogeno.

Leggi fenomenologiche per il corpo nero: la legge di Wien e le varianti di Thiesen e di Rayleigh

Era ben noto che, alle alte frequenze, le isoterme decadono a zero esponenzialmente al crescere della frequenza. Dunque, in virtù della legge di spostamento di Wien secondo cui la dipendenza dalla temperatura si presenta nella forma $f(\nu/T)$, anche le isocromatiche devono decadere esponenzialmente come $1/T$ al decrescere della temperatura. In conclusione, per alte frequenze e/o basse temperature, la legge di radiazione deve avere la forma[16]

$$u_\nu(T) \simeq c_1 \nu^3\, e^{-c_2 \nu/T} . \tag{3.1.6}$$

È questa la legge che chiameremo anche *legge di Wien di decadimento esponenziale* (come distinta dalla legge di spostamento di Wien). In effetti Wien aveva dato anche degli argomenti di tipo euristico che egli considerava come una deduzione di tale legge, ritenendo addirittura che tale legge avrebbe dovuto valere in tutto il dominio delle variabili ν, T, e non solo per alte frequenze (e/o basse temperature). Questo fatto veniva confermato da F. Paschen,[17] ma poi smentito dalle osservazioni di Lummer e Pringsheim[18] relative a lunghezza d'onda più lunghe e temperature più alte che nelle osservazioni precedenti. Infatti, secondo la legge di decadimento di Wien, su ogni isocromatica (ν fissato), al crescere della temperatura l'energia deve tendere a una costante dipendente da ν mentre i nuovi dati,

[16] W. Wien, *Wied. Ann.* **58**, 662 (1898).
[17] F. Paschen. *Sitz. Akad. Wiss. Berlin*, pag. 959 (1899).
[18] O.Lummer. E.Pringsheim, *Verh. Deuts. Phys. Ges.* **1**, 215 (1899).

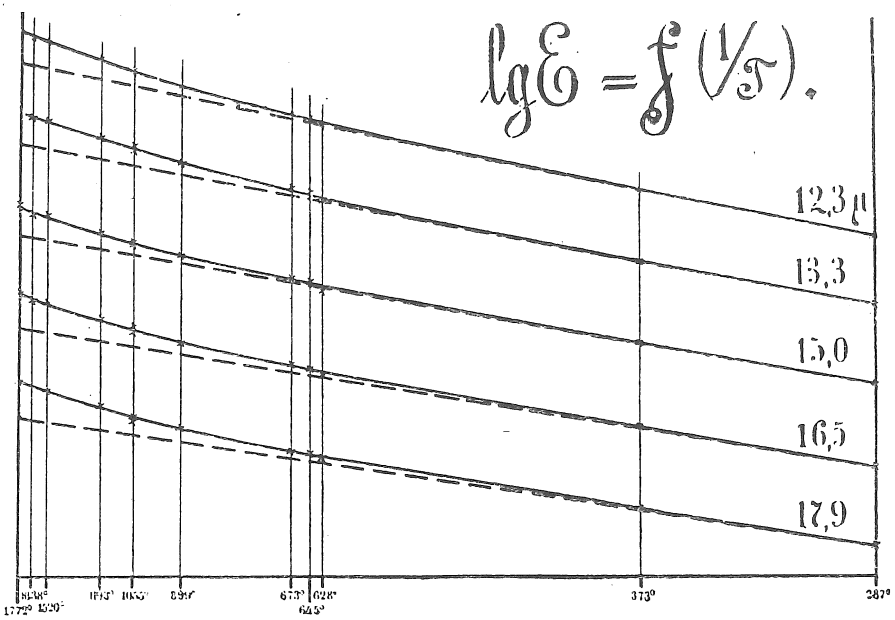

Figura 3.2: Spettro del corpo nero. Le isocromatiche di Lummer e Pringsheim del 1900 per cinque frequenze nell'infrarosso. Grafico del logaritmo dell'intensità in funzione di $1/T$. Le linee spezzate sono i fit con la legge di Wien (3.1.6). L'evidente scostamento sistematico a bassi valori di $1/T$ convinse definitivamente Planck ad accettare che la legge di Wien valesse solo asintoticamente per alti valori di $1/T$ (in effetti, di ν/T)

riportati alla figura (3.2), mostravano una deviazione sistematica alle alte temperature (a sinistra, nel grafico, che ha in ascissa $1/T$) e alle basse frequenze (perché le isosterme si riferiscono a righe infrarosse, caratteristiche dei cristalli ionici[19]).
.

Intermezzo. Per orientarsi, può essere utile anticipare le cose, dicendo che la legge corretta risulterà quella di Planck

$$u(\nu,T) = C\nu^2 \frac{h\nu}{e^{h\nu/k_BT}-1} = C\nu^2 k_B T \, \frac{h\nu/k_BT}{e^{h\nu/k_BT}-1}$$

da cui, con uno sviluppo di Taylor al prino ordine in $h\nu/k_BT$, si ottiene

$$u(\nu T) \simeq C\nu^2 k_B T \, .$$

Dunque, per ν costante (cioè su una isocromatica) per alte temperature si ha divergenza. Invece, secondo la legge di Wien la curva dovrebbe tendere a una costante dipendente

[19]Si tratta dei celebri "raggi restanti" o *residual rays*, in tedesco *Reststrhalen*.

dalla frequenza. Dai risultati di Lummer e Pringsheim si vede bene la crescita lungo curve esponenziali, fino a temperature non troppo alte.[20]

Vennero allora proposte delle altre leggi semi fenomenologiche. Dapprima da Thiesen, e poi da Rayleigh (maggio 1900). In particolare la legge proposta di Rayleigh ha la forma

$$u_\nu(T) \simeq c_1 \nu^2 \, k_B T \, e^{-c_2 \nu / T} \ ,$$

e fu proposta sulla base del fatto che egli accettava fenomenologicamente il decadimento esponenziale di Wien per alte frequenze e/o basse temperature. Invece, sulla base degli argomenti statistici di Maxwell e Boltzmann illustrati nel precedente capitolo, per alte temperature e/o per basse frequenze ("the graver modes", i modi più gravi, nelle sue parole) ammetteva che si dovesse avere l'equipartizione. ($k_B T$ per ogni modo).

Argomentazioni di questo tipo, che si rifacevano ai metodi della meccanica statistica, vennero invece sempre rifiutati da Planck, fino alla sua memoria del 14 dicembre, al punto che egli neppure mai citare il lavoro di Rayleigh pubblicato fn dal mese di maggio. Infatti Planck, che aveva dapprima sostanzialmente postulato o almeno ammesso la legge di Wien (lo dice egli stesso ripetutamente), produce poi un argomento di tipo termodinamico che sembra costituirne una dimostrazione. La storia poi cambierà quando il 7 ottobre, Rubens gli confermerà in maniera evidente le indicazioni di Lummer e Pringsheim sulla nonvalidità della legge di Wien ad alte temperature. Allora Planck proporrà la sua legge il 19 ottobre come pura legge fenomenologica, giungendo infine il 14 dicembre a una "giustificazione teorica" basata su argomenti di tipo statistico "alla Boltzmann".

3.2 I lavori di Planck

A differenza di Rayleigh, che applicava i metodi della meccanica statistica al campo elettromagnetico in maniera analoga a quanto si può fare per ogni campo, tipicamente la corda vibrante, inizialmente Planck ha come idea centrale quella di cercare di ottenere la legge del corpo nero con metodi termodinamici.[21] A tal fine egli considera un sistema composto di radiazione elettromagnetica da una

[20] Vale anche la pena di osservare che da uno sviluppo al secondo ordine nella legge di Planck si trova

$$u(\nu T) \simeq C\nu^2 k_B T \left(1 - \frac{1}{2}\frac{h\nu}{k_B T}\right) = C\nu^2\left(k_B T - \frac{1}{2}h\nu\right) ,$$

la quale mostra che per alte temperature si ottiene la legge di equipartizione $k_B T$ solo se si aggiunge la celebre energia di punto zero

$$E_{zp} = \frac{1}{2}h\nu$$

che fu introdotta da Planck nel 1911 e ritrovata poi da Heisenberg nel luglio 1925 e confermata da Schrödinger nel gennaio 1926.

[21] Sembra che egli manifestasse una opposizione di principio alle teorie cinetiche o statistiche, e si parla talvolta di una sua prima, e poi di una seconda, conversione.

parte, e dall'altra di un numero arbitrario di oggetti macroscopici identici, che possono oscillare a una ben definita frequenza ν, e che egli chiama "risonatori materiali". Egli pensa agli oggetti macroscopici che si usano (analogamente a quanto avviene in acustica) per misurare l'intensità della radiazione alla frequenza ν, considerata come un parametro, al variare del quale si ottengono informazioni sulle proprietà del campo. Partendo da una situazione di equilibrio tra i due sottosistemi (campo e risonatori) egli considera una nuova situazione in cui si ha una fluttuazione fuori dall'equilibrio, in cui l'energia del sistema di risonatori subisce una scostamento, diciamo ΔU (e naturalmente l'energia del campo subisce uno scostamento opposto). Il secondo principio della termodinamica richiede che il sistema torni in equilibrio con una certa crescita dell'entropia, e Planck si propone di trarre informazioni quantitative sullo spettro a partire da questa circostanza.

Questa idea venne sviluppata in una serie di lavori, di cui quelli rilevanti per la nostra discussione sono due, che vennero pubblicati all'inizio del 1900 sugli Annalen der Physik. Il primo, dal titolo *Sui processi irreversibili di radiazione*,[22] diviso in tre parti, è il riassunto di cinque precedenti contributi, comunicati all'Accademia Prussiana delle Scienze tra il 1897 e il 1899, con il medesimo titolo. Il secondo, dal titolo *Entropia e temperatura della radiazione termica*[23], è il più importante. In questi lavori Planck pervenne a dare una dimostrazione della legge di Wien, che, come detto, sembrava descrivere bene i dati sperimentali allora disponibili. Ma qualche mese dopo le osservazioni estese all'infrarosso mostrarono senza dubbi che la legge di Wien era qualitativamente errata alle basse frequenze e/o alle grandi temperature, e allora, suo malgrado, Planck si vide costretto a ripensare da capo al problema, ed ottenne la sua legge, prima (19 ottobre) per interpolazione dei dati sperimentali, e poi (14 dicembre) con argomenti di tipo statistico, ispirati a quelli di Boltzmann. I due primi lavori del 1900 si trovano, tradotti in italiano, nel libro di Paolo Campogalliani,[24] [25] dove sono riprodotti anche i lavori del 19 ottobre e del 14 dicembre, in cui Planck prima introdusse, e poi "dimostrò", la sua legge.

3.3 Deduzione di Planck della legge di Wien

Relazione tra spettro del corpo nero ed energia di un risonatore materiale

Anzitutto Planck stabilisce una relazione che svolge un ruolo fondamentale nella teoria del corpo nero, usata da lui stesso e poi da Einstein in maniera essenziale.

[22] M. Planck, *Ann. d. Phys.* **1**, 69–122 (1900).

[23] M. Planck, *Ann. d. Phys.* **1**, 69 (1900).

[24] P. Campogalliani, *Max Planck e la teoria della radiazione termica*, F. Angeli (Milano, 1999)

[25] Ringraziamo di cuore Paolo Campogalliani per averci messo a disposizione il suo libro, con la traduzione dei lavori rilevanti di Planck sul corpo nero. Infatti, stranamente, le Opere di Planck non paiono essere disponibili in rete.

Supponiamo di avere un risonatore materiale che possa oscillare con frequenza ν, di dimensioni (su scala macroscopica) piccolissime, che funga da rivelatore dell'intensità della radiazione in un punto specificato, e denotiamo con U la sua energia interna. Poniamolo in un certo punto di una cavità in cui vi sia della radiazione a una data temperatura. Allora esso scambierà energia con la radiazione fino a giungere ad uno stato di equilibrio in cui esisterà una determinata relazione tra l'energia $U(T)$ del risonatore e la densità di energia $u_\nu(T)$ della radiazione. Planck stabilisce che la relazione ha la forma [26]

$$u_\nu(T) = \frac{8\pi\nu^2}{c^3}\, U(T)\ , \qquad \text{ovvero} \qquad U(T) = \frac{c^3}{8\pi\nu^2}\, u_\nu(T)\,. \tag{3.3.1}$$

Nel dedurre questa relazione Planck opera una separazione tra quantità lentamente variabili (quelle macroscopiche) e quantità rapidamente variabili (quelle microscopiche). Questo è un punto molto importante. Una deduzione forse più limpida è quella data nella prima edizione (che esiste solo nella lingua originale) del Wärmestrahlung del 1906, primo capitolo. Questa è la deduzione cui fa riferimento Einstein nel suo noto lavoro con Stern del 1913.

Dunque, facendo riferimento alla formula (3.3.1) è possibile determinare lo spettro $u_\nu(T)$ se è conosciuta la energia media $U(T)$ di un risonatore materiale (formula a sinistra della (3.3.1)), o viceversa.

Come vedremo più avanti, il primo atteggiamento (indurre lo spettro del corpo nero dalla energia media del risonatore) fu quello tenuto da Planck nei suoi lavori. Invece Einstein, nel suo lavoro del 1907 sul calore specifico dei solidi, prende il secondo atteggiamento, determinare U da u_ν. Assumendo come sperimentalmente nota la legge del corpo nero, Einstein inferisce l'energia media dei risonatori materiali – tutti della stessa frequenza – mediante i quali egli modellizza il solido (e quindi l'energia del solido) in funzione della temperatura e della frequenza, e dunque ricava anche, per semplice derivazione rispetto a T, il corrispondente calore specifico.

Abbiamo già ricordato il procedimento con cui Rayleigh (nel suo articolo del giugno del 1900) formulò la proporzionalità (3.3.1) tra u_ν ed U. La differenza tra i due procedimenti è notevole: quello di Planck, che prende pagine di conti, riguarda in realtà il problema di come si possa "misurare" la quantità u_ν (la densitá di energia del campo). Invece, il calcolo di Rayleigh, che prende poche righe, è di carattere puramente meccanico–statistico, relativo al solo campo elettromagnetico libero, e si riduce al conteggio del numero di modi normali di oscillazione aventi frequenza nell'intervallo $(\nu, \nu + \mathrm{d}\nu)$. [27]

[26] Si tratta della seconda formula sotto la (41) del primo lavoro, riprodotta in alto a pag. 89 della traduzione di Campogalliani. In tale formula, al posto di u_ν compare J_ν (chiamato J_0) uguale a $(8\pi/3)u_\nu$.

[27] COMMENTO: DA RIVEDERE. Le dimostrazioni della relazione fondamentale (3.3.1) date da Planck fanno uso sostanziale della formula di Larmor per l'emissione di radiazione da parte di cariche accelerate, che viene impiegata per le antenne, e qui sembra venire estesa alle cariche microscopiche. Su questo Planck non è chiarissimo. È quindi lecito sollevare dei dubbi sulla validità di questa relazione. Ciò è dovuto alla cancellazione dell'emissione microscopica, che si produce in virtù della identità di Wheeler e Feynman. D'altra parte la relazione fondamentale viene usata

Nota critica. Lavori antichi e lavori moderni. Sembrerebbe spontaneo ritenere che una relazione importante come la (3.3.1), che stabilisce le condizioni di equiliobrio termico tra materia e radiazione, sia stata dedotta modernamente in maniera "rigorosa", nell'ambito della QED (elettrodinamica quantistica), e quindi si trovi esposta in tutti i moderni manuali. Semnbra tuttavia che le cose non stiano così. Il problema sembra essere ignorato, forse a causa del fatto che, come mette in enfasi Dirac stesso nella ultima pagina del suo libro (Priciples ..., edizione del 1954) la formulazione attuale della QED é adatta allo studio di situazioni di scattering (come nell'effetto Compton), e non allo studio dell'interazione radiazine–materia *in bulk*.

La prima deduzione della legge di Wien

Come abbiamo già detto, nei suoi primi due lavori del 1900 Planck considera il campo elettromagnetico in equilibrio termico con un sistema di risonatori. Quindi il sistema totale (in un dominio di volume V fissato) avrà una energia della forma

$$E^{tot} = \int_V u \, \mathrm{d}V + \sum_i U_i$$

dove U_i è l'energia di un singolo risonatore, e u l'energia, per unità di volume, del campo. Analogamente il sistema avrà anche una entropia totale della forma

$$S^{tot} = \int_V s \, \mathrm{d}V + \sum_i S_i$$

dove S_i è l'energia di un singolo risonatore e s l'entropia, per unità di volume, del campo. In seguito Planck si limiterà a considerare quantità relative ad una frequenza fissata ν, che svolge il ruolo di un parametro, che non viene neppure indicato esplicitamente.

Riduzione alla termodinamica dei risonatori materiali

A questo punto Planck compie un astuto procedimento tramite il quale riesce ad esprimere l'energia u_ν e l'entropia s_ν (per unità di volume e di frequenza) del campo attraverso l'energia U e l'entropia S di un singolo risonatore.[28] Sostanzialmente le cose vanno nel modo seguente.

Planck ammette che vi sia uno stato di equilibrio termodinamico, in cui l'entropia totale del sistema presenta un massimo. Se ora con una fluttuazione si produce (nel sistema isolato) uno scostamento ΔU dell'energia di un risonatore, l'entropia del sistema totale assumerà un valore minore. Allora, partendo

in maniera essenziale nel lavoro di Einstein del 1907 sui calori specifici, dove produce risultati in ottimo accordo con i dati sperimentali. La soluzione di questo paradosso potrebbe essere legata alla distinzione che si deve compiere in ambito classico tra energia meccanica ed energia termodinamica dei risonatori. L'energia interna termodinamica dovrebbe essere la porzione dell'energia meccanica che viene trasferita nei processi termodinamici.

[28] Vedremo che sarà invece l'entropia del campo la quantità su cui Einstein concentrerà la sua attenzione nell'articolo del 1905 (sui fotoni), mentre vedremoqui sotto che nel suo procedimento Planck elimina l'entropia del campo, esprimendola in termini dell'entropia dei risonatori.

dal nuovo stato, il sistema tenderà a ritornare ad uno stato cui competa il valore massimo dell'entropia, e dopo un tempo $\mathrm{d}t$ l'entropia totale del sistema sarà cresciuta, diciamo di una quantità $\mathrm{d}S^{tot}$. Planck calcola questo incremento, proporzionale al prodotto $\Delta U\,\mathrm{d}U$, mediante uno sviluppo in serie di S^{tot} al secondo ordine attorno al punto di equilibrio.

In particolare, dalla condizione di massimo egli deduce un importante risultato per l'entropia. Egli trova infatti che s_ν ed S sono mutuamente proporzionali, esattamente tramite il medesimo fattore che si incontrava nella relazione fondamentale di proporzionalità tra u_ν ed U. Ovvero si ha

$$s_\nu = \frac{8\pi\nu^2}{c^3} S \,. \tag{3.3.2}$$

Ora, per la stazionarietà, i termini coinvolgenti le derivate prime si compensano, sicché i primi termini dello sviluppo di Taylor dell'incremento dell'entropia totale coinvolgono solo le derivate seconde, anzi (per la proprietà di proporzionalità sopra menzionata), coinvolgono solo la derivata seconda dell'entropia del risonatore materiale. Precisamente si trova

$$\mathrm{d}S^{tot} = \frac{3}{5}\,\mathrm{d}U\,\Delta U\,\frac{\mathrm{d}^2 S}{\mathrm{d}U^2} \,. \tag{3.3.3}$$

A questo punto Planck ha già ottenuto un risultato interessantissimo di carattere generale, che noi non abbiamo mai trovato da nessuna altra parte, Si tratta del fatto che la proprietà di massimo dell'entropia totale comporta che si abbia (poiché $\mathrm{d}U$ e ΔU hanno evidentemente segni opposti)

$$\frac{\mathrm{d}^2 S}{\mathrm{d}U^2} \leq 0 \,, \tag{3.3.4}$$

cioè l'entropia $S(U)$ del risonatore materiale è convessa (ovvero concava verso il basso).[29] Si controlla subito che questa proprietà è equivalente alla proprietà che il calore specifico sia positivo.

In particolare, indipendentemente dall'ultima osservazione, risulta che per studiare la termodinamica del sistema totale basta studiare quella che *formalmente* è la termodinamica di un sistema di risonatori materiali, tutti della stessa frequenza.

Le "formule termodinamiche" per un singolo risonatore materiale

Il problema è ora di determinare la termodinamica di un singolo risonatore, avente frequenza fissata ν, che seguendo Planck neppure indichiamo.

[29] In particolare, in generale l'entropia del risonatore non ha massimi o minimi. Quella che ha un massimo è l'entropia del sistema totale rispetto a cambiamenti degli stati dei costituenti.

Naturalmente, come sappiamo dalla termodinamica, a tal fine basta assegnare la forma analitica (detta *relazione fondamentale*) di una funzione termodinamica, che può essere una delle seguenti:

$$U = U(T)\,, \quad S = S(U)\,,$$

o delle loro inverse. Naturalmente, dalla prima legge della termodinamica (riferita al caso in cui la frequenza è fissata, analogo al caso in cui è fissato il volume per un gas) si ha che, essendo $\delta Q = T\mathrm{d}S$, vale

$$\mathrm{d}U = T\mathrm{d}S\,, \quad \text{ovvero} \quad \frac{\mathrm{d}S}{\mathrm{d}U} = \frac{1}{T}\,, \tag{3.3.5}$$

sicché come relazione fondamentale si può anche scegliere la relazione funzionale $T = T(U)$ o anche

$$\frac{1}{T} = \frac{1}{T}(U)\,. \tag{3.3.6}$$

La scelta arbitraria di Planck nel primo lavoro

Di fatto fu quest'ultima la forma scelta da Planck. Infatti, indotto dalle sue considerazioni riguardanti il massimo dell'entropia, egli assegnava la forma funzionale della funzione[30]

$$S''(U) \equiv \frac{\mathrm{d}^2 S}{\mathrm{d}U^2} = \frac{\mathrm{d}}{\mathrm{d}U}\frac{1}{T}\,.$$

Che dalla conoscenza della funzione $S''(U)$ si ricostruisca $U = U(T)$ si vede subito. Infatti, in virtù della seconda delle (3.3.5) si ha allora

$$\frac{\mathrm{d}}{\mathrm{d}U}\frac{1}{T} = S''(U)\,.$$

Per integrazione si ottiene

$$\frac{1}{T} = \int S''(U)\mathrm{d}U \equiv G(U)\,,$$

da cui per inversione si ottiene infine $U = U(T)$

Dunque ad ogni scelta della funzione $S''(U)$ corrisponde una ben definita funzione $U(T)$. Sicuramente procedendo per tentativi (lo dice egli stesso) Planck aveva osservato che la scelta [31]

$$S''(U) = -\frac{\alpha}{U} \tag{3.3.7}$$

[30] Si veda il primo lavoro, paragrafo 17, pag. 102 della traduzione. Si legge forse meglio nel secondo lavoro (paragrafo 5, pag. 132 della traduzione).

[31] In effetti, egli aveva scritto la relazione in maniera equivalente ma formalmente diversa, ammettendo che fosse

$$S(U) = -\frac{U}{a\nu}\log\frac{U}{eb\nu}$$

dove $e = 2.718\ldots$ è ovviamente la base dei logaritmi naturali (l'equivalenza si controlla subito con una doppia integrazione, a meno di una costante che viene scelta nulla).

dove α è una costante (rispetto a T) positiva produce la legge di Wien.

L'esercizio è banalissimo. Infatti, integrando la relazione

$$\frac{\mathrm{d}}{\mathrm{d}U}\frac{1}{T} = -\frac{\alpha}{U} ,$$

si ottiene

$$\frac{1}{T} = -\alpha \log U + c ,$$

ovvero

$$U = \gamma\, e^{-1/(\alpha T)} .$$

Si usa poi il fatto che, in virtù della relazione di proporzionalitá tra u_ν e U, la legge di spostamento di Wien prende per un singolo risonatore la forma

$$U = \nu f(\nu/T) , \tag{3.3.8}$$

sicché si ottiene

$$U = b\nu\, e^{-\nu/(aT)} ,$$

ovvero con notazioni moderne,

$$U = h\nu\, e^{-h\nu/(k_B T)} ,$$

che corrisponde alla legge di Wien

$$u_\nu = \frac{8\pi h\nu^3}{c^3}\, e^{-h\nu/k_B T} . \tag{3.3.9}$$

La deduzione della legge di Wien nel secondo lavoro del 1900

Dunque, nel primo lavoro Planck aveva semplicemente introdotto a mano la forma funzionale della funzione $S''(U)$ relativa ad un singolo risonatore, che produceva la legge di Wien, assolutamente senza alcuna giustificazione. Nel secondo lavoro Planck fornisce una dimostrazione di quella forma funzionale, passando a considerare le condizioni di equilibrio per un sistema costituito dal campo e da un numero arbitrario di risonatori della stessa frequenza. Si veda il paragrafo 6 del secondo lavoro, pag. 133 della traduzione di Paolo Campogalliani. Riportiamo qui l'argomento di Planck come a noi attualmente riesce di leggerlo.

Dopo le considerazioni svolte per il caso di un singolo risonatore, Planck passa a considerare un sistema di n risonatori in equilibrio con il campo alla temperatura T, ammettendo che essi compiano i medesimi processi, **indipendentemente l'uno dall'altro**. È questo l'elemento di profonda analogia con la termodinamica del gas perfetto, che sarà portata a compimento nell'articolo di Einstein del 1905 (relativo al corpo nero, nell'approssimazione di Wien) in cui egli studia la dipendenza dell'entropia dal volume.

Planck continua a denotare con U l'energia *media* di un singolo risonatore, e con S la sua entropia, mentre denota con U_n l'energia del sistema di risonatori e con con S_n la corrispondente entropia. Egli riesce a stabilire due relazioni

funzionali che legano S_n ed S, che gli permettono poi di dedurre la legge di Wien.

Le due relazioni funzionali sono

$$S_n'' = S'' \,, \tag{3.3.10}$$

dove l'apice denota derivazione rispetto all'argomento della funzione, e

$$S_n''\Big|_{nU} = \frac{1}{n}\, S''\Big|_U \,. \tag{3.3.11}$$

Da queste segue immediatamente una relazione funzionale per la funzione S'', ovvero

$$S''\Big|_{nU} = \frac{1}{n}\, S''\Big|_U \,. \tag{3.3.12}$$

Dimostrazione. Queste relazioni ruotano attorno alla relazione (3.3.3), che esprime l'incremento di entropia che subisce il sistema totale durante il processo di ritorno all'equilibrio, quando sia accaduta una fluttuazione che ha portato il sistema fuori dall'equilibrio con uno scambio ΔU di energia e successiva variazione $\mathrm{d}U$ nel tempo $\mathrm{d}t$.

Nella relazione (3.3.3) si sta considerando il caso in cui vi sia un solo risonatore, e l'incremento di entropia totale del sistema sia espresso in termini dell'incremento di entropia del singolo risonatore. Se invece di un singolo risonatore ci trovassimo in presenza di un sistema di n risonatori ad energia U_n il cui incremento di energia fosse ΔU_n, analogamente si avrebbe

$$\mathrm{d}S^{tot} = \frac{3}{5}\, \mathrm{d}U_n\, \Delta U_n\, S_n'' \,,$$

dove S_n è l'entropia del sistema di risonatori. Allora Planck assume che l'incremento di entropia del campo sia il medesimo che si avrebbe nel caso di un singolo risonatore avente la stessa energia U_n che ha il sistema di risonatori e le stesse energie scambiate ΔU_n, $\mathrm{d}U_n$: importa quanta è l'energia che il campo scambia con il sistema di risonatori materiali, e non il dettaglio di quanti siano i costituenti del sistema materiale.[32] Dunque si ha

$$\frac{3}{5}\, \mathrm{d}U_n\, \Delta U_n\, S_n''\Big|_{U_n} = \frac{3}{5}\, \mathrm{d}U_n\, \Delta U_n\, S''\Big|_{U_n}$$

ovvero $S_n''\Big|_{U_n} = S''\Big|_{U_n}$, ovvero la (3.3.10).

La seconda relazione riguarda ancora l'incremento di entropia del sistema totale quando il sistema di n risonatori, ad energia U_n, scambia energia ΔU_n con il campo. Ma ora il confronto è con l'incremento di entropia che si ha per il sistema totale quando i singoli oscillatori siano ad energia $U = U_n/n$ e scambino ciascuno una quantità di energia $\Delta U = \Delta U_n/n$. Allora l'ipotesi è che l'incremento di entropia ΔS_n del sistema di risonatori sia n volte l'incremento di entropia ΔS del singolo:

$$\mathrm{d}U_n\, \Delta U_n\, S_n''\Big|_{U_n} = n\, \mathrm{d}U\, \Delta U\, S''\Big|_U \quad \text{se} \quad U_n = nU,\ \Delta U_n = n\Delta U \,.$$

[32] Questo argomento sembra l'analogo di quello che si ha nel gas perfetto, quando si considera la pressione totale. Si veda più avanti

Segue quindi $S''_n\Big|_{U_n} = (1/n)\, S''\Big|_U$, ovvero la (3.3.11).

Nelle parole di Planck (con qualche aggiustamento nostro), questo è spiegato nel modo seguente. *"D'altra parte questo aumento di entropia (quello relativo a uno scambio ΔU_n, e il conseguente $\mathrm{d}U_n$ con un sistema di n risonatori) è comunque uguale a n volte la espressione (3.3.3) relativa allo scambio ΔU (e corrispondente $\mathrm{d}U$) con un singolo risonatore, se si pone $U_n = nU$. Poiché n processi uguali tra loro hanno luogo contemporaneamente e indipendentemente l'uno dall'altro."* [33]

Deduzione della legge di Wien.

La relazione (3.3.12) stabilisce una proprietà di omogeneità per la funzione S''. Poiché potrebbe creare confusione il fatto che con l'apice stiamo denotando la derivazione rispetto all'argomento, chiamiamo con il nome F la funzione S'', cioè poniamo

$$F \equiv S'' \ .$$

Allora la (3.3.12) si esprime come

$$F(nU) = \frac{1}{n} F(U)$$

Ma per n grandi si può pensare ad n come una variabile continua, sicché questa relazione esprime la proprietà che F è una funzione omogenea di grado -1, ovvero $F(U) = \alpha'/U$. Tenendo conto del fatto che sappiamo che la funzione $S(U)$ è concava verso il basso, concludiamo dunque che si ha

$$S''(U) = -\frac{\alpha}{U} \ , \quad (\alpha > 0) \ ,$$

che abbiamo già mostrato condurre alla legge di Wien.

Esercizio. Dimostrare[34] che in corrispondenza con la legge di Planck l'entropia risulta essere data da (scriviamo $a \equiv k$, $b \equiv h\nu$)

$$\frac{S}{k_B} = (1 + \frac{U}{h\nu}) \log(1 + \frac{U}{h\nu}) - \frac{U}{h\nu} \log \frac{U}{h\nu} \ . \tag{3.3.13}$$

Esercizio. Mostrare che questa entropia è uguale a quella data da Cercignani.[35]

[33] La motivazione generale che Planck indicato è riportata alla pagina 133 della traduzione di Campogalliani, e prende quasi mezza pagina. Egli dice: *"Un significato fisico determinato lo possiede ... l'aumento di entropia dell'intero sistema ... perché questo costituisce la misura per l'irreversibilità del processo o per la trasformazioni non compensate di lavoro in calore, e su questa grandezza in effetti si può considerare e portare a termine conseguentemente una corrispondente conclusione. Infatti la relazione fisica necessaria tra la variazione di energia e l'aumento di entropia apparirebbe difficilmente comprensibile se non si assumesse che l'aumento di entropia degli n risonatori nella radiazione considerata è completamente determinato dalla loro energia U_n, dalla loro differenza ΔU_n dal valore stazionario e dalla variazione di energia $\mathrm{d}U_n$ nel tempo $\mathrm{d}t$, e di conseguenza che l'aumento di entropia si ottiene se nell'espressione (3.3.3)*— che ricordiamo dà l'incremento dell'entropia totale del sistema corpo nero più singolo risonatore — *si pone ovunque U_n dove appare U."*

[34] NOTA PER GLI AUTORI: verificare

[35] NOTA PER GLI AUTORI. Riportare l'argomento di Cercignani dal lavoro Cercignani, Galgani, Scotti del 1972.

Complementi: Alcuni commenti sulla legge di Wien e sulle ipotesi fatte da Planck per dedurla

- **Rilettura dell'ipotesi (3.3.11) al modo di Einstein. Estensività e legge dei grandi numeri.** La giustificazione della relazione (3.3.11) è limpidissima nella rilettura che ne dà Einstein nei suoi lavori successivi (soprattutto in quello del 1911), e può essere riassunta nel modo seguente. Anzitutto si osserva che la quantità significativa per la termodinamica è il calore specifico, che sostanzialmente è l'inverso di $S''(U)$. Infatti si ha

$$\frac{\mathrm{d}^2 S}{\mathrm{d}U^2} = \frac{\mathrm{d}}{\mathrm{d}U}\frac{1}{T} = -\frac{1}{T^2}\frac{\mathrm{d}T}{\mathrm{d}U} = -\frac{1}{T^2}\frac{1}{C} ,$$

dove abbiamo denotato con C il calore specifico,

$$C \stackrel{\text{def}}{=} \frac{\mathrm{d}U}{\mathrm{d}T} .$$

Dunque la relazione (3.3.10) prende la forma

$$C_n(nU) = nC(U) , \tag{3.3.14}$$

che esprime la estensività del calore specifico nella maniera consueta della termodinamica.

Tra l'altro, questa relazione termodinamica di estensività è coerente con una proprietà fondamentale della teoria delle probabilità (la legge dei grandi numeri), quando si tenga conto della relazione di Einstein–Boltzmann di proporzionalità tra calore specifico C e varianza dell'energia σ_E^2 nell'insieme canonico,

$$\sigma_E^2 = -\frac{\partial U}{\partial \beta} = k_B T^2 C ,$$

di cui parleremo più avanti. Secondo la legge dei grandi numeri, sotto condizioni molto generali la varianza di una somma di variabili aleatorie è proporzionale al numero di addendi, che è proprio la relazione (3.3.14).

- **Ancora sulla relazione funzionale (3.3.11): indipendenza mutua dei risonatori, e analogia col gas perfetto. Commento al modo di Ehrenfest.** Abbiamo visto come la relazione (3.3.11) venga giustificata pensando che la variazione di entropia del campo elettromagnetico sia le medesima, indipendentemente dal numero di risonatori col quale interagisce, per un fissato scambio di energia. Questa ipotesi sembra costituire una formalizzazione dell'ipotesi che, per quanto riguarda l'entropia, i risonatori si comportino come se fossero indipendenti.

Si pensi all'analogia con il caso dei gas perfetti, in relazione alla pressione p. Poiché, fissato il volume, la pressione è proporzionale alla somma delle energie cinetiche del baricentri delle molecole, formalmente sembra aversi che la pressione p_n esercitata da un sistema di n molecole avente energia nU sia uguale alla pressione esercitata da un unico "particellone" avente quella medesima energia nU:

$$p_n(nU) = p_1(nU) \equiv p(nU) .$$

L'analogia con il corpo nero è che anche in tal caso la pressione dipende dall'energia u (per unità di volume), e non dalla distribuzione dell'energia tra le varie frequenze. Infatti vale la legge di Stefan–Boltzmann $p = u/3$. Questa osservazione si trova esposta in una nota a piè di pagina di un articolo di Paul Ehrenfest (del 1914, riportato nel libro di van der Waerden, pag. 90), dove dice: *"Both examples [gas perfetto e radiazione di corpo nero] have the following property in common. The pressure depends only on the energy of the system: it is independent of the distribution of the energy over the different normal modes of vibration or over the molecules. In a cyclic compression, catalytic process, dilatation, adiabatic process, the same amount of work is given to the system or taken from it. For general systems this is no more the case."*

- **Legge di Planck vs indipendenza dei risonatori.** Dunque la legge di Wien sembra descrivere l'approssimazione in cui i costituenti materiali (i risonatori) sono considerati indipendenti. Pertanto la correzione introdotta successivamente da Planck nella sua comunicazione del 19 ottobre 1900 sarebbe quella necessaria per tenere conto di una mutua interazione (o misteriosa mutua interazione, per usare le parole di Einstein del 1925) che sembra esistere tra i risonatori. Si potrebbe presumere che essa sia indotta dalla interazione che ogni risonatore ha con il campo elettromagnetico, il quale interviene come mediatore, ma è a sua volta influenzato dai risonatori stessi.[36] Se si leggono i lavori di Planck, ci si rende conto che egli lavorava nella approssimazione in cui ogni risonatore singolarmente interagisce con il campo, pensato come campo esterno, che è presente in una maniera già costituita, come se i risonatori non esistessero. Sembra più ragionevole concepire il campo e i risonatori come un unico sistema fisico globale. [37] Il tener conto di questo aspetto globale ha condotto a significativi risultati recenti.[38]

- **Il valore positivo della legge di Wien** Tuttavia, il fatto che Planck deduca (e in maniera sorprendentemente semplice e significativa) la legge di Wien costituisce un contributo fondamentale alla teoria del corpo nero, perché sembra mettere in luce l'argomento cruciale, di carattere puramente termodinamico, che stabilisce il decadimento esponenziale sulle alte frequenze. È questo l'elemento che garantisce che l'energia totale del campo, integrata su tutte le frequenze, sia *finita*. Invece gli argomenti statistici, alla Rayleigh, pur fornendo una ottima spiegazione della equipartizione alle basse frequenze, a livello classico producono la cosiddetta *catastrofe ultravioletta*, ovvero la divergenza dell'energia totale (integrata su tutte le frequenze). Si noti che questa catastrofe in un certo senso sparisce nella trattazione quantistica, ma in effetti vi riappare addirittura in una forma più virulenta, perché il contributo dell'energia di punto zero $h\nu/2$ conduce ad una divergenza come ν^3 anziché come ν^2.

[36] NOTA PER GLI AUTORI. Rivedere in questa luce l'espressione della entropia di Cercignani.

[37] A. Carati, L. Galgani, *A critical remark on Planck's model of black–body*, *Int. Journ. of Mod. Phys.* **B 18**, 549-553 (2004).

[38] A. Carati, L. Galgani, *Non radiating normal modes in a classical many-body model of matter-radiation interaction*, Nuovo Cimento **118 B**, 839-849 (2003); A. Lerose, A. Sanzeni, A. Carati, L. Galgani, *Classical microscopic theory of polaritons in ionic crystals, Eur. Phys. J. D* **68**, 35 (2014).

3.4 La comunicazione del 19 ottobre 1900

La comunicazione del 19 ottobre, dal titolo *Su un miglioramento dell'equazione di Wien per lo spettro*, consta di due pagine e poche righe.

Nel primo paragrafo Planck prende atto che i recenti risultati sperimentali sulle isocromatiche del corpo nero nell'infrarosso (riprodotta qui nella Figura 3.2) hanno mostrato che su ogni isocromatica la legge di decadimento di Wien non è valida in generale, come si era creduto, ma ha il carattere di una legge limite valida per alte frequenze e/o le basse temperature. E aggiunge: *"Poiché io stesso ho espresso davanti a questa società l'opinione che la legge di Wien deve essere necessariamente vera, mi sia permesso di spiegare brevemente la relazione tra la teoria elettromagnetica da me sviluppata, e i dati sperimentali."*

Egli allora ricorda come sia sufficiente stabilire la legge per la energia media $U(T)$ di un singolo risonatore (oscillatore materiale) di assegnata frequenza, o equivalentemente la forma funzionale della $S''(U)$. Ricorda poi come egli avesse dato argomenti che sembravano giustificare la relazione (3.3.7), ovvero la

$$S''(U) = -\frac{\alpha}{U} ,$$

che produce la legge di Wien

$$U(\nu, T) = c_1 \nu e^{-c_2 \nu / T} , \quad c_2 > 0 . \tag{3.4.1}$$

Ma i recenti esperimenti condotti sulle isocromatiche nell'infrarosso avevano mostrato che la legge di Wien era *qualitativamente* errata alle alte temperature.

Infatti secondo la legge di Wien (3.4.1) l'energia dovrebbe tendere a una costante quando T tende all'infinito, mentre le nuove misure del 7 ottobre confermavano in maniera irrevocabile una crescita dell'energia (come d'altra parte era stato anche predetto da Rayleigh).

Planck si rende dunque conto che deve rivedere gli argomenti che lo avevano condotto ad ammettere la (3.3.7), ma non ha ancora elementi chiari. Allora prosegue: *"Following this suggestion I have finally started to construct completely arbitrary expressions for the entropy.... I was especially attracted by one of the expressions thus constructed which is nearly as simple as Wien's expression and which would deserve to be investigated since Wien's expression is not sufficient to cover all observations. We get this expression by putting"* (Planck scrive α invece di a, e β invece di b)

$$S''(U) = -\frac{a}{U(b+U)} . \tag{3.4.2}$$

In tal modo, come mostrato subito sotto, egli ottiene per l'energia U del risonatore una espressione che egli neppure scrive, e che noi scriviamo nella forma coerente con la forma usata nel suo secondo lavoro per la legge di corpo nero, ovvero

$$U(T) = \frac{h\nu}{e^{h\nu/k_B T} - 1} . \tag{3.4.3}$$

Infine egli scrive la legge per la densità di energia del corpo nero, che qui riportiamo nella forma in cui egli la mette nel secondo lavoro, ovvero

$$u_\nu = \frac{8\pi\nu^2}{c^3}\,\frac{h\nu}{e^{h\nu/k_BT}-1}\,. \tag{3.4.4}$$

Per ottenere la legge di Planck, si procede esattamente come fatto per la legge di Wien, evidentemente usando

$$\frac{1}{U(b+U)} = \frac{1}{b}\Big[\frac{1}{U} - \frac{1}{b+U}\Big]\,.$$

Dunque si ha

$$\frac{1}{T} = \frac{a}{b}\log\frac{b+U}{U} + c$$

dove c è una costante, che deve essere nulla per soddisfare la condizione

$$U \to \infty \quad \text{per} \quad T \to \infty\,.$$

Segue allora

$$e^{b/aT} = \frac{b+U}{U}\,,$$

ovvero, risolvendo rispetto ad U,

$$U = \frac{b}{e^{b/aT}-1}\,.$$

Dunque si hanno le identificazioni

$$b = h\nu\,, \quad a = k_B\,. \tag{3.4.5}$$

3.5 La comunicazione del 14 dicembre 1900

Leggiamo come Planck stesso descrive i quasi tre mesi trascorsi tra le due comunicazioni.[39] A proposito della sua formula empirica ottenuta per interpolazione, egli dice: *"If, however, the radiation formula should be shown to be absolutely exact, it would possess only a limited value, in the sense that it is a fortunate guess at an interpolation formula. I have been trying to give it a physical meaning, and this problem led me to consider the relation between entropy and probability, along the lines of Boltzmann's ideas. After a few weeks of* the most strenuous work of my life, the darkness lifted and an unexpected vista began to appear."

Infatti nel suo lavoro del 14 dicembre Planck dice: *"I suspected that one should evaluate this quantity [la distribuzione di energia tra molti risonatori] by introducing probability considerations ..., the importance of which for the second law of thermodynamics was originally discovered by Mr. Ludwig Boltzmann. This suspicion*

[39] Da *The origin and the development of the quantum theory*, in M. Planck, *A survey of physical theories*, Dover (New York, 1960), pag 106.

has been confirmed. I have been able to derive deductively ... the energy distribution in a stationary radiation state, that is, in the normal spectrum."

Confronto tra il procedimento di Planck e quello di Boiltzmann. Il procedimento di Planck si ispira ai procedimenti statistici di Boltzmann, ma è ben diverso. Infatti Boltzmann si interessava della probabilità degli stati macroscopici, definiti da una successione di numeri di occupazione, e ricercava lo stato macroscopico più probabile, e infine determinava il valore medio dell'energia per quello stato, in funzione della temparatura. Invece Planck si disinteressa delle successioni di numeri di occupazione, e ricerca direttamente quale sia l'energia media. Più precisamente, considera un sistema composto di sottosistemi di oscillatori di diversa frequenza, e ricerca quale sia la piú probabile distribuzione di energia tra questi, a temperatura fissata. La connessione fra il procedimento di Planck del 14 dicembre 1900 e quello di Boltzmann verrà poi resa esplicita da Bose nel suo lavoro del 1924.

Il procedimento "alla Boltzmann" per dedurre la legge di Planck, compiuto da Einstein nel suo lavoro del 1907 sul il calore specifico. Il procedimento di Einstein può essere descritto, a posteriori, nel modo seguente. Egli considera lo spazio μ di un singolo oscillatore di frequenza ν, riferendo tale spazio a coordinate energia e angolo, anziché posizione e momento, e sceglie cellette che prendono tutte le fasi tra 0 e 2π con altezze tutte uguali, pari ad $h\nu$. Questo naturalmente è equivalente a quantizzare l'energia asssumendo livelli di energia $\epsilon_k = k\,h\nu$. In effetti, a parte la scelta del valore $h\nu$ di discretizzazione. questo procedimenbto era usato normalmente da Boltzmann). Allora la funzione di partizione $Z(\beta)$ si calcola subito, come serie geometrica di una opportuna ragione q. Infaffti si ha in quanto

$$Z(\beta) = \sum_{k=0}^{i} nf\,tyk\,e^{-\beta k h\nu} = \sum_k q^k \ , \quad q = e^{-\beta h\nu} \ .$$

Dunque si ha $Z(\beta) = 1/(1-q)$. e un calcolo immediato per $u = \partial_\beta \log Z$ fornisce la legge di Planck.

Invece il procedimento di Planck è ben diverso, e verrà ripreso da Bose quasi in una forma non del tutto benevola confronti di Planck, fino a che verrá ripreso da Einstein direttamente, ma con diversi simboli, nel nel suo secondo lavoro del 1924.

Vediamo dunque come procede Planck. Egli considera un sistema di N risonatori materiali (*Resonatoren*) tutti della stessa frequenza ν, aventi energia totale E; vi saranno poi N' risonatori di frequenza ν' ed energia E', poi N'' risonatori di frequenza ν''' ed energia E'', dove $E + E' + E'' \ldots = E_0$ è l'energia totale del sistema di risonatori.[40]

Come abbiamo anticipato, egli ricerca poi la più probabile distribuzione di energia $E, E', E'' \ldots$ tra i diversi sistemi di risonatori. Nelle sue parole: *"Dobbiamo assegnare la distribuzione di energia tra i diversi gruppi di risonatori, e anzitutto la distribuzione dell'energia E tra gli N risonatori di frequenza ν. Se E è considerata*

[40]In effetti, Planck considera che il sistema totale comprende anche la radiazione. Ma poi nei calcoli si concentra solo sull'energia dei risonatori, e per questo abbiamo qui semplificato la notazione, ignorando la radiazione che egli stesso in effetti tralascia di considerare.

una quantità infinitamente divisibile, allora tale distribuzione è possibile in infiniti modi. Consideriamo tuttavia – e questo è il punto essenziale di tutto il calcolo – che E sia composta di un numero ben definito di parti uguali [leggi, numero di quanti o di fotoni] *e ci serviamo a tal fine della costante $h = 6.55\,10^{-27}$ erg sec. Tale costante, moltiplicata per la frequenza comune ν dei risonatori, dà l'elemento di energia ϵ in erg, e dividendo E per ϵ si ha il numero P di elementi di energia* [41], [42] *che devono essere suddivisi tra gli N risonatori. Se il rapporto così calcolato non è un intero, prendiamo per P un intero prossimo. È chiaro che la distribuzione di P elementi tra N risonatori può avere luogo solo in un certo numero finito di modi. Ognuno di questi modi di distribuzione lo chiameremo una "complessione". Se denotiamo i risonatori con i numeri $1, 2, \ldots, N$ e li scriviamo su una riga, abbiamo per ogni complessione un simbolo del tipo seguente:*

1	2	3	4	5	6	7	8	9	10
7	38	11	0	9	2	20	4	4	5

se $N = 10$ e $P = 100$. Il numero di tutte le possibili complessioni è chiaramente uguale al numero di tutti i possibili insiemi di numeri che si possono ottenere in tal modo per la successione inferiore, per assegnati N e P. Per eliminare ogni equivoco, osserviamo che due complessioni devono essere considerate diverse se le corrispondenti successioni contengono gli stessi numeri, ma in ordine diverso.[43] *Dal calcolo*

[41] Oggi diremmo "numero di quanti".

[42] Si veda anche il contributo di Einstein alla Conferenza Solvay del 1911, pag 250 dell'edizione italiana, dove, per la parte quantistica delle fluttuazioni, introduce l'energia *totale* dei fotoni anziché quella del singolo risonatore (parla di $E = nN\bar{E}$ anziché di $\bar{E}$. Inoltre aggiunge "Ciò che vi è di inesatto, nella forma attuale della teoria dei quanti, discende forse dal fatto che in essa questa limitazione del numero degli stati possibili è stata considerata come una proprietà dei gradi di libertà presi *individualmente*."

[43] Questo è il punto in cui il conteggio differisce da quello di Boltzmann relativo alle molecole. Oggi diremmo che Boltzmann considera le molecole distinguibili, mentre Planck considera gli elementi di energia – oggi diremmo i fotoni – indistinguibili. Ciò è ben naturale dal punto di vista di Planck. Perché in termini classici, quando diciamo che l'energia totale $P\epsilon$ del considerato sistema di risonatori si distribuisce tra gli N risonatori, non pensiamo di dire che la porzione di energia che va ad esempio al terzo risonatore sia distinguibile dalla porzione di energia che va ad esempio al quinto, se non per il suo valore (quanti fotoni, e non quali fotoni). La relazione con il conteggio di Boltzmann apparirà più chiara quando commenteremo il procedimento di Bose del 1924.

combinatorio si ha per il numero di tutte le possibili complessioni[44, 45, 46]

$$R = \frac{(N+P-1)!}{(N-1)!P!} \tag{3.5.1}$$

ovvero, con sufficiente approssimazione

$$R = \frac{(N+P)^{N+P}}{N^N P^P} \ . \tag{3.5.2}$$

Compiamo lo stesso calcolo per i risonatori degli altri gruppi, determinando per ogni gruppo di risonatori il numero di possibili complessioni relativo all'energia data al gruppo. Il prodotto di tutti i numeri così ottenuti ci dà il numero totale di tutte le complessioni relativo all'assegnata distribuzione di energia tra tutti i risonatori. Occorre dunque determinare la distribuzione $P, P', P''\ldots$ che massimizza il prodotto $RR'R''\ldots$ (o, equivalentemente, il suo logaritmo) con il vincolo che sia fissata l'energia totale[47] $NP\epsilon + N'P'\epsilon + N''P''\epsilon\ldots$*; per il noto metodo del moltiplicatore di Lagrange, basta allora massimizzare ...*"

Nel lavoro del 14 dicembre viene solo dato il risultato del calcolo della massimizzazione. Comunque il calcolo è banalissimo. Usiamo la notazione $\{P_i\}$ in luogo di $P, P', P''\ldots$, $\{N_i\}$ in luogo di $N, N', N''\ldots$, ed $\{\epsilon_i\}$ in luogo di ϵ, ϵ', ϵ'', ... Tenendo conto del vincolo $\sum_i P_i\epsilon_i = E_0$, usando il metodo dei moltiplicatori di Lagrange si deve allora di massimizzare l'espressione

$$\sum\left[(N_i+P_i)\log(N_i+P_i) - N_i\log N_i - P_i\log P_i - \beta P_i\epsilon_i\right]$$

rispetto alla successione P_i e un immediato calcolo fornisce $\log(N_i+P_i) - \log P_i = \beta\epsilon_i$, ovvero

$$P_i = \frac{N_i}{e^{\beta\epsilon_i}-1} \tag{3.5.3}$$

[44]Questa formula appare nel lavoro di Boltzmann del 1877. Si tratta del numero totale di celle nello spazio Γ a fissata energia totale E. Per Boltzmann questo è dunque il fattore di normalizzazione da impiegarsi per la probabilità di una successione di numeri di occupazione, a fissata energia totale del sistema considerato. Si tratta della formula $J = ..$ in fondo a pag. 181 vol.II, lavoro n. 42 delle opere. Lui denote $E \equiv \lambda'$, $P \equiv p = E/\epsilon$. Si tratta del classico fattore binomiale, e B. ritiene ovvia la dimostrazione. NOTA PER GLI AUTORI. Studiare questa parte di B. Si noti che nel lavoro di Einstein 1904 $\log R(E)$ sarà l'entropia microcanonica.

[45]Dobbiamo distribuire P oggetti identici (indistinguibili), i grani di energia, in N celle. Una distribuzione (o complessione) viene individuata da P palline disposte in fila, con interposte tra di loro $N-1$ sbarrette, che definiscono la collocazione delle N celle. A partire da una fissata complessione, ne ottengo altre $(N+P-1)!$ permutando quegli $N+P-1$ oggetti (palline e sbarrette). Ma tra tali permutazioni, alcune (proprio in numero di $P!(N-1)!$) sono evidentemente indistinguibili dalla prima.

[46]Questo modo di dimostrare la formula combinatoria di Planck fu illustrato in un lavoro di Ehrenfest dedicato specificamente a tale calcolo. Sembra quasi incredibile. Si confronti il modo in cui tale formula viene ritrovata nell'articolo di Poincaré del 1912.

[47]Planck scrive nU in luogo di $P\epsilon$ e così via. Ma subito dopo sostituisce $NU = P\epsilon$ e così via.

dove β viene identificato con la temperatura inversa.[48] Dunque abbiamo ottenuto il numero P_i di elementi di energia (di quanti, o di fotoni) di frequenza ν_i, e anche l'energia $P_i \epsilon_i$. Infine, in virtù della legge generale di spostamento di Wien, si pone $\epsilon_i = h\nu_i$.

Moltiplicando ora la (3.5.3) per il singolo "elemento di energia" $\epsilon_i = h\nu_i$, si ottiene l'energia U_N del sistema degli N_i risonatori di frequenza ν_i, ed avremo (sopprimendo l'indice i)

$$U_N = N \frac{h\nu}{e^{\beta h\nu} - 1} \,. \tag{3.5.4}$$

Possiamo poi considerare l'energia U di un singolo risonatore",

$$U \stackrel{\text{def}}{=} \frac{U_N}{N} \,,$$

e si avrà allora

$$U = \frac{h\nu}{e^{\beta h\nu} - 1} \,. \tag{3.5.5}$$

Nel caso del corpo nero sappiamo che $u(\nu, T)$, la densità di energia (per unità di volume) rispetto alla frequenza, si ottiene moltiplicando la energia U (3.5.5) del singolo "risonatore macroscopico" per il noto fattore

$$N = \frac{8\pi\nu^2}{c^3} \mathrm{d}\nu$$

determinato da Planck, sicché avremo

$$u(\nu, T) = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{\beta h\nu} - 1} \,. \tag{3.5.6}$$

Formalmente, questa relazione viene solitamente letta "alla maniera di Rayleigh", interpretando U non come energia di un risonatore macroscopico, ma come energia media di un singolo "oscillatore del campo elettromagnetico" di frequenza ν (pensata come energia microscopica, analoga dell'energia di una molecola), moltiplicandola poi per il numero N, interpretato come numero di oscillatori del campo di frequenza tra ν e $\nu + \mathrm{d}\nu$.

Commento. Il procedimento di Planck venne ripreso in maniera sostanzialmente identica nei lavori di Bose e di Einstein del 1924–1925. Da un certo punto di vista, tali lavori potrebbero dunque trovare posto in questo stesso capitolo, ma uttavia li rimandiamo al prossimo capitolo. In effetti, in mezzo tra il lavoro di Planck e quelli di Bose ed Einstein si trova un nuovo passaggio logico, cioè quello di considerare il campo elettromagnetico come costituito di particelle identiche (i fotoni), sicché la legge del corpo nero viene pensata da Bose come la legge di un gas di particelle identiche di un tipo particolare. Einstein poi estenderà questa idea al caso di particelle identiche materiali, e nel 1926 Fermi

[48] NOTA PER GLI AUTORI. Spiegare come Planck identifica l'entropia S sicché la derivata rispetto ai P_i (come rispetto all'energia) dà la temperatura e dunque la costante β risulta essere la temperatura inversa, uguale per tutti i diversi gruppi di risonatori (con frequenze diverse.

e Dirac faranno l'analogo per sistemi di particelle identiche (i fermioni) soddisfacenti il principio di esclusione di Pauli.

Dunque il passaggio intermedio è quello di concepire il corpo nero come un gas di particelle identiche. Questo in effetti è il contributo del celebre lavoro di Einstein del 1905. Vedremo tuttavia che questa analogia (corpo nero, gas di fotoni) è legittima soltanto nel limite in cui il corpo nero viene trattato nella approssimazione di Wien. Il "vero" corpo nero, come sistema analogo a un gas costituito di particelle identiche, studiato da Bose, corrisponde ad un sistema di particelle che, usando le parole di Einstein stesso, presentano una "misteriosa" interazione mutua. Altrimenti si troverebbe la legge di Wien.

Commento: i dati sperimentali disponibili sulla legge di Planck per il corpo nero. Può apparire sorprendente il fatto che i dati sperimentali di laboratorio sulla legge del corpo nero sono piuttosto rari. Gli ultimi dati raccolti in maniera sistematica risalgono al 1921 (Rubens e Michel), e si riteneva che rispetto alla legge teorica presentassero scarti dell'ordine dell'uno per cento. Invece è stato mostrato (L. Crovini, L. Galgani, *Lett. Nuovo Cim.*) che lo scarto è dell'ordine del tre percento. La situazione è ancora più strana per quanto riguarda la costante di Stefan. I dati fino a qualche tempo fa scartavano di qualcosa come il 5 percento. Più recentemente è stata eseguita una misura che sembrerebbe dare uno scarto molto più piccolo. Una rassegna dei dati sperimentali per il corpo nero di laboratorio può essere trovata in un articolo pubblicato diversi anni fa.[49]

Una situazione diversa si presenta per i dati sperimentali sulla radiazione di fondo dell'universo, che sembrano seguire in maniera buonissima la legge di Planck. Questo avviene per i dati ottenuti da satellite, perché precedenti dati ottenuti da Woody e Richards (articolo su PRL) da palloni indicavano una fortissima discrepanza sul picco (che gli autori attribuivano a difficoltà nel fissare la taratura, o calibrazione, dei rivelatori alle diverse frequenze – si noti che qui si ha a che fare con una isoterma invece che con delle isocromatiche). Naturalmente, anche per i dati ottenuti da satellite esiste un delicato problema di calibrazione o taratura. Una rassegna di tali metodi sperimentali dovrebbe essere contenuta in un lavoro di Giorgio Sironi, presentato nel 1916 a Princeton in occasione del centenario della scoperta della radiazione di fondo.

La costante di Planck, la velocità delle reazioni chimiche (legge di azione di massa) e la costante additiva dell'entropia. * Parte ancora da scrivere, in cui si discutono i lavori di Sackur e Tetrode del 1911–1913 (compresa la loro premonizione del principio di indeterminazione di Heisenberg). Si veda anche il lavoro di Fermi del 1923 sulla costante additiva dell'entropia, lavoro n. 16, pag. 114, del primo volume delle opere.

[49] L. Galgani, *Annales de la Fondation L. de Broglie*.

Capitolo 4

Einstein e i quanti. Inizia la dualità onda corpuscolo. I calori specifici dei solidi. Le statistiche di Bose–Einstein (e di Fermi–Dirac)

L'ipotesi dei quanti "spiegava" in qualche modo lo spettro del corpo nero, ma da una parte era in palese contrasto con tutta la tradizione della fisica, dall'altra era tutt'altro che una teoria, era un'ipotesi appunto, ristretta ad un fenomeno particolare. Tra il 1900 ed il 1911 questa ipotesi fu estesa ad altri ambiti di meccanica statistica, riuscendo a "spiegare" molta fenomenologia. Molti di questi risultati furono dovuti ad Einstein, ed in questo capitolo esporremo i suoi contributi più rilevanti.

4.1 Introduzione

Si possono rintracciare diversi filoni nella ricerca di Einstein sulla teoria dei quanti di Planck. *Un primo filone*, relativo al lavoro del 1905, riguarda l'idea che il campo elettromagnetico possa essere concepito come costituito da particelle identiche (oggi dette fotoni), analogamente a quanto avviene per un gas. In effetti, questo fatto viene dimostrato solo nel caso limite di alte frequenze e/o basse temperature, in cui la legge di Planck si riduce a quella di Wien. *Un secondo filone* (lavoro del 1907) riguarda la teoria dei calori specifici dei solidi; in particolare, fu in questo ambito che Einstein concepí la deduzione più compatta della legge di Planck. *Un terzo filone*, culminato nel contributo alla conferenza Solvay del 1911 e iniziato nel 1906 o addirittura nel 1904, riguarda il ruolo delle fluttuazioni dell'energia in meccanica statistica, in particolare in relazione alla legge di Planck.

In questo terzo ambito egli descrive anche come ritenga che la quantizzazione sia solo una via d'uscita provvisoria, nel senso che *"tutto va come se il sistema fosse quantizzato"*, opinione che egli tenne fino alla fine della sua vita. *Un quarto filone*, sviluppato nel 1925, riguarda la formulazione della cosiddetta statistica di Bose–Einstein per particelle identiche, che egli stesso (nel suo secondo lavoro del 1925) espone riprendendo quasi alla lettera il procedimento del lavoro di Planck del 14 dicembre 1900. *Un quinto filone* fu sviluppato in un lavoro del 1917 in cui, lavorando nell'ambito della concezione dell'atomo introdotta da Bohr nel 1913, egli ritrova la legge di Planck come condizione di equilibrio dinamico tra campo elettromagnetico e materia. In particolare vi introduce le sue celebri probabilità di transizione, spontanee e indotte, tra livelli, deducendone anche la seconda legge ($\Delta E = h\nu$) che Niels Bohr aveva postulato nel 1913 come relazione tra frequenza emessa o assorbita e salto quantico di energia. Inoltre, generalizzando il suo risultato del 1905, in tale lavoro egli mette in luce come ai fotoni si debba attribuire non solo un'energia, ma anche un momento, come in seguito fu osservato sperimentalmente nell'effetto Compton (1923).

4.2 Il lavoro del 1905, sui quanti del campo elettromagnetico. Inizia la scoperta della dualità onda–corpuscolo

Potremmo esporre i risultati di questo lavoro di Einstein in forma compatta, descrivendoli con parole nostre. Preferiamo scegliere la strada un po' più lunga che consiste nel seguire quasi passo passo la sua esposizione, come possibile stimolo alla lettura diretta del lavoro originale.[1]

Genesi e riassunto del lavoro

Ci pare utile, tuttavia, dare almeno un brevisimo riassunto del procedimento, che è il seguente. Einstein pare mosso da una qualche intuizione che egli non illustra. Da quanto si legge nel riassunto che egli fa nella sua autobiografia scientifica,[2] è plausibile che egli volesse dare concretezza ai quanti di energia del corpo nero (ovvero del campo elettromagnetico), che erano stati introdotti da Planck solo in maniera formale, immaginando che potessero essere assimilati agli atomi di un gas. Deve allora confrontare le formule termodinamiche che descrivono i "due fluidi", quello del gas perfetto e quello del corpo nero. Dato che era bravo, concepisce che le formule giuste siano quelle dell'entropia, o meglio dell'entropia S nella sua dipendenza dal volume V. Tutti sapevano da sempre la formula per il gas perfetto di n atomi, ovvero

$$S(n, V) = k_B n \log V + C$$

[1] Si veda anche la descrizione di questo lavoro data nel libro di Born e Jordan, Sect. 72, pag. 386 e seguenti.

[2] In Schilpp ed., Einstein philosopher scientist.

(che egli comunque ritrova parafrasando un argomento probabilistico già usato da Boltzmann per il caso più complicato del gas di Van der Waals). Lasciamo al lettore di cercare di trovare una formula analoga nel caso del corpo nero, relativo a un intervallo di frequenze tra ν e $\nu + \mathrm{d}\nu$, con energia E, in un volume V. Se tale calcolo apparirà non banale, la ragione potrà essere compresa dalla lettura di un prossimo paragrafo di questo capitolo, relativo al successivo contributo che Einstein diede in relazione alle fluttuazioni di energia. Avendo certamente compiuto tale calcolo senza venirne a capo, Einstein dovette rendersi conto che la situazione è più semplice se si considera il "fluido corpo nero" nell'approssimazione di Wien, valida per alte frequenze o basse temperature, e quindi si limita a considerare tale caso limite. In tal modo perviene alla formula

$$S(E,V) = k_B \frac{E}{h\nu} \log V + C \ .$$

Dal confronto delle due formule egli conclude che, nella approssimazione di Wien, il corpo nero di frequenza ν si comporta come un gas perfetto di particelle aventi ciascuna energia $h\nu$.[3] Nel celebre lavoro del 1917, precisamente nella prima parte del lavoro, egli attribuirà poi ai quanti di luce (detti oggi fotoni) anche un impulso **p**, di modulo $h\nu/c$.

Ilustrazione del lavoro

Nell'*Introduzione* Einstein osserva che vi è una grande distinzione qualitativa tra campo elettromagnetico, trattato come un continuo, e sistema di un numero anche illimitato di atomi ed elettroni. Infatti, *"nel caso della luce l'energia deve essere concepita come una funzione spaziale continua, ... mentre ... l'energia di un corpo ponderabile ... non può sparpagliarsi in parti arbitrariamente numerose e arbitrariamente piccole."* Invece egli intende mostrare che le osservazioni sul corpo nero, e sui fenomeni di generazione e trasformazione della luce, *"appaiono più comprensibili nell'ipotesi di una distribuzione spaziale discontinua dell'energia luminosa. Secondo l'ipotesi che sarà qui considerata, quando un raggio luminoso uscente da un punto si propaga, l'energia non si distribuisce in modo continuo in uno spazio via via più grande; essa consiste invece di un numero finito di quanti di energia, localizzati in punti dello spazio, i quali si muovono senza dividersi e possono essere assorbiti e generati solo nella loro interezza."*

Il primo paragrafo si legge con grande piacere, e dovrebbe essere ben comprensibile al lettore che ci ha seguito finora. Egli considera una regione dello spazio in cui si trovano molecole, e particelle cariche attratte da molle verso punti fissi (i risonatori che ben conosciamo). Vi è poi naturalmente la radiazione, che egli ammette essere in equilibrio con i risonatori, se *"supponiamo che vi siano presenti risonatori per tutte le frequenze da prendere in considerazione"* (si ricordi il modo in cui l'equilibrio dinamico tra risonatori e campo viene concepito da

[3] NOTA PER GLI AUTORI: aggiungere la citazione dalla utobiografia scientifica.

Planck). Ricorda poi il principio di equipartizione (che egli tuttavia non nomina), e come questo condurrebbe alla legge che oggi chiamiamo di Rayleigh–Jeans. Scrive dunque tale legge, che qui riproduciamo sia nelle nostre notazioni, sia in quelle dell'articolo di Einstein, ovvero

$$u_\nu = \frac{8\pi\nu^2}{c^3} k_B T \ , \qquad \rho_\nu = \frac{R}{N}\frac{8\pi\nu^2}{L^3} T \ ,$$

perché egli denota

$$\rho_\nu \equiv u_\nu \ , \quad L \equiv c \ , \quad N \equiv N_A \ , \ \frac{R}{N} \equiv k_B \ .$$

Osserva poi che questa legge corrispondente all'equipartizione condurrebbe evidentemente per il campo a una energia per unità di volume divergente.

Nel secondo paragrafo egli assume che valga la legge di Planck, considerandola come un fatto di esperienza, *"senza presupporre una teoria per la generazione e la propagazione della radiazione"*. Scrive dunque la legge di Planck, che qui riproduciamo usando le due notazioni, ossia

$$u_\nu = \frac{8\pi h\, \nu^3/c^3}{\exp(h\nu/k_B T) - 1}, \qquad \rho_\nu = \frac{\alpha \nu^3}{\exp(\beta_E \nu/T) - 1}$$

dove abbiamo denotato con β_E quello che Einstein denota con β, per distinguerlo dal nostro $\beta = 1/k_B T$. Dunque si ha

$$\alpha = \frac{8\pi h}{c^3} \ , \qquad \beta_E \equiv \frac{h}{k_B}$$

Egli poi osserva che per grandi lunghezze d'onda o per grandi densità di radiazione si ottiene quella che chiamiamo legge di Rayleigh–Jeans.

Infatti, come ben sappiamo, si ha che la legge di Planck per un oscillatore si riduce a $k_B T$, ovvero alla legge "classica" corrispondente al principio di equipartizione, quando la quantità adimensionale

$$x \stackrel{\text{def}}{=} \frac{h\nu}{k_B T}$$

tende a zero (cioè per basse frequenze e/o per alte temperature):

$$\frac{h\nu}{e^{h\nu/k_B T} - 1} \simeq k_B T \qquad \text{per} \quad \frac{h\nu}{k_B T} \ll 1 \ . \tag{4.2.1}$$

Infine, attraverso la conoscenza sperimentale dei coefficienti α, β_E, egli determina il valore del numero di Avogadro come $6.17 \cdot 10^{23}$, concludendone che *"un atomo di idrogeno pesa* $1/N = 1.62 \cdot 10^{-24}$ *g. Questo è precisamente il valore trovato da Planck, e concorda in modo soddisfacente con i valori di questa grandezza trovati per altre vie."*

Nel terzo paragrafo, riproducendo un argomento di Wien, egli mostra come si presenta per il corpo nero l'analogo della relazione termodinamica che abbiamo ripetutamente usato discutendo il primo lavoro originale di Planck, ovvero

$$\frac{\partial S}{\partial U} = \frac{1}{T} \, .$$

Il procedimento è il seguente. Nel caso del corpo nero, analogamente a quanto avviene per l'energia, anche l'entropia sarà data da una entropia per unità di volume moltiplicata per il volume (che Einstein denota con v anziché con V). Egli compie poi una decomposizione in frequenze analoga alla $u = \int u_\nu \mathrm{d}\nu$, e ora introduce s_ν (che egli denota con $\phi(\nu)$) come l'analoga della u_ν. Ammette poi che la componente della densità di entropia dipenda anche dalla densità di energia u_ν (per lui ρ_ν) (analogamente a come in termodinamica S dipende da U), per cui scrive (qui teniamo ancora le due notazioni)

$$S = V \int s_\nu(u_\nu) \mathrm{d}\nu \, , \qquad S = v \int \phi_\nu(\rho_\nu) \mathrm{d}\nu \, .$$

Qui la dipendenza dal volume V è essenziale. È proprio questo il punto che mancava nei lavori di Planck del 1900, e che permette ad Einstein di rendere piena la analogia con il gas perfetto (che ora sarà un gas di fotoni).[4]

Con passaggi semplici, seguendo un procedimento di Wien che qui non riportiamo, egli ottiene allora l'analoga della $\frac{\partial S}{\partial U} = 1/T$, che evidentemente prende la forma

$$\frac{\partial s_\nu}{\partial u_\nu} = \frac{1}{T} \, , \qquad \frac{\partial \phi_\nu}{\partial \rho_\nu} = \frac{1}{T} \, .$$

Il quarto paragrafo tiene un posto centrale in tutto il lavoro. Egli considera la legge di Planck nel caso limite di grandi ν/T, in cui essa si riduce alla legge di Wien,

$$u_\nu = \alpha \nu^3 \exp(-h\nu/k_B T) \, , \qquad \rho_\nu = \alpha \nu^3 \exp(-\beta_E \nu/T) \, .$$

Da questa, prendendo il logaritmo, si ricava immediatamente[5]

$$\frac{1}{T} = -\frac{k_B}{h\nu} \log \frac{u_\nu}{\alpha \nu^3} \, , \qquad \frac{1}{T} = -\frac{1}{\beta_E \nu} \log \frac{\rho_\nu}{\alpha \nu^3} \, .$$

A questo punto egli segue esattamente il procedimento di Planck. Per integrazione della $\frac{\partial s_\nu}{\partial u_\nu} = 1/T$ rispetto ad u_ν, ottiene immediatamente

$$s_\nu(u_\nu) = -\frac{u_\nu}{h\nu/k_B}\left[\log \frac{u_\nu}{\alpha \nu^3} - 1\right] , \qquad \phi_\nu(\rho_\nu) = -\frac{\rho_\nu}{\beta_E \nu}\left[\log \frac{\rho_\nu}{\alpha \nu^3} - 1\right] .$$

[4]Nel libro del 1906 (paragrafo 125, pag. 125), Planck considera una variazione di volume, ma poi si limita allo scambio di energia tra radiazione e risonatore, a V fissato. Qui Einstein completa il discorso.

[5]Si ricordi $\int \log x = x \log x - x$.

Da questa, moltiplicando per $V\mathrm{d}\nu$, ottiene allora l'entropia S della radiazione nell'intervallo $(\nu, \nu + \mathrm{d}\nu)$, ovvero (si ha qui un abuso di notazione,perché poco sopra E es S denotavano l'energia e l'entropia del sistema comprendente tutte le frequenze, anziché solo quelle tra ν e $\nu + \mathrm{d}\nu$; notazioni come E_ν ed S_ν potrebbero essere più appropriate)

$$S = V s_\nu \mathrm{d}\nu \tag{4.2.2}$$

e analogamente ottiene l'energia E della radiazione nello stesso intervallo di frequenze, ovvero

$$E = V u_\nu \mathrm{d}\nu \,. \tag{4.2.3}$$

Si noti che a questo punto sembrerebbe che Einstein già disponesse della dipendenza dell'entropia dal volume. Ma l'espressione in questione è puramente formale, e non ha significato termodinamico. Qui Einstein sembrerebbe compiere uno dei suoi tipici voli pindarici. In effetti, il punto consiste semplicemente nel fatto che il sistema termodinamico di interesse è quello del campo (contenuto nel volume V) con frequenze tra ν e $\nu + \mathrm{d}\nu$, la cui energia ed entropia sono date dalle esperessioni E ed S riportate sopra. L'astuzia di Einstein consiste allora nell'esprimere u_ν in funzione di E, ovvero nel sostituire u_ν con $E/(V\mathrm{d}\nu)$. In tal modo egli ottiene per la funzione $S = S(E, V, \nu)$ l'espressione

$$S = -k_B \frac{E}{h\nu}\Big[\log \frac{E}{V\alpha\nu^3\mathrm{d}\nu} - 1\Big] \;\equiv\; -\frac{E}{\beta_E \nu}\Big[\log \frac{E}{V\alpha\nu^3\mathrm{d}\nu} - 1\Big],$$

Dunque, pensando ad S come funzione di V dipendente parametricamente da E e da ν, si ha

$$S = k_B \frac{E}{h\nu} \log V + \mathrm{cost}$$

Infatti aggiunge:*"Se ci limitiamo a studiare la dipendenza dell'entropia dal volume occupato dalla radiazione e denotiamo con S_0 l'entropia della radiazione quando occupa il volume V_0, otteniamo"*

$$S - S_0 = k_B \frac{E}{h\nu} \log \frac{V}{V_0} \,, \tag{4.2.4}$$

dove noi abbiamo scritto h/k_B in luogo di β_E. La fondamentale conclusione è allora: *"Questa equazione mostra che l'entropia di una radiazione monocromatica di densità abbastanza piccola varia con il volume secondo la medesima legge con cui varia l'entropia di una gas perfetto o di una soluzione diluita."*

Infatti, **nel paragrafo quinto** Einstein ricorda la ben nota formula dell'entropia del gas perfetto o di una soluzione diluita, che egli scrive, alla fine del paragrafo, nella forma (stiamo mescolando le notazioni)

$$S - S_0 = k_B \; n \; \log \frac{V}{V_0} \,, \tag{4.2.5}$$

dove n è il numero di punti costituenti il sistema (gas perfetto o soluzione diluita). Come si vede, le formule (4.2.4) e (4.2.5) coincidono se si sostituisce $E/h\nu$ con n, ovvero se si assume

$$E = nh\nu \,. \tag{4.2.6}$$

Nel sesto paragrafo ottiene la conclusione fondamentale del lavoro. Noi la formuliamo confrontando il modo in cui dipendono dal volume, da una parte l'entropia di un sistema di n particelle costituenti un gas perfetto o una soluzione diluita, formula (4.2.5), e dall'altra l'entropia della radiazione di fissata frequenza, formula (4.2.4), come dedotta dalla legge di Wien.

Dal confronto concludiamo, con Einstein (quasi alla fine del paragrafo), che : *""Sotto il profilo della teoria del calore, una radiazione monocromatica di piccola densità (nel dominio di validità della legge di Wien)* ***si comporta come se consistesse di quanti di energia tra loro indipendenti, di grandezza*** $h\nu$*.".*

Intermezzo: Deduzione della (4.2.5) in termini probabilistici mediante il principio di Boltzmann. A dire il vero, un punto cruciale del paragrafo 5 è che Einstein deduce la formula (4.2.5) non nella abituale maniera termodinamica macroscopica, ma muovendosi in ambito probabilistico, pensando al gas perfetto come a un sistema di n punti indipendenti, con distribuzione di probabilità uniforme (nello spazio fisico racchiuso nel volume V), di modo che, come egli stesso dice alla fine del paragrafo, si vede che *"per ricavare questa legge non è necessario fare alcuna ipotesi sulle leggi di moto delle molecole".* Questo punto è fondamentale, perché non saranno le leggi del moto a produrre l'analogia tra molecole e fotoni, ma le leggi probabilistiche (indipendenza, e uniforme distribuzione spaziale).

Riportiamo qui questa sua interessantissima deduzione di tipo probabilistico, in cui egli segue quasi alla lettera la trattazione data da Boltzmann per il caso più complesso del gas di van der Waals[6]. Il metodo d Boltzmann consiste nel calcolare l'entropia S di uno stato, relativamente a uno stato di riferimento, attraverso le probabilità relativa W del primo stato rispetto al secondo,[7] secondo la formula

$$S - S_0 = k_B \log W \,.$$

Seguiamo ora le sue parole. "*In un volume V_0 siano presenti n punti mobili .... Al sistema considerato competa una certa entropia S_0. ... Immaginiamo ora che tutti gli n punti mobili migrino in una parte V del volume V_0 ... . A questo stato compete evidentemente un diverso valore S dell'entropia. Ci chiediamo: quale è la probabilità del secondo stato rispetto allo stato originario? Ovvero, quale è la probabilità che, in un istante scelto a caso, tutti gli n punti, che si muovono indipendentemente l'uno dall'altro in un volume V_0 assegnato, si trovino (per caso) nel volume V? Questa probabilità ha evidentemente il valore*[8]

$$W = \left(\frac{V}{V_0}\right)^n \,,$$

[6] L. Boltzmann, Lectures on gas theory, Dover 1964, Part II, Section 61, pagina 371.

[7] Nel senso che W è il rapporto delle probabilità dei due stati.

[8] NOTA PER GLI AUTORI. Aggiungere questo punto. Ricordare la legge binomiale e come questa abbia due limiti, a seconda del valore di Np, dove N è il numero di prove ripetute. Se $N \gg 1/p$ si ha Gauss, se $N \ll 1/p$ si ha Poisson. Vedi Doob o Sinai.

e dunque, applicando il principio di Boltzmann, si ottiene $S - S_0 = k_B\, n \log\big(V/V_0\big)$*, ovvero la (4.2.5).*

Nei rimanenti tre paragrafi egli fa poi tre applicazioni fisiche di questa sua concezione della luce come costituita di grani di energia. La seconda applicazione, all'effetto fotoelettrico, è nota a tutti. [9]

La prima riguarda *la regola di Stokes* sulla fotoluminescenza. Il problema è che viene assorbita luce di una certa frequenza ν_1 e ne viene poi emessa dell'altra di frequenza ν_2. Egli osserva che, *"secondo il principio di conservazione dell'energia, l'energia di un quanto generato non può essere maggiore dell'energia di un quanto di luce eccitante: deve quindi valere la disuguaglianza* $h\nu_2 \leq h\nu_1$*, ovvero*

$$\nu_2 \leq \nu_1 \,,$$

che è la ben nota regola di Stokes." La terza applicazione riguarda infine la ionizzazione dei gas mediante luce ultravioletta. Osservando che *"l'energia di ionizzazione di una molecola non può essere superiore all'energia di un quanto di energia luminosa assorbita"*, e ricordando che, *"secondo le misure di Lenard la più grande lunghezza d'onda efficace per l'aria è di circa* $1.9 \cdot 10^5$ *cm"*, egli ottiene *"una limitazione superiore per l'energia di ionizzazione"*.

Osservazione: Legge di Wien e indipendenza; legge di Planck e dipendenza, prescrizione di simmetrizzazione. Abbiamo fatto presente che la legge di Wien è l'analoga della legge dei gas perfetti, nelle quale le molecole del gas sono indipendenti l'una dall'altra. Questo fatto è rispecchiato, come abbiamo detto, nella deduzione della legge dei gas perfetti di tipo probabilistico, che Einstein dà nel suo articolo, in cui le molecole sono supposte indipendenti ed uniformemente distribuite entro il volume V. Quando Einstein ritornerà su questo problema nel 1925, ancora ripeterà che la legge di Wien corrisponde alla indipendenza, sicché la legge di Planck in qualche modo tiene conto in maniera implicita di una certa "misteriosa" influenza reciproca tra i quanti di luce. Abbiamo già detto che, in termini moderni, nel caso della statistica di Bose–Einstein questa influenza reciproca si rispecchia nella prescrizione che le funzioni d'onda del sistema di n particelle sia simmetrizzata.[10]

[9] NOTA PER GLI AUTORI Spiegarla rapidamente e anticipare la somiglianza con Franck e Hertz. Ci sono i granuli di energia del campo e.m., e in Frank ed Hertz ci sono gli stati energetici corrispondenti ai livelli discreti del singolo atomo. N.B. Come mai sembra si debba parlare del singolo atomo o del singolo fotone, mentre Einstein stesso nel 1911 a Solvay dice di no? Vedere se inserire qui una subsection che ora si trova idove si parla del primo congresso Solvay 1911.

[10] NOTA PER GLI AUTORI (sui fotoni). Jeans, nel suo libro *Dynamical theory of gases*, a pag. 379 discute due difficoltà sul concetto di fotoni: 1) interferenza dei fotoni, dovrebbero estendersi per lunghezze dell'ordine di un metro (qualche piede), 2) passano uno alla volta (come riottenuto nei lavori che hanno portato il Nobel attorno al 2012). Normalmente si supera la seconda difficoltà ammettendo (come Dirac nelle prime pagine del suo libro) che ogni fotone interferisce con se stesso. Ma questo si scontra con la difficoltà che i fotoni non sono indipendenti – vedi Planck vs Wien, misteriosa interazione nelle parole di Einstein, dovuta alla simmetrizzazione.

La dualità onda–corpuscolo

Dunque con questo lavoro del 1905 ha inizio la scoperta della dualità onda corpuscolo. Einstein ha mostrato che, nell'approssimazione di Wien, il corpo nero appare come costituito da particelle indipendenti (come un gas perfetto), quelli che oggi chiamiamo fotoni, con energia legata alla frequenza dalla relazione $E = h\nu = \hbar\omega$. Questa scoperta verrà da lui generalizzata in due modi. Il primo modo riguarda il suo lavoro del 1916–1917, noto soprattutto per la deduzione della legge di Planck con il metodo delle probabilità di transizione tra livelli energetici in un sistema materiale. Infatti nella seconda parte del lavoro egli completa la relazione tra energia e frequenza con quella tra momento p e numero d'onda $k = 2\pi/\lambda$, o meglio tra vettore $\mathbf{p}$ e vettore d'onda $\mathbf{k}$. Ricordiamo che, per le onde elettromagnetiche, la cosiddetta *relazione di di dispersione* (ovvero la funzione $\omega = \omega(\mathbf{k})$ ha la forma $\omega = c|\mathbf{k}|$ dove c è la velocità della luce. Ora, nella seconda parte del lavoro Einstein mostra che la luce di frequenza ω può essere pensata come un sistema di particelle che non solo hanno energia $E = \hbar\omega$, ma anche hanno momento (o quantità di moto) avente una definita direzione, e modulo dato da $\hbar\omega/c$. Dunque per i fotoni, pensati come particelle di energia E e momento $\mathbf{p}$, si hanno le relazioni

$$E = \hbar\omega\,, \quad \mathbf{p} = \hbar\mathbf{k}\,,$$

che, in termini relativistici, esprimono una proporzionalità tra i corrispondenti quadrivettori p^μ e k^μ. Con questi semplici elementi nel 1921 Compton interpretò l'effetto che porta il suo nome, come poi avvenne nel 1922 per l'effetto Raman.

Invece, nella sua comunicazione alla prima conferenza Solvay del 1911, Einstein mostrerà come entrambi gli aspetti corpuscolare ed ondulatorio contribuiscano alle fluttuazioni di energia del corpo nero, quando questo sia correttamente descritto dalla legge di Planck, anziché dalla legge limite di Wien. Questa costituisce la seconda generalizzazione del laoro del 1905.

Come è ben noto, e vedremo nella seconda parte, l'idea della dualità onda–corpuscolo verrà ripresa da de Broglie nel 1924, ma in maniera invertita, perché allora saranno le particelle a presentare anche aspetti ondulatori, in aggiunta a quelli corpuscolari che le caratterizzano. Questo porterà nel 1927 alla scoperta della diffraziome degli elettroni dai cristalli (Davisson e Germer) e successivamente alla diffrazione di neutroni (che vennero scoperti solo nel 1932). Ma soprattutto l'idea della dualità onda–corpuscolo condusse Schrödinger, nel gennaio del 1926, a formulare la meccanica quantistica nella cosiddetta forma di meccanica ondulatoria, un aspetto questo che non era stato messo in luce dall'originario procedimento di quantizzazione alla Heinsenberg (luglio 1925), noto come metodo delle matrici.

In effetti il rimbalzare tra particelle e onde nella fisica dei solidi è alquanro affascinante: non soltanto le onde vengono in pratica trattate come particelle, ma anche le particelle costituenti un cristallo possono essere trattate come onde

(onde di densità). Si veda il capitolo 24 del manuale di Ashcroft e Mermin, specialmente nella parte *Wave picture of the interaction of radiation with lattice vibrations*, pag. 482 e seguenti.

4.3 Il lavoro del 1907 sui calori specifici

Dal corpo nero ai solidi

Il lavoro consta di tre parti. La prima è molto interessante di per sé, perché Einstein vi presenta una deduzione della legge di Planck per l'energia media di un oscillatore, con un metodo che è il più economico di tutti quelli possibili, ed è in effetti quello riportato in tutti i manuali. Tale calcolo, che egli compie all'inizio del suo lavoro, viene qui riprodotto alla fine del presente paragrafo.

Da questo punto in poi abbandoniamo la notazione di Planck per l'energia media di un oscillatore, ritornando alla notazione termodinamica consueta, ovvero denoteremo con U l'energia interna di un sistema di N oscillatori, e con $u = U/N$ la corrispondente energia specifica, intesa per singolo oscillatore. Einstein ritorna alla consueta concezione meccanico–statistica in cui l'oscillatore è un oggetto miscroscopico inosservabile, per cui la sua energia risulta inaccessibile. Non ha dunque senso parlarne, se non come energia "per oscillatore" $u = U/N$.[11]

Einstein considera la legge per l'energia media dell'oscillatore materiale come una conseguenza della legge di Planck per la densità di energia u_ν del corpo nero, dovuta all'equilibrio dinamico che si instaura tra radiazione elettromagnetica e risonatori materiali. Infatti, la formula (5) dell'articolo di Einstein è proprio la (3.3.1) di Planck, che nelle nostre notazioni si scrive ora

$$u = \frac{c^3}{8\pi\nu^2}\, u_\nu$$

e gli permette di dedurre la legge

$$u = \frac{h\nu}{e^{\beta h\nu} - 1} \quad , \tag{4.3.1}$$

a partire dalla legge del corpo nero.

Proprio dalla concezione che esista questo equilibrio dinamico tra campo e risonatori scende l'idea centrale del lavoro, ovvero che anche i *"portatori di calore"* del solido debbano seguire l'analoga (4.3.1) della legge del corpo nero, e quindi presentare un calore specifico corrispondente. Il fatto che poi Einstein presenti anche un'altra deduzione indipendente della legge di Planck per un oscillatore, portando a compimento l'idea della discretizzazione dell'energia dell'oscillatore introdotta nel secondo lavoro di Planck, costituisce in qualche modo una aggiunta, anche se di grande interesse, al contributo centrale del lavoro.

[11] Viceversa, per Planck l'energia del risonatore, pensato come oggetto macroscopico, era sperimentalmente accessibile, ed è per questo che, nella notazione, essa era stata promossa ad U.

Nella seconda parte Einstein viene al vero problema fisico che gli interessa, ovvero estendere la legge di Planck dal corpo nero al calore specifico dei solidi. Qui egli compie un volo pindarico. Infatti, dopo avere osservato come Planck abbia dovuto modificare *"la teoria cinetico–molecolare del calore [cioè il principio di equipartizione] per quanto riguarda la interazione fra radiazione e materia in relazione al problema del corpo nero"*, poco sotto egli aggiunge: *"Se la teoria della radiazione di Planck coglie il nocciolo del problema, è presumibile che anche in altri settori della teoria del calore si presentino contraddizioni tra l'attuale teoria cinetico–molecolare e l'esperienza, che si possano rimuovere procedendo per la strada intrapresa. A mio parere ciò effettivamente avviene, come cercherò di mostrare nel seguito."*

Dunque egli ricorda che, se ci rappresentiamo i portatori di calore in un solido come oscillatori armonici (o risonatori), allora per il calore specifico di una mole di solido il principio di equipartizione predice il valore $3R$,[12] ovvero 5.94 calorie, qualunque sia la temperatura. Poi egli ricorda (abbiamo già riportato questo fatto in un precedente paragrafo) che, per quanto riguarda l'interazione tra campo elettromagnetico e materia, ovvero per la fenomenologia della dispersione della luce, secondo Drude si devono ricondurre le frequenze proprie infrarosse di un solido a oscillazioni degli atomi. Dunque, *"Se consideriamo i portatori del calore nei corpi solidi come strutture in oscillazione periodica di frequenza indipendente dall'energia di oscillazione [cioè come oscillatori armonici], stando alla teoria della radiazione di Planck non possiamo attenderci che il calore specifico [di una mole] possieda sempre il valore* 5.94 *[cioè il valore* $3R$ *in calorie]. Dobbiamo anzi porre"* per l'energia interna di ogni oscillatore 3 volte la formula di Planck (4.3.1).

Si noti l'inciso *"stando alla teoria della radiazione di Planck"*. Questo vuol dire che Einstein non sta semplicemente quantizzando l'energia dell'oscillatore secondo quella che oggi chiamiamo la prescrizione della meccanica quantistica, come se questa concernesse l'energia di un sistema meccanico che ubbidisce all'equazione di Schroedinger anziché a quella di Newton. Per lui l'oscillatore materiale eredita la quantizzazione dalla interazione col campo elettromagnetico. Addirittura, nel 1911 Einstein giunge a dire che si tratta di una proprietà *mutua* del sistema campo elettromagnetico ed oscillatore materiale. È a causa della loro interazione che essi sono quantizzati, o meglio che, ai fini dei calori specifici, essi si comportano come se fossero quantizzati.

In ogni modo, la formula per il calore specifico C_V riferita ad un oscillatore pensato come un oscillatore su una retta (sicché, per un atomo, si dovrà assegnare un oscillatore per ciascuna delle tre direzioni spaziali), si ottiene immediatamente derivando rispetto alla temperatura la formula (4.3.1) per l'energia media di un oscillatore. Si trova allora (usando $C_V = -\frac{\partial u}{\partial \beta}/k_B T^2$)

$$C_V = \frac{1}{k_B T^2} \frac{(h\nu)^2 e^{h\nu/k_B T}}{\left(e^{h\nu/k_B T} - 1\right)^2} \quad . \tag{4.3.2}$$

[12] Abbiamo tralasciato il fattore n di Einstein, numero di atomi in una molecola.

Nella terza parte Einstein confronta la curva teorica del calore specifico in funzione della temperatura con diversi dati sperimentali, trovando un accordo qualitativo, e anche quantitativo (almeno parzialmente), molto buono. Egli stesso fa presente (pag 198–199 dell'edizione italiana) che in effetti si trova una certa discrepanza alle frequenze molto alte (o temperature molto basse) dove la curva teorica decade a zero esponenzialmente, molto più rapidamente rispetto ai dati osservati. Questo fatto, ben noto, è oggi spiegato al modo di Debye, facendo riferimento agli oscillatori virtuali corrispondenti ai modi normali del cristallo, da cui si deduce che il calore specifico decade a zero come T^3 (si tratta di una legge analoga a quella di Stefan–Boltzmann per il corpo nero).

Ma questi sono dettagli. L'idea di applicare al calore specifico dei solidi la legge di Planck del corpo nero, trasportandola agli oscillatori materiali in virtù di un equilibrio dinamico tra radiazione e materia, costituisce un grande volo pindarico. Il confronto con i dati sperimentali è emozionante, e meriterebbe di essere commentato paragrafo per paragrafo. Così è anche per l'applicazione che Einstein fa al diamante, la cui frequenza infrarossa non era nota, e viene allora indotta da Einstein attraverso la conoscenza dei valori del calore specifico.

Osservazione: portatori di calore e modi normali. Debye come Rayleigh; Debye vs Raman. Aggiungiamo un ulteriore commento. Oggi sembra si dia per scontato che il calore specifico dei solidi debba essere stimato a tutte le temperature mediante il metodo di Debye, almeno nell'approssimazione armonica. Ovvero, si calcolano i modi normali del cristallo nell'approssimazione in cui si trascurano le interazioni non lineari tra gli atomi o gli ioni costituenti il cristallo, e si assegna ad ogni modo normale (caratterizzato da una sua certa frequenza) il calore specifico secondo la formula di Einstein relativa a quella frequenza. Ma concretamente la procedura che si segue è quasi indipendente dal calcolo dei modi normali (se non per la parte di basse frequenze, in cui si usa la prescrizione di Debye, usando l'approssimazione del passaggio al continuo). Infatti, si prendono in considerazione le frequenze osservate sperimentalmente (per effetto Raman ad esempio, oppure tramite scattering di neutroni od altro), che in generale dipendono dalla temperatura, e poi ad ognuna di tali frequenze si assegna il contributo di Einstein, nel range di temperatura considerato.[13].

Un atteggiamento di questo tipo fu tenuto per tutta la sua vita anche dal fisico indiano Raman. Gran parte del sesto e ultimo volume in cui sono raccolte le sue opere è dedicato ad una controversia che egli sempre tenne sostanzialmente contro tutta la comunità scientifica (compreso Max Born, la cui tesi era sostenuta perfino dal curatore stesso delle opere di Raman). Raman sosteneva che il procedimento consueto, basato sul metodo di Debye, non fosse corretto, e che invece si debba applicare la formula di Einstein, ma soltanto per le alte frequenze, precisamente quelle relative alla cella elementare di un cristallo.[14]

[13] Si veda ....

[14] Raman, Opere, volume 6.

Deduzione meccanico–statistica della legge di Planck per un singolo oscillatore

Illustriamo infine la deduzione della legge di Planck di tipo "meccanico– statistico" (cioè, fatta senza riferimento all'interazione con la radiazione), che Einstein dà nella prima parte del lavoro. Nonostante che all'inizio del lavoro egli si riferisca alla meccanica statistica di un sistema macroscopico costituito da un numero enorme di sottosistemi (sulla linea di due suoi precedenti lavori), egli infine viene a considerare un sistema costituito di un solo oscillatore, con distribuzione di Maxwell–Boltzmann, e ricorda come allora si ottenga per la sua energia media il valore $k_B T$ (in accordo col principio di equipartizione). Egli poi rilegge il procedimento del lavoro di Planck del 14 dicembre 1900 in qualche modo forzandolo, cioè assumendo che l'oscillatore possa veramente assumere solo le energie $\varepsilon_n = nh\nu$ (come oggi fanno tutti i manuali, riferendosi agli autovalori dell'hamiltoniana di un oscillatore armonico).[15]

Allora Einstein assume che la probabilità di tali livelli sia data formalmente ancora da una distribuzione di Maxwell–Boltzmann, ristretta però ai solo valori ammessi dell'energia, ovvero sia data da[16]

$$p_n = \frac{e^{-\beta\varepsilon_n}}{Z(\beta)} \quad \text{con} \quad \mathscr{Z}(\beta) = \sum_n e^{-\beta\varepsilon_n} \ .$$

Infine calcola il valor medio dell'energia

$$u \equiv \langle E \rangle \stackrel{\text{def}}{=} \sum_n \varepsilon_n p_n$$

con un procedimento alquanto curioso, che egli ripeterà esattamente identico nel suo contributo alla conferenza Solvay del 1911. Qui riportiamo il calcolo che si compie oggi su tutti i manuali. Si osserva che, come per il caso continuo, il valor medio u è evidentemente dato dalla consueta formula $u = -\partial_\beta \log \mathscr{Z}$, sicchè basta calcolare $\mathscr{Z}$, o addirittura solo il modo in cui $\log \mathscr{Z}$ dipende da β, trascurando eventuali termini additivi costanti. Per fortuna succede che la serie si somma immediatamente, essendo una serie geometrica di ragione a con

$$a = e^{-\beta h\nu} \ ,$$

[15]In effetti, gli autovalori sono $\varepsilon_n = nh\nu + h\nu/2$. Il termine additivo $h\nu/2$ contribuisce all'energia media per un termine $h\nu/2$ che costituisce la cosiddetta "energia di punto zero" (dal tedesco *nullspunkt*, energia allo zero assoluto). Ma tale termine, non dipendendo dalla temperatura, non contribuisce al calore specifico ed è pertanto irrilevante per la termodinamica. Sembra però che abbia qualche rilievo nella teoria dei solidi e nel cosiddetto effetto Casimir, anche se questo punto è tutt'altro che chiaro. La situazione rispetto a questa energia di punto zero è molto ambigua, e non abbiamo qui il tempo di discuterne.

[16]In effetti, Einstein si esprime in termini che sembrano essere equivalenti all'ammettere che la densità di probabilità nello spazio μ sia data da una somma di funzioni δ di Dirac centrate sui valori E_n. Su questo punto ritorneremo quando discuteremo il lavoro di Poincaré sulla necessità della quantizzazione.

e pertanto si ha

$$\mathscr{Z} = \sum_{n=0}^{\infty} a^n = \frac{1}{1-a} = \frac{1}{1-e^{-\beta h\nu}} \ .$$

Dunque $\log \mathscr{Z}(\beta) = -\log(1-e^{-\beta h\nu})$ e

$$u = \frac{h\nu\, e^{-\beta h\nu}}{1-e^{-\beta h\nu}}$$

sicché infine, moltiplicando numeratore e denominatore per $e^{\beta h\nu}$, si ha la (4.3.1). [17]

4.4 Einstein e le fluttuazioni di energia, prima conferenza Solvay del 1911, *La théorie du rayonnement et les quanta*. "Tutto va come se" il sistema fosse quantizzato

Veniamo ora ad uno dei temi dominanti nella ricerca di Einstein, quello legato alle fluttuazioni. Tutti abbiamo sentito parlare dell'articolo del 1905 sul moto browniano, in cui egli studia le fluttuazioni negli spostamenti di una particella. Ma si possono considerare anche gli spostamenti (le fluttuazioni, appunto) del valore dell'energia di un sistema in interazione con un termostato. Queste sono le fluttuazioni che qui considereremo. Il primo contributo di Einstein risale al 1904, al suo secondo lavoro sui fondamenti della meccanica statistica, in cui mise in rilievo una relazione tra calore specifico e fluttuazioni dell'energia, già presente in Boltzmann, che noi chiameremo relazione di Einstein–Boltzmann, a motivo dell'amplissimo rilievo che le attribuì Einstein, facendone uso essenziale

[17] NOTA PER GLI AUTORI. Riportare la dimostrazione in forma compatta. CONFRONTO Einstein–Planck. Einstein procede come Boltzmann (naturalmente, senza passare al continuo). Tra l'altro, citare anche cella nello spazio μ come analoga del principio di indeterminazione, alla Sackur–Tetrode. Boltzmann avrebbe avuto la difficoltà che l'indice i delle cellette e' arbitrario e non si riesce ad avere una espressione di $\epsilon = \epsilon_i$ in funzione di i. Quindi facendo come Boltzmann e calcolando l'energia media si troverebbe

$$u = C \sum \epsilon_i e^{-\beta\epsilon_i} \ .$$

e non si riuscirebbe a fare la somma. Invece Einstein distingue le celle con l'energia stessa e, come Planck, tiene il volume finito. Quindi riesce a fare la somma. Planck, fa diversamente, ma ha un trucco per passare all'energia media e calcolarla. Vedere pero' come nel lavoro del 1925 Einstein riprende il metodo di Planck. Il motivo è che in tal modo evita di fare le cellette nello spazio μ. A colpo si calcola l'energia media U (seguendo Planck). Sempre, la probabilità nella spazio μ e' data dall'esponenziale di MB. Ma vi compare ϵ_i, cioe' i livelli quantistici. Planck evita di parlare dei livelli, e Einstein lo segue quando fa i bosoni. Calcola l'energia media di un pacchetto di oscillatori o sottosistemi, con la formula combinatoria di Planck, che in effetti era già di Boltzmann. N.B. Einstein e' colpito dal lavoro di Bose, perché lui parla delle energie medie dei pacchetti senza conoscere i livelli. Planck si interessa alle energie medie dei pacchetti $U = P$ disinteressandosi dei livelli.

in molti dei suoi lavori. Ad esempio, questa relazione venne reinterpretata in maniera interessantissima nel lavoro del 1908, in cui venne stabilito il legame delle fluttuazioni dell'energia con il valore medio di questa, nel caso della legge di Planck. Questi temi vennero poi ripresi in maniera sistematica nel suo contributo alla conferenza Solvay del 1911. Questo contributo (lavoro 10 della raccolta italiana delle opere) è veramente interessantissimo, perché Einstein vi condensa tutti i suoi pensieri sulla legge di Planck. In particolare, vi manifesta in maniera esplicita e ferma la sua opinione che terrà per tutta la vita (come ribadito nella sua autobiografia scientifica), e che determinò il suo quasi completo isolamento nella comunità scientifica. Ovvero, che la quantizzazione di un *singolo* sottosistema non è necessaria, ma costituisce soltanto una *"via d'uscita provvisoria"*.

Fluttuazioni di energia e calore specifico: la lettura dinamica di Einstein

Fin dal suo lavoro del 1904 sui fondamenti della termodinamica statistica Einstein aveva messo in luce la relazione fondamentale tra calore specifico e varianza dell'energia. che chiameremo relazione di Einstein–Boltzmann,

$$\sigma_E^2 = k_B T^2 C_V \ . \tag{4.4.1}$$

Essa stabilisce il fatto che, per un sistema macroscopico, il calore specifico è sostanzialmente uguale (a meno del fattore $k_B T^2$) alla varianza σ_E^2 della sua energia. Ricordiamo che questa è definita da

$$\sigma_E^2 \stackrel{\text{def}}{=} \langle (E-U)^2 \rangle \tag{4.4.2}$$

dove E è l'energia, considerata come una variabile aleatoria (random variable), ed $U \stackrel{\text{def}}{=} \langle E \rangle$ il suo valor medio.

La relazione (4.4.2) potrebbe essere dimostrata sotto ipotesi molto deboli, e comunque viene ottenuta in mezzo minuto se si assume che si abbia un distribuzione di probabilità canonica (o di Gibbs) nello spazio delle fasi del sistema totale (spazio Γ, il cui punto generico denotiamo con z), ovvero

$$\rho(z) = \frac{e^{-\beta E(z)}}{Z(\beta)} \ ,$$

dove E è l'energia totale e Z il fattore di normalizzazione (funzione di partizione). Formalmente, questa legge è l'analogo della distribuzione di MB nello spazio μ. In tal modo, come sappiamo, riferendosi alla espressione che definisce il valor medio di E in termini di ρ, si verifica immediatamente che si ha

$$U = -\partial_\beta \log Z \ ,$$

e allo stesso modo si verifica subito che si ha

$$\partial_\beta U = -\sigma_E^2$$

Infine, si considera il calore specifico (o la capacità termica – qui confondiamo le due nozioni) definito come la derivata dell'energia media, $\partial_T U$, e si usa

$$k_B T^2 \partial_T = -\partial_\beta \ .$$

Si noti che la relazione (4.4.2), che abbiamo ricavato a partire dalla distribuzione di Gibbs o di MB, vale anche per la corrispondente distribuzione nel discreto, cioè quando la distribuzione viene ristretta ai soli valori quantizzati dell'energia.

Il punto su cui vogliamo mettere l'accento è che questa relazione potrebbe interpretarsi in maniera puramente cinematica, ovvero: nella distribuzione di Maxwell–Boltzmann (o di Gibbs), l'energia non ha un valore fissato, ma è una *random variable* che ha un suo valor medio $< E > \equiv U$ ma può assumere tutti gli altri valori, e lo sparpagliamento di tali valori è caratterizzato dalla varianza σ_E^2. Nella mente di Einstein, invece, tale sparpagliamento ha una origine dinamica, ed è dovuto al fatto che il sistema considerato è in interazione con un termostato con cui scambia energia. Per questo il valore *istantaneo* della sua energia varia col tempo. Nelle parole del suo lavoro del 1904 (edizione italiana, pag. 115), *"In genere il valore* istantaneo *dell'energia E differisce dal suo valor medio di una certa quantità che chiameremo fluttuazione dell'energia"*. E continua: Ponendo $E =< E > +\epsilon$ abbiamo

$$< \epsilon^2 > \equiv \sigma_E^2 = k_B T^2 \ \frac{\partial < E >}{\partial T} \ .$$

Dunque dice valore "istantaneo", ovvero valore assunto al variare del tempo, e non "valore possibile", come si dice quando si parla di una *random variable*.

Questo aspetto è fortemente ribadito in un punto dell'articolo del 1911, dove dice (pag. 251, abbiamo talora cambiato le notazioni). "*Consideriamo un sistema K di capacità termica C che sia in uno stato di scambio continuo con un ambiente di capacità termica infinita a temperatura T. A causa dell'irregolarità dei fenomeni termici elementari, l'energia di K presenta fluttuazioni attorno a un valore medio U, discostandosi da esso di una quantità variabile* $\epsilon \equiv E - U$. *Dal principio di Boltzmann risulta che la media quadratica* σ_E^2 *di questa fluttuazione è data da* $\sigma_E^2 = k_B T^2 C$."

A noi piace leggere questa frase di Einstein come una premonizione del risultato di Kubo sul calore specifico, di cui parleremo in un prossimo capitolo.

Relazione tra fluttuazioni di energia e valor medio per un sistema di oscillatori secondo Planck, secondo Wien e secondo Maxwell–Boltzmann

Vogliamo ora discutere la relazione tra fluttuazioni di energia (varianza σ_E^2) e valor medio U per un sistema di oscillatori. Risulta che questa relazione presenta anche una interessante dipendenza sia dal numero N di sottosistemi, sia dal numero di quanti di energia. Per evidenziare questo duplice aspetto può essere

conveniente considerare dapprima il caso di in singolo oscillatore e solo dopo passare al caso del sistema di oscillatori.

Consideriamo dunque il caso di un singolo oscillatore, con energia media u data dalla formula di Planck

$$u = \frac{h\nu}{e^{h\nu/k_B T} - 1} \ .$$

Per trovare la relazione cercata basta prendere in esame l'espressione (4.3.2) del calore specifico (per particella) che fu data da Einstein nel 1907, e qui riscriviamo nella forma

$$\sigma_E^2 \equiv k_B T^2 C_V = \frac{(h\nu)^2 \, e^{h\nu/k_B T}}{\left(e^{h\nu/k_B T} - 1\right)^2} \ .$$

Seguiamo il procedimento suggerito da Einstein stesso nel suo contributo alla conferenza Solvay (pag. 249 dell'edizione italiana), dove dice *"eliminando T mediante l'espressione di u"*. Dunque, ricordando l'espressione data sopra del valor medio u è evidente che per trovare la relazione con u conviene al numeratore scrivere

$$e^{h\nu/k_B T} = \left(e^{h\nu/k_B T} - 1\right) + 1 \ ,$$

sicché si ottiene subito

$$\sigma_E^2 = h\nu \, u + u^2 \ . \tag{4.4.3}$$

Ora, nei casi limite $h\nu/k_B T \gg 1$ e $h\nu/k_B T \ll 1$ la legge di Planck si riduce rispettivamente a quella di Wien e a quella cosiddetta classica di equipartizione,

$$u^{Wien} = h\nu \, e^{-h\nu/k_B T} \ , \qquad u^{cl} = k_B T \ ,$$

sicché il calore specifico è dato rispettivamente da

$$C_V^{Wien} = \frac{(h\nu)^2}{k_B T^2} \, e^{-h\nu/k_B T} \ , \qquad C_V^{cl} = k_B \ ,$$

e dunque senza alcun calcolo si trova rispettivamente

$$\text{Wien}: \ \sigma_E^2 = h\nu \, u \ , \quad \text{classica}: \ \sigma_E^2 = u^2 \ . \tag{4.4.4}$$

Per confronto con la (4.4.3) si vede dunque come la relazione tra varianza e valor medio dell'energia nel caso della legge di Planck costituisca una interpolazione (con pesi identici) delle analoghe relazioni che valgono nei casi limite di alte e basse frequenze (o basse temperature ed alte temperature). Commenteremo più sotto come Einstein rilegge a questa luce il procedimento di interpolazione seguito da Planck nell'ottobre 1900, aggiungendogli una valenza dinamica che è assolutamente mancante in Planck.

Veniamo finalmente al **caso "fisico" di N oscillatori** (o, se vogliamo, meglio al caso di N/N_A moli), in cui l'energia totale[18] ha la forma

$$U = N \frac{h\nu}{e^{h\nu/k_B T} - 1} \, .$$

In tal caso il calore specifico, come anche la varianza, hanno le espressioni trovate sopra, ora però moltiplicate per N. È dunque evidente che in quello che diventerà il secondo termine viene a mancare un fattore N per potervi introdurre uno dei due fattori U. In conclusione, invece della formula (4.4.3) si trova ora

$$\sigma_E^2 = h\nu \, U + \frac{1}{N} U^2 \, . \tag{4.4.5}$$

Ancora più significativa è la formula per lo scarto quadratico medio relativo σ_E^2/U^2, che prende la forma

$$\frac{\sigma_E^2}{U^2} = \frac{h\nu}{U} + \frac{1}{N} \quad \text{ovvero} \quad \frac{\sigma_E^2}{U^2} = \frac{1}{N_q} + \frac{1}{N} \, , \tag{4.4.6}$$

"dove con

$$N_q = \frac{U}{h\nu}$$

abbiamo denotato il numero di quanti di Planck presenti in media nel corpo" (Einstein, pag. 249) Tra l'altro, come subito si vede, i due termini sono proprio quelli relativi al caso limite di Wien e al caso limite di Maxwell–Boltzmann. Dunque, la dipendenza dal numero di oscillatori (o di moli) è completamente diversa nei due casi limite, e nel caso della formula di Planck si ha una situazione in qualche modo intermedia.

Le parole di Einstein (articolo del 1908, pag 249) sono le seguenti. *"L'equazione (4.4.6) mostra che le fluttuazioni relative dell'energia del sistema dovute all'agitazione termica irregolare dipendono da due cause distinte, corrispondenti ai due termini del secondo membro. La fluttuazione relativa corrispondente al secondo termine, l'unica che dovrebbe esistere in base alla meccanica ordinaria, deriva dal fatto che il numero di gradi di libertà del sistema è finito; essa esprime l'esistenza degli atomi ed è indipendente dall'energia contenuta nel sistema. La fluttuazione relativa corrispondente al primo termine non dipende in alcun modo dal numero dei gradi di libertà, ma solo dalla frequenza propria e dalla quantità di energia presente in media, e si annulla quando questa energia diventa molto grande. Data*

[18] È alquanto significativo che, a differenza che nel lavoro del 1907, qui Einstein non nomini mai l'oscillatore singolo, ma consideri sempre soltanto il caso macroscopico di N oscillatori. Per lui la distribuzione di MB per il singolo sottosistema semplicemente non esiste. L'unica eccezione sembra essere quella dell'inserto nel lavoro del 1907, ma solo nell'inserto, non nella parte centrale cui è veramente rivolto il lavoro. È plausibile che da poco tempo si fosse reso conto di come si possa dedurre la legge di Planck nello spazio μ introducendo i livelli di energia, con un procedimento che poi condusse Bohr alla concezione degli analoghi livelli ad esempio nel caso dell'atomo di idrogeno.

la sua forma, questa fluttuazione corrisponde esattamente all'ipotesi che l'energia sia composta di quanti di grandezza $h\nu$ che vengono trasferiti indipendentemente l'uno dall'altro". Riprenderemo più avanti il seguito di questa citazione.

Nota: Fluttuazioni di energia e fluttuazioni di pressione. Nel caso del corpo nero le fluttuazioni di energia, con la presenza dei due termini indipendenti di Wien e di Maxwell–Boltzmann, hanno un corrispondente nelle fluttuazioni di pressione, che ancora presentano due termini analoghi. Einstein è tornato ripetutamente su questo fatto. Nel contributo alla conferenza Solvay egli vi dedica tutto il paragrafo 7.[19]

L'interpolazione di Planck rivisitata. Anticipazione del classical program: "Tutto va come se" l'energia fosse quantizzata

Quando abbiamo illustrato il modo in cui Planck aveva ottenuto la sua legge mediante interpolazione, essendoci proposti di seguire per quanto possibile almeno le linee del suo procedimento originale siamo stati costretti a tenere un andamento in qualche modo elaborato e tortuoso. Nel lavoro del 1911 Einstein dà di tale interpolazione una rilettura che a posteriori è illuminante.

Si ricorderà che la legge di Planck era stata ottenuta per integrazione di una equazione del primo ordine nell'incognita $U = U(T)$. Il secondo membro di tale equazione era stato inventato da Planck, modificando l'espressione che conduceva alla legge di Wien. Ora Einstein parte dalla constatazione che aveva appena fatto, ovvero, che se vale la legge di Planck per l'energia media U allora il calore specifico (o equivalentemente la varianza dell'energia, in virtù del teorema di Boltzmann) è la somma dei due termini che conosciamo, espressi in funzione di U, uno proporzionale ad U, l'altro proporzionale ad U^2. Ma poiché il calore specifico è la derivata di U rispetto a T, è allora evidente che quella relazione può essere letta come una equazione differenziale (ordinaria) nella incognita $U = U(T)$, che, risolta, fornisce proprio la legge di Planck.

Nelle sue parole. *"Ci si può domandare se l'equazione appena ottenuta per le fluttuazioni esaurisca il contenuto termodinamico della formula della radiazione di Planck, o della formula che fornisce l'energia media di un oscillatore. Si vede facilmente che è proprio così. Se in effetti nell'equazione delle fluttuazioni si sostituisce σ_E^2 con l'espressione che abbiamo trovato dal teorema di Boltzmann*

$$\sigma_E^2 = k_B T^2 C_V = k_B T^2 \frac{\partial U}{\partial T} ,$$

otteniamo per integrazione la formula di Planck". [20]

[19]NOTA PER GLI AUTORI. Nell'edizione italiana, a pag. 211 si parla di "uno specchio libero di muoversi in direzione ortogonale alla sua normale". Forse si potrebbe trattare di uno specchio libero di muoversi in direzione ortogonale alla sua giacitura. Controllare.

[20]Si usa

$$\frac{1}{h\nu U + U^2} = \frac{1}{U(U + h\nu)} = \frac{1}{h\nu}\left(\frac{1}{U} - \frac{1}{h\nu + U}\right) .$$

sicché si deve integrare l'equazione

$$\mathrm{d}(\beta h\nu) = \left(\frac{1}{h\nu + U} - \frac{1}{U}\right)\mathrm{d}U .$$

Dunque ora appare manifesto che si sta compiendo una interpolazione tra i due casi limite di Wien e di Maxwell–Boltzmann.

E immediatamente Einstein aggiunge la sua conclusione, che a nostro avviso costituisce un documento fondamentale per illustrare il suo punto di vista. Egli dice:

"Una meccanica *compatibile con l'equazione delle fluttuazioni dell'energia di un corpo solido ideale deve dunque condurre necessariamente alla formula di Planck per l'energia di un oscillatore."*

In altri termini. Se qualcuno è capace di trovare una meccanica che produca delle fluttuazioni di energia della forma adatta, allora avrà dedotto la formula di Planck. Nella premonizione che Einstein ha di questo fatto, si confrontano due aspetti. Il primo è di tipo potremmo dire cinematico–statistico: si tratta della relazione di Einstein–Boltzmann tra calore specifico e varianza. Il secondo dovrebbe essere di carattere dinamico,[21] e dovrebbe fornire la relazione funzionale tra varianza ed energia media. Quando illustreremo il teorema di fluttuazione dissipazione di Kubo relativo al calore specifico, vedremo che i due aspetti sono in effetti mescolati, e l'aspetto dinamico si manifesterà in maniera formale attraverso l'autocorrelazione temporale dell'energia del sistema di oscillatori. In tal modo la eventuale caoticità dei moti renderà possibile ottenere il valore "classico" mediante l'annullarsi dell'autocorrelazione, mentre la presenza di moti ordinati potrebbe condurre a calori specifici "non classici".

Queste considerazioni di Einstein sembrano indicare che egli non fosse convinto della necessità della quantizzazione, nel senso che riteneva che le quantità microscopiche fossero descrivibili tramite funzioni continue. Ad esempio nella conferenza Solvay, a pag 256 dell'edizione italiana egli dice: *"Quando un corpo scambia energia mediante un meccanismo quasi periodico di frequenza ν, le proprietà statistiche del fenomeno* sono le stesse che se *l'energia si trasferisse per quanti interi di grandezza $h\nu$".* D'altra parte, nella sua autobiografia scientifica si sono diverse frasi che testimoniano come egli abbia tenuto questa posizione fino alla fine della sua vita.[22]

Si vede poi che la costante additiva di integrazione deve essere posta uguale a zero, affinché U diverga al divergere di T.

[21] Si ricordi la citazione di Einstein, in cui dice esplicitamente che la variazione di energia del sistema studiato deve pensarsi indotta dall'interazione dinamica con il termostato.

[22] NOTA PER GLI AUTORI. Vedere anche pag. 250, *"Non è indispensabile ipotizzare l'esistenza di quanti"* e pag. 208, 203 e 185.

4.5 Il lavoro sulle probabilità di transizione del 1917. Nuova deduzione della legge di Planck. Attribuzione di un momento (o quantità di noto) al fotone

Questo lavoro, dal titolo *La teoria quantistica della radiazione* (lavoro 12 della raccolta italiana delle opere) consta di due parti. La seconda (dal paragrafo 4 alla fine) riguarda il fatto che ai quanti di luce di frequenza ν si deve attribuire, non solo l'energia $h\nu$, ma anche il momento (o quantità di moto) come un vettore di modulo $p = h\nu/c$. Questo è molto importante, perché ci dice che in questo senso il fotone si comporta come una particella di ben definito momento. Dunque si potranno studiare gli urti con altre particelle utilizzando la conservazione sia dell'energia sia del momento nella maniera familiare. Fu utilizzando questa idea che in seguito (nel 1923) Compton descrisse ed osservò quello che chiamiamo effetto Compton.

Qui ci occuperemo della prima parte del lavoro, che è molto breve e semplicissima, e che ebbe una grandissima influenza sui lavori del 1925 con cui Heisenberg, Born, Jordan e Dirac fondarono la meccanica quantistica.

Rispetto ai lavori precedenti il 1912 che abbiamo appena illustrato, questo lavoro tiene conto della nuova concezione, introdotta da Niels Bohr nel 1913, che è nota a tutti e discuteremo più ampiamente nella seconda parte di queste note. Si tratta della concezione che

- ogni atomo o ogni molecola o in generale ogni sistema possa esistere indisturbato in certi stati, che Einstein denota con Z_n, aventi energia ε_n (generalizzando la concezione di Planck–Einstein per i livelli dell'oscillatore armonico);

- che il sistema possa "saltare" (è questo il "*quantum jump*") tra due stati con energia ε_n, ε_m, dove ad esempio

$$\varepsilon_m > \varepsilon_n$$

  emettendo o assorbendo un quanto avente la corrispondente energia $\varepsilon_m - \varepsilon_n$.

- Bohr aggiunse poi l'ipotesi che il quanto avesse frequenza $\nu = \nu_{mn}$ data dalla relazione

$$h\nu_{mn} = |\varepsilon_m - \varepsilon_n| \,, \quad \text{ovvero} \quad \nu_{mn} = \frac{|\varepsilon_m - \varepsilon_n|}{h} \,. \tag{4.5.1}$$

  Ciò è in sintonia con la relazione $|\Delta\varepsilon| = h\nu$ per i salti di energia dell'oscillatore armonico di frequenza ν: creare un fotone è equivalente a far saltare di un livello l'energia del corrispondente oscillatore. È questa la cosiddetta seconda legge di Bohr.

Einstein assume le due prime ipotesi e deduce sia la legge di Planck sia la legge di Bohr (4.5.1) che lega frequenza e salto di energia. A tal fine egli considera un sistema costituito di sistemi identici, con i loro livelli di energia, in interazione con la radiazione di corpo nero, e ricerca quali siano le condizioni di equilibrio (a fissata temperatura T) nel processo in cui il sistema atomico e la radiazione si scambino energia (e momento) attraverso i fotoni emessi od assorbiti.

Ovviamente, per procedere occorrerebbe assegnare una dinamica microscopica. Come di consueto, Einstein usa invece un procedimento statistico, e anzitutto assume che all'equilibrio gli stati atomici siano distribuiti con la legge di Maxwell–Boltzmann.[23] Cominciamo considerando il caso semplice (o di mancanza di degenerazione – livelli tutti distinti) in cui lo stato con energia ε_n ha probabilità

$$p_n = Ce^{-\beta\varepsilon_n}$$

dove C è il ben noto fattore di normalizzazione, il cui preciso valore sarà irrilevante nel ragionamento che si farà.

Poi egli assume che vi sia una situazione di equilibrio nello scambio di energia che si produce tra ogni coppia di livelli $\varepsilon_m, \varepsilon_n$ (nelle notazioni, seguiamo Einstein, ponendo $\varepsilon_m > \varepsilon_n$). Si tratta di una ipotesi che solitamente va sotto il nome di *"ipotesi del bilancio dettagliato"*. Questa, nel caso di sistemi discreti in cui si producono salti con assegnata probabilità, é l'analoga dell'ipotesi di reversibiltà per i sistemi retti da equazioni differenziali.[24]

Precisamente, per quanto riguarda la interazione fra sistema atomico e radiazione egli introduce le seguenti ipotesi.

- **Emissione spontanea.** Il sistema atomico nello stato di energia ε_m può spontaneamente cadere su uno stato di energia inferiore ε_n. *"La probabilità che ciò avvenga effettivamente nel tempo* $\mathrm{d}t$ *è data da*

  $$\mathrm{d}W = A^n_m \mathrm{d}t \ ,$$

  dove A^n_m indica una costante caratteristica della combinazione di indici considerata." Con questo processo si ha trasferimento di energia dal sistema atomico al campo. Nelle parole di Einstein, *"La legge statistica ipotizzata corrisponde a quella di una radiazione radioattiva, e il processo elementare presupposto corrisponde ad una reazione in cui vengano soltanto* emessi *raggi γ"*.

- **Emissione indotta ed assorbimento indotto.** Tuttavia nell'interazione tra sistema atomico e campo, il campo (elettrico) compie lavoro sul sistema atomico, e questo *"a seconda delle fasi del risonatore e del campo"* può essere sia positivo sia negativo. In altri termini, il sistema atomico può emettere

[23] Qui, a dire il vero, come abbiamo già discusso a proposito dell'articolo del 1907 sui calori specifici, Einstein considera un caso un poco più generale, ovvero una distribuzione analoga a quella di MB nello spazio μ, ma ristretta ai soli livelli di energia.

[24] Si veda il celebre articolo di Kolmogorov del 1936.

energia, ma può anche assorbirla. Si hanno dunque sia una probabilità di assorbimento sua una probabiità di emissione, date rispettivamente da[25]

$$\mathrm{d}W = B_n^m \rho_\nu \mathrm{d}t \qquad \text{assorbimento}$$

$$\mathrm{d}W = B_m^n \rho_\nu \mathrm{d}t \qquad \text{emissione}$$

dove B_n^m e B_m^n sono opportune costanti e $\rho_\nu \equiv u_\nu$ è la densità di energia del campo relativa alla frequenza ν associata al salto $|\varepsilon_m - \varepsilon_n|$ considerato.

La condizione di equilibrio per la singola coppia m, n (equilibrio dettagliato) viene dedotta da Einstein nel terzo paragrafo del suo lavoro. Basta richiedere che nel tempo $\mathrm{d}t$ gli eventi di emissione e quelli di assorbimento siano dello stesso numero. Naturalmente, il numero di eventi ad esempio di assorbimento è dato dalla probabilità di essere nello stato di energia inferiore moltiplicata per la probabilità di transizione (compiere il salto) verso l'alto. Analogamente per il numero di salti con emissione. La condizione di equilibrio è dunque

$$e^{-\beta\varepsilon_n} B_n^m \rho_\nu = e^{-\beta\varepsilon_m} (B_m^n \rho_\nu + A_m^n) \, .$$

Da questa si deduce anzitutto che si deve avere

$$B_n^m = B_m^n$$

come si vede immediatamente prendendo il limite $T \to \infty$ (in cui gli esponenziali tendono a 1) e richiedendo come di consueto che anche ρ_ν diverga. Si ha dunque immediatamente

$$\rho_\nu = \frac{A_m^n / B_m^m}{e^{-\beta(\varepsilon_m - \varepsilon_n)} - 1} \, .$$

Usando ora la legge di spostamento di Wien si ottengono due conseguenze. Anzitutto si deduce la seconda legge di Bohr $\Delta E = h\nu$, ovvero la (4.5.1). Poi si ricava la relazione

$$\frac{A_m^n}{B_m^m} = \alpha \nu^3$$

dove α è una costante che potrebbe essere determinata dalla condizione che al limite di basse frequenze e/o alte temperature valga la legge di Rayleigh–Jeans. In particolare, si vede così che i coefficienti B di emissione o assorbimento indotti sono determinati dal coefficiente di emissione spontanea A (e viceversa).

Esercizio. Verificare che le stesse conclusioni si traggono se la legge di Maxwell–Boltzmann viene assunta in una forma più generale, ovvero

$$p_n = C \, g_n \, e^{-\beta\varepsilon_n} \, ,$$

dove g_n è un fattore positivo (detto *fattore di degenerazione*), che definisce la molteplicità del livello E_n. Si pensi ad esempio al caso in cui vi siano g_n livelli di energia vicinissima, e li si tratti come se fossero tutti uguali.

[25] Per molti, risulta difficilissimo ricordare la convenzione di Einstein per le notazioni. **L'indice in basso denota lo stato di partenza**, e quindi quello in alto denota lo stato di arrivo

Seconda parte del lavoro: attribuzione di una quantità di moto $\mathbf{p} = \hbar\mathbf{k}$ al fotone

Si ha poi una seconda parte (molto meno nota) del lavoro, che qui non abbiamo tempo di illustrare, in cui Einstein viene ad attribuire al fotone, oltre all'energia $E = \hbar\omega$, anche un momento $\mathbf{p} = \hbar\mathbf{k}$.[26]

4.6 I lavori di Bose ed Einstein del 1924–1925

Abbiamo visto come nel 1905 Einstein interpretò la legge di Planck (nel limite di Wien) in termini di fotoni (come si dice oggi), mostrando come, "forzando" il procedimento formale di Planck, davvero si deve considerare la radiazione elettromagnetica di frequenza tra ν e $\nu + d\nu$ come costituita di quanti di energia $h\nu$.

Compiamo ora un salto, e passiamo all'anno 1923, in cui il fisico indiano (di Dacca) S.N. Bose ritrovò la legge di Planck in un modo molto interessante, che diede luogo alla cosiddetta statistica di Bose–Einstein. Il fatto curioso è che egli inviò il suo breve articolo alla rivista *Philosophical Magazine* che ne rifiutò la pubblicazione. Allora Bose mandò il manoscritto ad Einstein, che lo studiò e lo generalizzò, come ora diremo. Inoltre Einstein stesso tradusse l'articolo in tedesco, e lo fece pubblicare sulla rivista *Zeitschrift für Physik*, dove apparve nel 1924, con l'aggiunta della seguente *"Nota del traduttore: La derivazione di Bose della formula di Planck rappresenta a mio parere un importante progresso. Il metodo qui usato pone le basi anche per la teoria quantistica del gas ideale, come intendo esporre in altra sede"*. Infatti Einstein stesso estese il ragionamento di Bose applicandolo al caso del gas ideale monoatomico, mostrando che anche il gas deve essere quantizzato, e che ciò comporta la celebre condensazione di Bose–Einstein. Nel suo articolo (del medesimo anno 1924), a parte la differenza tra fotoni e particelle di gas, Einstein riprodusse sostanzialmente lo stesso calcolo di Bose (con la sola differenza di dovere fissare anche il numero di atomi oltre che il valore dell'energia).

L'articolo ricevette diverse critiche, particolarmente da Paul Ehrenfest, riguardo il modo di eseguire i conteggi: il problema era se le particelle dovessero essere considerate distinguibili o indistinguibili, indipendenti o dipendenti. Questo problema fu discusso da Einstein in un secondo articolo, in cui fece anche il confronto tra il metodo di conteggio quantistico e quello classico. In particolare, in questo secondo lavoro egli eseguì il conteggio del numero di complessioni non come nel primo articolo, nel quale riproduceva il calcolo di Bose, ma in un modo diverso, che invece è del tutto uguale a quello originario della deduzione di Planck del 14 dicembre 1900 (che abbiamo illustrato nel capitolo precedente), con la sola differenza che Einstein usa nomi diversi. Per semplicità di esposizione,

[26] Egli si esprime dicendo che se la luce cede a una particella una energia $h\nu$, allora le cede anche un momento di direzione ben definita e modulo $h\nu/c$, come richiesto in ambito relativitico.

sia per il contributo di Bose sia per quello di Einstein riprodurremo qui il secondo metodo di Einstein, cioè quello originario di Planck, usando le notazioni di Planck.[27]

Bose

Vediamo dunque la prima parte del procedimento di Bose per ritrovare la legge di Planck del corpo nero. Bose critica la dimostrazione di Planck in quanto farebbe ancora riferimento a nozioni classiche riguardanti il numero di oscillatori tra ν e $\nu+\mathrm{d}\nu$, che sarebbe *"una ipotesi sul numero dei gradi di libertà dell'etere"* che invece può *"essere dedotto solo per via classica"*. Analoga critica egli fa anche all'uso della legge di spostamento di Wien.

Dunque egli prende sul serio l'idea di Einstein che il campo elettromagnetico debba pensarsi (almeno nel limite di Wien) composto di quanti di luce indipendenti, concepiti come particelle[28] aventi una posizione nello spazio, e inoltre momento $\mathbf{p}$ ed energia E con $p^2 = E^2/c^2$, come previsto dalla teoria della relatività.[29] Inoltre, l'energia del quanto è legata alla frequenza da $E = h\nu$ (come era stato mostrato da Einstein), e dunque la relazione tra energia e modulo del momento si esprime come

$$p_x^2 + p_y^2 + p_z^2 = \frac{h^2\nu^2}{c^2} \ .$$

Si noti bene che il campo sparisce e restano solo i fotoni, come particelle. Stranamente, il fattore di Planck e Rayleigh $8\pi\nu^2/c^3$ (che già in tali autori aveva origini molto diverse) riappare quasi in una maniera ancora completamente diversa, che è la seguente. Egli suddivide i quanti di luce (i fotoni) in gruppi, ciascuno relativo ad un intervallo di frequenza tra ν e $\nu+\mathrm{d}\nu$. Poiché fissare la frequenza tra ν e $\nu+\mathrm{d}\nu$ è come fissare l'energia di una molecola tra E ed $E+\mathrm{d}E$ (con $E = h\nu$), egli calcola quale è il volume nello spazio delle fasi (spazio μ) relativo a quanti di frequenza tra ν e $\nu+\mathrm{d}\nu$. Evidentemente, la regione dello spazio delle fasi μ definita da $E < h\nu$, ovvero da $p^2 < (h\nu/c)^2$, ha volume $(4/3)\pi(h\nu/c)^3$.

[27] Gli articoli di Bose e di Einstein, insieme con un successivo lavoro di Schrödinger, sono raccolti e tradotti in italiano nel libro S.N. Bose, A. Einstein, E. Schrödinger, *La statistica quantistica e le onde di materia*, a cura di P. Bernardini, Bibliopolis (Napoli, 1986).

[28] L'idea di assegnare energia a questi "quanti" è sostanzialmente implicita nel lavoro di Planck, e venne rese esplicita nel lavoro si Einstein del 1905. Invece, l'idea che essi fossero anche muniti di momento, davvero come delle particelle, fu avanzata da Einstein nel suo articolo del 1917. In seguito Compton scoprì l'effetto noto con il suo nome, che viene appunto interpretato come corrispondente a collisioni tra particelle ordinarie e fotoni.

[29] Si ricordi che per una particella di massa a riposo (*rest mass m* si ha

$$\left(\frac{E}{c}\right)^2 - p^2 = m^2c^2 \ .$$

Qui si hanno i fotoni, che devono avere massa a riposo nulla ($m = 0$) perché si ammette che si muovano alla velocità della luce.

Dunque, differenziando, si ha che lo strato (*shell*) compreso tra ν e $\nu + \mathrm{d}\nu$ ha volume

$$V 4\pi \left(\frac{h\nu}{c}\right)^2 \frac{h}{c} \mathrm{d}\nu = h^3 V 4\pi \frac{\nu^2}{c^3} \mathrm{d}\nu \ ,$$

dove V è il volume nello spazio *fisico* in cui è racchiusa la radiazione. Egli moltiplica poi per il consueto fattore due che tiene conto delle possibili polarizzazioni. Infine, divide per il volume h^3 di una celletta elementare[30] nello spazio μ, e in tal modo riottiene il noto fattore di Planck–Rayleigh che ora denotiamo con $g_\nu \mathrm{d}\nu$, dove

$$g_\nu = V \frac{8\pi\nu^2}{c^3} \ , \tag{4.6.1}$$

interpretato però ora in un modo completamente diverso, ovvero come numero di celle in cui possono trovarsi i quanti di luce aventi energia tra E ed $E + \mathrm{d}E$ (o frequenza tra ν e $\nu + \mathrm{d}\nu$). Oggi si preferisce dire che il numero (4.6.1) definisce la *degenerazione* del livello di energia $E = h\nu$. Cioè vi sono $g_\nu \mathrm{d}\nu$ livelli aventi tutti energia E ed $E + \mathrm{d}E$, dove g_ν è dato dalla (4.6.1).

Siamo quindi ricondotti alla situazione di Planck. Planck aveva gruppi di N_i oscillatori di frequenza ν_i, (essendo $N_i = (8\pi\nu_i^2/c^3), \mathrm{d}\nu$ il numero di oscillatori di ogni gruppo), e uno stato macroscopico era una successione $\{P_i\}$ dove P_i è il numero di quanti (fotoni) di ogni gruppo. Fissata l'energia totale, egli trovava allora lo stato macroscopico di equilibrio come quello di massima probabilità. Qui si cambiano i nomi. Vi sono gruppi di cellette, con $g_\nu \mathrm{d}\nu$ cellette ciascuno, il medesimo che per Planck, e uno stato macroscopico è una successione $\{P_i\}$ di fotoni in ogni gruppo di cellette. La probabilità di ogni stato macroscopico è data ancora dalla formula di Planck, e quindi si giunge al medesimo risultato.

Einstein

L'idea di Einstein è semplicemente (a posteriori!) di sostituire i fotoni con le molecole di un gas, e per semplificare la trattazione egli considera un gas monoatomico di atomi puntiformi, in cui non esistono dunque gradi di libertà interni (vibrazioni e rotazioni) come avverrebbe per molecole poliatomiche.

Dunque, implicitamente Einstein sta trattando le molecole come se fossero indistinguibili, come avveniva per i quanti (fotoni) di Planck. Fino ad allora i metodi di Planck erano stati applicati ad oscillatori (che si riferivano al corpo nero o ad oscillatori "materiali" come le molecole di un solido), ed estesi anche ai moti rotatori delle molecole poliatomiche. Ma era stato suggerito da Nernst che fenomeni "di degenerazione" dovessero verificarsi anche per le molecole monoatomiche (che hanno solo gradi di libertà traslazionali), perché questo era richiesto dal terzo principio della termodinamica, da lui stesso formulato nel 1906.

[30] Che il procedimento di Planck avesse come cuore proprio il fatto di considerare lo spazio μ come suddiviso in celle elementari di volume h^μ (se μ è il numero di gradi di libertà di un sottosistema) era un fatto acquisito da diverso tempo, certamente dopo il 1911, anno in cui Sackur e Tetrode pubblicarono i loro noti lavori.

Ora, la relazione tra energia E e modulo per momento p di un una molecola monoatomica è quella di una particella puntiforme,

$$p_x^2 + p_y^2 + p_z^2 = 2mE \ .$$

Dunque il volume della regione nello spazio μ avente energia minore di E è ora (formula (1a) del primo lavoro di Einstein)

$$V \frac{4}{3} \pi (2mE)^{3/2}$$

dove V è il volume della regione della spazio fisico in cui è contenuto il gas. Dunque il volume della regione tra E ed $E + \Delta E$ si ottiene differenziando, ed è dato da

$$V 2\pi (2m)^{3/2} E^{1/2} \Delta E \ ,$$

e dunque il numero N di cellette elementari (di volume h^3) contenuto in tale regione è dato da (formula (2) del primo lavoro)

$$N = \frac{1}{h^3} V 2\pi (2m)^{3/2} E^{1/2} \Delta E \ .$$

Einstein commenta: *"Per $\Delta E/E$ fissato piccolo a piacere, si può sempre scegliere V tanto grande che N sia un numero molto grande."*

Questa è anche la formula (2a) del secondo lavoro di Einstein. Per il numero di complessioni egli riporta esattamente la formula di Planck (a parte la scelta dei simboli, di cui si dirà qui sotto). Compie allora la massimizzazione del logaritmo della probabilità con i vincoli $\sum P_i E_i = E_0$ e $\sum P_i = n$ dove n è il numero totale di particelle, e trova $P_i = P_i^*$ dove

$$P_i^* = \frac{N_i}{e^{\alpha + \beta E_i} - 1} \ .$$

Notazioni moderne. Le notazioni di Einstein nel secondo lavoro sono $n_i = P_i$ per il numero di particelle nel gruppo i–esimo di celle, $z_i = N_i$ per il numero di celle nel gruppo i–esimo, di energia E_i. Oggi (seguendo il libro *Atomic Physics*) di Born ,solitamente il numero di celle nel gruppo i–esimo, $z_i = N_i$, viene piuttosto denotato con g_i e detto talvolta *fattore di degenerazione*. Dunque la formula di Planck per il numero di complessioni relativo ad uno stato macroscopico viene scritta

$$W[n_1, n_2, n_3, \ldots] = \frac{(g_i + n_i - 1)!}{g_i! n_i!}$$

e la distribuzione corrispondente allo stato di equilibrio viene scritta

$$n_i(\alpha, \beta) = \frac{g_i}{e^{\alpha + \beta E_i} - 1} \ .$$

Il problema della distinguibilità

Nel secondo lavoro Einstein prende atto di di una critica sollevata da Ehrenfest (ed altri) al suo primo lavoro e a quello di Bose, in cui veniva messo in luce che essi stavano eseguendo dei conteggi non standard (per così dire) riguardo la distinguibilità degli oggetti considerati. A dire il vero, nel caso di Boltzmann e Planck, la cosa era implicita, ma ovvia. Già Boltzmann considerava P "grani" di energia che suddivideva tra molecole (mentre Planck li suddivideva tra N risonatori). E a nessuno sarebbe venuto in mente di considerare distinguibili due grani di energia. L'energia veniva suddivisa in grani (evidentemente da considerarsi indistinguibili) solo per comodità di conteggio.

Ma quando Einstein trasporta il conteggio alle molecole e, parafrasando Bose, conta i modi diversi di distribuire N molecole tra g celle dello spazio μ (celle aventi tutte la medesima energia), allora è evidente che in un conteggio standard si dovrebbero considerare le particelle distinguibili, come già faceva Boltzmann. Il fatto di non distinguerle costituisce un nuovo assioma, che deve essere dichiarato esplicitamente. Questo Einstein lo riconosce. Questo nuovo assioma corrisponde ad una sorta di ipotesi di natura fisica, Einstein dice ripetutamente, e dirà così anche Fermi. Si tratterebbe del fatto che anche un gas perfetto monoatomico presenta una sorta di mutua interazione degli atomi (in probabilità si direbbe di mutua dipendenza), che Einstein stesso chiama "misteriosa". Naturalmente, una tale denotazione, di tipo in qualche modo romantico, scompare assolutamente nella successiva formulazione assiomatica, in cui la "dipendenza" tra le particelle di gas perfetto viene descritta come corrispondente a un procedimento di simmetrizzazione del prodotto di funzioni d'onda per la statistica di Bose–Einstein e di antisimmetrizzazione per la statistica di Fermi–Dirac.

Per comprendere questa situazione, è un utilissimo esercizio rendersi conto di come la formula di Bose–Einstein per il gas perfetto monoatomico ridiventa quella di Boltzmann se si ammette la distinguibilità delle particelle. Il conto è fatto da Einstein nel paragrafo 7 del suo secondo lavoro, prima della formula (29b). Noi riproduciamo il calcolo in un modo leggermente diverso. Pensiamo che i $g_E \mathrm{d}E$ livelli di energia tra E ed $E + \mathrm{d}E$ siano tutti diversi, anche se vicinissimi. Si applica allora il conteggio di Boltzmann (con il consueto contributo $N!/(n_1!n_2!\ldots)$ e si ottengono i numeri di occupazione di MB. Poi, per ottenere l'energia tra E ed $E + \mathrm{d}E$ si moltiplica per il fattore di degenerazione $g_E \mathrm{d}E$ e per l'energia E di ogni livello (cella). Se si procede in maniera analoga per i fotoni, si ottiene la legge di Wien invece della legge di Planck.[31]

[31] NOTA PER GLI AUTORI. Scrivere meglio questa parte. Citare anche come Einstein descrive la condensazione, e la differenza rispetto al gas saturo. Discutere anche il lavoro di Schrödinger, in cui riproduce Einstein, ma introducendo un campo anche per le particelle, una sorta di anticipazione della seconda quantizzazione, ispirata dal campo di de Broglie (possibile analogia con il lavoro Bohr Kramers Slater ?). N.B. Riprendere le citazioni di Paolo Bernardini pag. 55: Hanle, (Arch. Hist.Ex. Sc. 17. 165 (1977)) , Landé (Z. f. Phys. 33, 571 (1925)) sulla Kohärenz della radiazione di corpo nero. N.B. Riprodurre il lavoro di Fermi per la statistica di Fermi–Dirac, con l'esempio del . gas di elettroni nei metalli. Accennare alla generalizzazione di Gentile.

Bosoni e fermioni

Abbiamo visto che Einstein, nella sua geniale (a posteriori, quasi banale) invenzione della termodinamica statistica di un gas, procede per analogia con il modo in cui Bose aveva trattato il corpo nero, mediante i fotoni (che Einstein aveva introdotto solo per il caso limite di validità della legge di Wien). Ma naturalmente i fotoni non seguono il principio di esclusione di Pauli, e per questo motivo si dice che Einstein trasportò il procedimento di Bose dal caso del campo elettromagnetico, con i suoi fotoni, al caso dei gas costituiti di particelle materiali identiche che sono bosoni, cioè particelle che non seguono il principio di esclusione di Pauli, e quindi possono trovarsi in un numero arbitrario su ogni livello di energia. Implicitamente è dunque a bosoni che si riferisce Einstein quando discute della condensazione, che infatti oggi viene chiamata condensazione di Bose–Einstein.

L'invenzione della termodinamica statistica dei gas costituiti da particelle che seguono il principio di esclusione di Pauli è un'altra invenzione geniale, che potrebbe apparire anch'essa, a posteriori, quasi banale. Basta ripetere il ragionamentto di Einstein, imponendo la restrizione che su ogni livello può trovarsi al più una particella. Appena dopo la pubblicazione dell'articolo di Pauli nel gennaio 1925,[32] questa estensione fu compiuta nel 1926, da Fermi, in un lavoro di poche pagine presentato il 7 febbraio, che lo rese per la prima volta celebre nel mondo.[33] [34] Nel lavoro, che è di una semplicità impressionante, egli segue proprio un procedimento come quello appena descritto, e fornisce immediatamente, come è nello stile della sua caratteristica concretezza, le formule per la termodinamica statistica del gas perfetto monoatomico, trovandone l'equazione di stato. Nella presente versione delle note non riportiamo i semplicissimi calcoli, che si trovano in tutti i manuali. ad esempio in relazione all'energia di Fermi nel caso tipico del gas di elettroni in un metallo.

Una trattazione di carattere più generale venne fissata da Dirac in un articolo inviato il 26 agosto dello stesso anno.[35] Per questo motivo la statistica di particelle identiche che seguono il principio di Pauli viene solitamente detta *statistica di Fermi–Dirac*. Bisogna dire tuttavia che una trattazione indipendente, e davvero molto interessante, venne data da Heisenberg[36] l'11 luglio dello stesso anno (quin-

[32] W. Pauli, *Zeit. f. Phys.*, **31**, 765 (1925).

[33] E. Fermi, *Sulla quantizzazione del gas perfetto monoatomico che segue il principio di Pauli*, e *Zur Quantelung des idealen einatomigen Gases*, lavori n. 30 e 31 delle *Note e Memorie*, Vol. I. Si veda anche il lavoro sulla costamte dell'entropia, n. 16, pag. 114.

[34] In tutti i manuali, il procedimento di Fermi viene riprodotto nel caso, che è effettivamente il più semplice e che è stato indicato da Fermi stesso all'inizio del suo lavoro, di un gas ideale racchiuso in una scatola cubica, per il quale i livelli energetici sono particolarmente semplici. Nel suo lavoro, Fermi sceglie in effetti il metodo sostanzialmenete equivalente di considerare un sistema di particelle attratte vero un punto fisso da una forza linerae, cioé un sistema di oscillatori armonici indipendenti.

[35] P.A.M. Dirac, *On the Theory of Quantum Mechanics*, *Proc. R. Soc. A*, **112**, 661 (1926).

[36] W. Heisenberg, *Mehrkörper Problem und Resonanz in der Quantenmechanik*, *Z. Phys.* **38**, 411 (1926), lavoro n. 3.5 pag. 456 del primo volume delle opere. Si veda anche il seguito, Parte II, n. 4.3, pag. 551 del primo volume delle opere. Il tratto caratteristico di questi due lavori di

di prima del lavoro di Dirac). In questo articolo di Heisenberg, rarissimamente citato, viene detto esplicitamente per la prima volta che nella MQ di particelle identiche si ha una situazione di indeterminatezza (*Umbestimmtheit*).[37] Infatti egli si rende conto che sono possibili diverse scelte (simmetrizzazione o antisimmetrizazione, come egli stesso dice, esprimendosi nel linguaggio di Schrödinger), e bisogna rivolgersi alla natura per sapere quale di queste in effetti si presenta.

4.7 Una anticipazione del "classical program" in relazione al principio di equipartizione

NOTA PER GLIL AUTORI; Raccogliere qui i diversi commenti di Einstein alla prioma conferenza Solvay ?

Heisenberg è che egli pensa che la non possibilità di passaggio da uno stato simmetrico a uno stato antiimmetrico abbia origine fisica, nella natura della interazione elettromagnetica. In tale contesto egli discute anche un paradosso su cui aveva attirato l'attenzione Einstein in relazione all'identità delle particelle (forse analogo al noto paradosso di Gibbs). Si veda la pag. 423 del primo lavoro, in cui egli sembrerebbe considerare possibili transizioni su scale di tempo lunghissime.

[37] È questo il più importante risultato del suo lavoro, come egli dice a pag. 416.

Capitolo 5

Poincaré 1912: sulla necessità della quantizzazione. La risposta di Nernst, 1916

L'idea che le quantità fisiche subiscano variazioni discontinue era veramente estranea a tutta la fisica teorica. Essa infatti, a partire da Galileo, basava il suo apparato matematico sul concetto di derivabilità delle osservabili, e dunque sulla loro regolarità. Quindi, nonostante gli indubbi successi dell'ipotesi dei quanti, rimaneva sempre in sospeso la domanda: ma questa ipotesi è veramente necessaria? Sostanzialmente in questi termini si esprimeva Poincaré all'inizio del suo lavoro[1] *"Sur la theorie des quanta"*, scritto al suo ritorno a Parigi da Bruxelles, sotto la profonda impressione delle lunghe discussioni svoltesi alla conferenza Solvay del 1911.[2] [3]

A tale domanda egli rispose in maniera affermativa. Comunque egli espresse anche dei dubbi che il suo teorema fosse la parola definitiva sulla questione, e più avanti vedremo come la situazione non sia così ovvia. Da una parte, infatti,

[1]H. Poincaré, *Sur la théorie des quanta*, Journal de phys. théor. et appliq. **2**, 5–34 (1912); *Oeuvres* IX, pag. 626–653.

[2]La testimonianza più diretta si ha leggendo le circa 20 pagine in cui sono riportate le discussioni che fecero seguito alla relazione di Einstein, pagg. 263–281 della edizione italiana delle opere scelte di Einstein. A questa discussione conribuì anche Poincaré con diversi interventi. Egli poi mandò anche una nota (inviata da Parigi, e pubblicata a pag. 280 dell'ediziome italiana delle opere scelte di Einstein) nella quale annunciava il risultato del lavoro pubblicato nel 1912, che qui discutiamo. Ricordiamo che Poincaré poi morì nello stesso anno 1912.

[3]Due importanti lavori sulla necessità della quantizzazione vennero compiuti anche da Paul Ehrenfest. Si veda particolarmente P. Ehrenfest, Ann. Phys. **36**, 91–118 (1911), dal titolo *Welche Zúge der Lichtquantenhypothese spielen in der Theorie del Wärmestrahlung eine wesentliche Rolle?* , riprodotto in P. Ehrenfest, *Collected Scientific Papers*, North-Holland (Amsterdam, 1959). Si veda anche l'articolo n. 31 (del 1814) dei Colledcted Papers. Bellissima è la introduzione ai Collected Scientific Papers scritta da Casimir, dove egli confronta lo stile di Ehrenfest con quello di Lorentz (che lo aveva chiamato poco più che trentenne a Leyden), e dove commenta che Ehrenfest era *a central figure in a happy era of physics that will not come again!* Si veda anche M.J. Klein, *Paul Ehrenfest: the making of a theoretical physicist*, North Holland (Amsterdam, 1970).

sembra che la conclusione di Poincaré sia basata su una ipotesi che, come vedremo, potrebbe non essere garantita. D'altra parte, dopo il 1963, con la scoperta dello *"strano attrattore di Lorenz"*[4] è stato mostrato che anche sistemi alquanto "semplici", retti da equazioni diffrenziali, e quindi con movimenti lisci, possono presentare asintoticamente (per tempi lunghi) un comportamento che sembrerebbe descrivibile come se il sistema si muovesse non su di una varietà liscia, ma piuttosto su di un insieme "strano", di tipo frattale.

NOTA DIDATTICA. Come accennato nella prefazione alla prima parte, nella presente versione non è ancora stato possibile illustrare dei nuovissimi risultati che mostrano come una ipotesi cruciale introdotta da Poincaré potrebbe non essere giustificata su basi dinamiche. Qui, ci limitiamo, come nella versione precedente, ad una esposizione abbastanza dettagliata del procedimento statistico seguito da Poincaré, basandosi sulla sua ipotesi.

5.1 Introduzione

Di primo acchito l'argomentazione di Poincaré che conduce alla necessitá della quantizzazione sembrerebbe avere il carattere di una dimostrazione per assurdo. Infatti egli ammette che il sistema fisico considerato (un sistema di n oscillatori interagenti con un gas di p atomi) sia descritto da un modello classico, retto da equazioni differenziali ordinarie (come quelle di Newton), per il quale si hanno orbite continue in uno spazio delle fasi ovviamente pensato come un continuo. A questo punto egli, introducendo l'ipotesi che le energie degli oscillatori siano variabli causali di tipo iid (indipendenti ed identicamente distribuite), si riduce a studiare il caso di un singolo oscillatore, e deduce quella che sembrerebbe essere la più generale espressione possibile per la sua energia media. Infine, ammettendo che tale energia coincida con quella data dalla legge di Planck, che rappresenta bene i dati sperimentali, utilizzando la forma in cui viene espressa l'energia media egli conclude che l'energia di un singolo oscillatore deve essere necessariamnente quantizzata.

Questo risultato fu inteso come una dimostrazione definitiva ed irrevocabile della necessità della quantizzazione, ed ebbe immediatamente un grandissimo impattto sulla cmunitá scientifica.

Davvero caratteristico fu l'atteggiamento di Jeans. Questi aveva partecipato alla conferenza Solvay come rappresentante di Rayleigh, illustrando il punto di vista che fosse possibile spiegare la legge di Planck su basi classiche, quando, seguendo Rayleigh e Boltzmann stesso, si tenga conto che possano essere lunghissimi i tempi richiesti per il raggiungimento dell'equilibrio (e quindi del-

[4]Questo Lorenz (che si pronuncia come in italiano Lorenzo) non ha nulla a che fare con il grande Lorentz, e nemmeno con l'altro Lorenz che, indipendentemente da Lorentz, e nello stesso anno, giunse a stabilire la cosiddetta "formula di Lorentz–Lorenz". Questa è una versione elettromagnetica della formula di Clausius–Mossotti, e svolge un ruolo rilevante nella teoria della dispersione della luce.

l'equipartizione) per i gradi di libertà interni alle molecole. Ebbene, dopo la pubblicazione del lavorodi Poincaré, Jeans fece una pubblica ritrattazione (proprio così) ed aderì al punto di vista quantistico. Si veda anche il libro di Jeans sulle teorie cinetiche, e la prefazione a un altro libro scritto intorno al 1917, nella prefazione del quale egli fece ancora sostanzialmente una ritrattazione.[5]

Del lavoro di Poincaré venne pubblicato nel 1921 un ampio riassunto da Lorentz, che evidentemente lo riteneva allo stesso tempo di grande importanza ma di difficile lettura.[6] Si può poi vedere il libro di Jammer, o i tanti articoli pubblicati ad esempio sulla autorevole rivista ufficiale per la didattica della fisica, l'American Journal of Physics.[7]

5.2 L'argomentazione semplificata: la quantizzazione del singolo oscillatore

Prima di illustrare il lavoro di Poincaré in dettaglio, esponiamo il suo ragionamento in forma semplificata, che si applica quando si ammetta di essersi ridotti allo studio dell'energia di un singolo oscillatore. Ci ambientamo dunque nello spazio delle fasi μ di un singolo oscillatore la cui energia, seguendo Poincaré, denoteremo con η.

La quantizzazione viene usualmente recepita come il fatto che l'energia η di un oscillatore possa assumere solo dei valori discreti, mentre Poincaré ragiona differentemente. Egli cerca di determinare la distribuzione di probabilità dell'energia, o la sua densità $w(\eta)$,[8] e mostrare che questa è "concentrata" sui livelli di energia. In altri termini, i valori di energia diversi da $nh\nu$ hanno probabilità nulla. Se i due approcci possono sembrare identici dal punto di vista logico, il secondo permette a Poincaré di seguire una via Statistico–Meccanica per dimostare la necessità della quantizzazione.

La dimostrazione della necessità della quantizzazione è basata dul fatto che la distribuzione di probabilità $\widetilde{w}(\eta)$ può essere determinata univocamente a partire dalla sua trasformata di Laplace

$$\mathscr{Z}(\beta) = \int e^{-\beta\eta} \mathrm{d}\widetilde{w} = \int e^{-\beta\eta} w(\eta) \mathrm{d}\eta \ , \tag{5.2.1}$$

che in Meccanica Statistica è detta funzione di partizione ed in probabilità funzione generatrice dei momenti. Pertanto, se vale che l'energia media u di un singolo oscillatore si può ottenere mediante l'analoga della usuale relazione della

[5]NOIA PER GLI AUTORI: Ritrovare il libro, che esiste nella biblioteca di fisica a Milano.

[6]Math Ann. ... ..., ristampato anche nelle opere di Poincaré.

[7]Si vedano le relazioni date in F.E. Irons, Am. J. Phys. **69** (9), 879 (2001) e J.J. Prentis, Am. J. Phys. **63**, 339 (1995), che seguono R. McCormmach, *Isis* **58**, 37 (1967).

[8]Ricordiamo che per distribuzione di probabilità di una random variable X si intende la funzione $F(x)$ definita da $F(x) = \{$probabilità che sia $X < x\}$. Invece la funzione $f(x) = F'(x)$ (dove l'apice ' denota derivazione) definisce la corrispondente densità di probabilità.

Meccanica Statistica

$$u = -\frac{\partial}{\partial \beta} \log \mathscr{Z} \,, \tag{5.2.2}$$

allore è evidente che i dati sperimentali di u in funzione della temperatura $1/\beta$ permettono di ricavare (per integrazione) la funzione di partizione e quindi $\widetilde{w}(\eta)$, come antitrasformata di Laplace, che è nota essere unica. Quello che farà Poincaré nel lavoro sarà proprio di dimostrare che vale la relazione (5.2.2), nell'ipotesi che le energie degli oscillatori possano essere considerate variabili causali iid.

D'altra parte, fin dal lavoro di Einstein del 1907 sui calori specifici era noto (in una forma ancora più diretta che nel lavoro originale di Planck), che la formula di Planck per l'energia media si ottiene se si ammette che gli unici valori possibili dell'energia dell'oscillatore sono quelli quantizzati, dati da

$$\eta_n = nh\nu \,, \quad n = 0, 1, 2, \dots \,, \tag{5.2.3}$$

e che essi abbianon probabilità p_n date da

$$p_n(\beta) = \frac{1}{\mathscr{Z}(\beta)} e^{-\beta \eta_n} \,, \qquad \mathscr{Z}(\beta) = \sum_{n=0}^{\infty} e^{-\beta \eta_n} \,. \tag{5.2.4}$$

Infatti in tal caso il valor medio u è ancora dato dalla formule (5.2.2) e (5.2.1), prendendo

$$w(\eta) = \sum_n \delta(\eta - \eta_n) \,, \tag{5.2.5}$$

dove $\delta(x)$ è la funzione delta di Dirac. Dunque, per l'unicità dell'antitrasformata di Laplace, si ottiene che le legge di Planck per l'energia media di un singolo oscillatore implica la necessità della quantizzazione.[9] [10]

In conclusione, se ci chiediamo se la discretizzazione o quantizzazione dell'energia sia necessaria quando si voglia dedurre con procedimenti statistici la legge di Planck, allora seguendo Poincaré otteniamo una risposta positiva.

Osservazione. Abbiamo già fatto presente che Einstein stesso, a partire dal 1911 (anzi, già dal 1908) in maniera decisa espresse l'opinione che la quantizzazione del singolo

[9] Con un semplicissimo calcolo (somma di una serie geometrica di ragione $e^{-\beta h\nu}$) si trova

$$\mathscr{Z}(\beta) = \frac{1}{1 - e^{-\beta h\nu}} \,,$$

e dunque (prendendo la derivata logaritmica)

$$u(\beta) = \frac{h\nu}{e^{\beta h\nu} - 1} \,.$$

[10] In effetti, sia Einstein all'inizio dell'articolo del 1907, sia Poincaré nel suo articolo (si veda pag. 641, a metà), utilizzano quella che poi diventerà la "funzione delta di Dirac" senza introdurre per essa una notazione ma usando, per definirla, le medesime parole che verranno poi usate da Dirac.

oscillatore (e in generale del singolo sottosistema) non è necessaria. Questo lo espresse vividamente nel celebre articolo del 1917, dove dice che la p_n che appare nella (5.2.4), che qui denotiamo con p_k, non è la probabilità che un singolo oscillatore si trovi nel livello k (come direbbe Bohr), ma è invece la corrispondente "frequenza relativa". Nella terminologia che abbiamo usato per descrivere il procedimento di Boltzmann, si ha

$$p_k = n_k/N \ ,$$

dove n_k è il "numero di occupazione" della cella k–esima, rappresentata dal "livello" ϵ_k", e N il numero totale di oscillatori.

5.3 Il lavoro di Poincaré

Veniamo dunque al lavoro di Poincaré. In effetti, su questo argomento della necessità della quantizzazione egli scrisse tre articoli. Il primo costituisce una discussione qualitativa, di tipo divulgativo, del problema. Il secondo lavoro, di una decina di pagine, scritto per fisici, contiene dei calcoli che prendono circa tre quarti del lavoro ma non sono completi. I calcoli dettagliati sono invece pubblicati nel terzo lavoro, di trenta pagine. Egli ne trasse la conclusione che la quantizzazione sia necessaria, e infatti il titolo del paragrafo 8 è proprio *Necessità dell'ipotesi di Planck*. Invece, in un altro lavoro Poincaré mostrò qualche perplessità su tale conclusione,[11] e comunque egli purtroppo morí in quello stesso anno 1912.

Il problema della misura invariante

Poincaré comincia a discutere quale tipo di modello si debba studiare. Egli ricorda come Nernst, l'organizzatore della conferenza di Bruxelles, avesse suggerito di prendere in considerazione modelli in cui si hanno particelle soggette anche a forze proporzionali ad $\dot{\mathbf{a}}$, cioè alla derivata temporale della accelerazione. Questo è un punto molto importante, perché questa è proprio la forma che presenta la forza di reazione di radiazione, ovvero la forza di frenamento (una specie di forza di attrito) che si esercita su ogni carica allorché essa irraggia energia elettromagnetica. Tale forza fu introdotta fenomenologicamente da Planck verso la fine dell'ottocento, poi studiata attorno al 1905 da Abraham e Lorentz, e infine, in maniera probabilmente definitiva, da Dirac nel 1938. Nella seconda parte di queste note vedremo come tale forza produca effetti qualitativi inaspettati, e come potrebbe condurre a situazioni analoghe a quelle trovate nel 1963 da Lorenz con il suo *"strange attractor"*.

Poi Poincaré prende in esame i modelli di interesse per la legge di Planck. Anzitutto si avrebbe il problema di studiare il corpo nero, e dunque l'interazione tra campo elettromagnetico e particelle cariche. Lasciando questo problema a una successivo lavoro (che purtroppo non ebbe il tempo di affrontare), egli infine si concentrò su un modello completamente meccanico, sostanzialmente un sistema

[11] NOTA PER GLI AUTORI. Riportare la frase.

composto di due sottosistemi di oscillatori, rispettivamente di bassa frequenza e di alta frequenza, come illustreremo poco più sotto.

Egli si restringe dunque a un modello meccanico, descrivibile mediante equazioni differenziali, della forma tipica

$$\dot{x} = X(x)$$

con $x \in R^n$, ed X un campo vettoriale in R^n. Per la dinamica questo è assolutamente standard. Il punto delicato è il fatto che ora si intende formulare una meccanica statistica, il che richiede di calcolare delle medie. Come lungamente discusso alla conferenza Solvay, particolarmente da Einstein nel suo primo intervento dopo la sua relazione, e come era ben noto a Poincaré[12], seguendo Boltzmann stesso egli ammette che le misure significative sono quelle definite dalla dinamica, attraverso i tempi di soggiorno.[13] In effetti, misure di tale tipo sono comunemente usate mella teoria ergodica, significativamente nel celebre lavoro del 1937 di Krylov e Bogoliubov.[14] Nel caso del modello di Lorenz del 1963, la misura così definita risulta proprio essere concentrata sullo strano attrattore, e quindi non può essere rappresentata come integrale di una densità rispetto alla misura di Lebesgue nello spazio delle fasi. Così avviene anche per le equazioni differenzialì che descrivono un punto soggetto alla forza di reazione di radiazione.[15] Comunque, Poincaré esclude situazioni di tale tipo, e infatti si limita al caso in cui la misura sia descritta da una densità (nello spazio delle fasi), che egli denota con W.

È noto che una misura μ definita mediante i tempi di soggiorno (rispetto a un "tempo finale" T sufficientemente lungo, eventualmente infinito) risulta essere una "misura invariante", ovvero ha la proprietà che, per ogni tempo t si ha $\mu(\Phi^t A) = \mu(A)$ per ogni insieme A misurabile. Qui $\Phi^t x$ denota l'evoluto, al tempo t, del punto x. È noto anche (si veda il capitolo su Kubo) che, per le misura invarianti che ammettono densità, questa deve soddisfare l'equazione di Liouville stazionaria

$$\sum_i \frac{\partial}{\partial x_i}(WX_i) = 0\,.$$

[12]NOTA PER GLI AUTORI. Citare il lavoro di P. in cui risolve una diatriba tra due noti scienziati dicenso che uno di loro aveva sbagliato perché non aveva calcolato le medie temporali.

[13]Si fissa un pempo T sufficientemente lungo. Fissato poi un dato iniziale x, e la corrispsondente orbita, che denotiamo con $\Phi^t x$, $t \in R$, per ogni regione V si introduce come misura $\mu(V)$ di V la frazione di tempo nel quale l'orbita si trova in V. Naturalmente, questa frazione di tempo (detta tempo di soggiorno) dipende dal punto iniziale x e dal tempo finale T. Ma ci si può attendere che la dipendenza dal tempo finale T genericamente scompaia. In molti casi scompare anche la dipendenza dal "dato iniziale" (è questo il cosiddetto "problema ergodico").

[14]Annals of Mathematics, 1937.

[15]Un caso é già stato messo in evidenza. Si tratta della collisione di una particella classica con una barriiera di potenziale, in cui si incontra un analogo classico dell'effetto tunnel. Si veda il lavoro Delzann, Carati, Galgani, Sassarini, il cui risultato é riassunto in un lavoro dedicato al problema di Bell.

Una densità che soddisfi tale equazione viene chiamata da Poincaré con il nome classico di *"ultimo moltipicatore"*. Evidentemente i sistemi hamiltoniani, avendo la proprietà div $X = 0$, ovvero $\sum_i \partial_i X_i = 0$, ammettono sempre un ultimo moltiplicatore, proprio $W = 1$. In altri termini, per i sistemi hamiltoniani la misura di Lebesgue nello spazio delle fasi è sempre invariante. Questa in effetti è proprio la misura a priori nello spazio delle fasi totale (spazio Γ) considerata fino dai tempi di Boltzmann. Tuttavia ben sappiamo che la misura di Lebesgue conduce alla statistica di Boltzmann e dunque all'equipartizione dell'energia, che ammettiamo essere adeguata per la statistica degli atomi puntiformi, ma inadeguata per gli oscillatori di alta frequenza. Quindi Poincaré si chiede se nei sistemi di interesse per la legge di Planck (oscillatori armonici) esista un ultimo moltiplicatore, diverso da quallo banale di Lebesgue, che possa eventualmente condurre alla statistica di Planck.

Il modello

Poincaré considera un sistema composto da due sottosistemi, uno costituito da p oscillatori tutti uguali, di bassa frequenza (o lungo periodo, come egli dice) , e l'altro costituito da n oscillatori tutti uguali di grande frequenza (o di piccolo periodo, o *oscillatori hertziani*, come egli anche li chiama). L'interazione tra i due sottosistemi avviene attraverso collisioni a due corpi tra componenti dei due sottosistemi. Gli oscillatori di bassa frequenza (sostanzialmente aventi frequenza nulla) possono pensarsi rappresentare un sistema di atomi puntiformi (perché l'energia di un oscillatore armonico diventa quella di una particella libera sulla retta quando si ponga nulla la frequenza dell'oscillatore), per i quali possiamo ammettere che valga il principio di equipartizione. I due sottosistemi verranno chiamati rispettivamente il sistema di atomi e il sistema di oscillatori (che Poincaré chiama anche sistema dei *risonatori*, usando la classica terminologia di Planck).

Le quantità di interesse ai fini di ottenere una termodinamica in termini microscopici (ovvero formulare una termodinamica statistica) sono allora l'energia media del sistema di atomi e quella del sistema di oscillatori che, seguendo Poincaré, denoteremo rspettivamente con pX e nY, mettendo in rilievo le corrispondenti energie *specifiche* X ed Y. Osserviamo che spesso quando si parla di quantità specifiche ci si riferisce ad una mole. Qui invece Poincaré si riferisce a quantità che potremmo chiamare (contrariamente alle convenzioni sin qui adottate per denotare le quantità specifiche) quantità specifiche *per atomo, o per oscillatore*, ovvero quantità relative ad un sistema (di atomi o di oscillatori), ma divisa per il numero di costituenti del sistema (numero p di atomi o numero n di oscillatori).

Nota: un possibile equivoco. Si potrebbe essere indotti a ritenere che Y definisca l'energia media di un singolo oscillatore, mentre essa definisce l'energia totale (del sistema di oscillatori), che per pura opportunità di notazione abbiamo diviso per il numero n di oscillatori. Essa coincide con l'energia media di un singolo oscillatore solo nel caso in

cui le energie dei singoli oscillatori siano variabili casuali indipendenti ed identicamente distribuite (iid). Questo è un punto rilevante sul quale ritorneremo.[16]

Ora, ben naturalmente Poincaré ammette che il valor medio X dell'energia specifica per atomo sia proporzionale alla temperatura $T \equiv X$. Anzi, Poincaré definisce la temperatura T in termini meccanici proprio mediante $X \equiv T$, ovvero prendendo unità di misura in cui si ha $k_B = 1$. Il problema che interessa è allora quale sia, come dice Poincaré, *la partizione dell'energia*, cioè quale sia, in funzione della temperatura (ovvero in funzione di X) l'energia media *specifica* per oscillatore Y.

In conclusione, il problema è di determinare la funzione $Y = Y(T)$, ovvero, in termini meccanico–statistici, la funzione $Y = Y(X)$.

Un problema preliminare di tipo generale: necessità di considerare il caso asintotico di un grandissimo numero di costituenti

Tuttavia, prima di studiare il problema della partizione dell'energia, Poincaré dedica una notevole parte del lavoro (le prime 15 pagine) a un problema prelminare più generale, che in effetti egli studia con tale intensità da dare l'impressione che questo sia il vero problema che gli sta a cuore (come dice egli stesso; vedi sotto). Si tratta del problema di come sia possibile ottenere una descrizione statistica che sia in accordo con una situazione fisica di equilibrio termico, facendo uso di misure invarianti diverse da quella di Lebesgue (ovvero usando densitá $W \neq 1$). Egli formalizza questa condizione richiedendo che la relazione tra le due energie specifiche medie (per atomo e per risonatore) X ed Y sia indipendente dal rapporto n/p (infatti tale relazione dovrebbe dipendere solo dalla temperatura, oltre che dal parametro ν, che rappresenta la frequenza comune degli oscillatori). Anticipiamo qui la conclusione del suo lungo lavoro su questo problema, che egli ottiene alla tredicesima pagina, dove sostanzialmente conclude che l'equilibrio termico può aversi solo nella situazione limite in cui i numeri n e p dei costituenti dei due sistemi siano molto grandi.

Infatti, dopo avere esaminato diversi casi particolari, egli così si esprime: *"È dunque in casi molto eccezionali che la legge di partizione dell'energia risulta indipendente dagli interi n, p. Sembra allora seguirne che non sia possibile alcun equilibrio termico, in contrasto con il secondo principio della termodinamica. Ma bisogna ricordarsi che i numeri n, p sono sempre molto grandi.*

Conviene dunque porre il problema in modo diverso: è la legge di partizione dell'energia indipendente dal rapporto n/p, quando gli interi sono molto grandi? Se questa indipendenza non avesse luogo, l'equilibrio termico sarebbe impossibile. ...Finché non venga stabilita tale indipendenza, possono restare dei dubbi sui ragionamenti di Planck, che si basano sull'esistenza dell'equilibrio e sul teorema di Boltzmann."

[16]Effettivamente, a pag 636 in basso. Poincaré stesso dice: *"Nel caso che gli oscillatori siano un sistema isolato, l'energia media Y è* ***l'energia di ogni singolo oscillatore"****.*

E conclude con la frase il cui senso avevamo anticipato sopra: *"Questo sarebbe sufficiente per giustificare il presente lavoro."*

Formalizzazione del problema

Seguendo Poincaré, denotiamo con $\eta_1, \dots \eta_n$ le energie degli oscillatori e con $\xi_1, \dots \xi_p$ le energie degli atomi. Ricerchiamo che forma deve avere la misura invariante nello spazio delle fasi totale. Assumiamo anche, con Poincaré, che la misura ammetta densità W (un ultimo moltiplicatore), e allora il problema sarà di stabilire se questa funzione è continua o "fatta a salti" [17]. Come coordinate nello spazio delle fasi si possono prendere le energie e le fasi di ogni atomo (pensato come un oscillatore di bassa frequenza), e quelle di ogni oscillatore (di grande frequenza). A questo punto Poincaré, occupandosi di misure di equilibrio (ovvero invarianti), ritiene ovvio anzitutto che la densità W non dipenda dalle fasi (che sono le variabili "veloci" del sistema), e anzi neppure dalle energie degli atomi (come si assume solitamente quando si discute il sistema dei soli atomi, e si sceglie la misura di Lebesgue, ovvero si prende $W = 1$ come ultimo moltiplicatore). Ammettiamo dunque che si abbia punt

$$W = W(\eta_1, \dots \eta_n) \, .$$

Poi Poincaré compie un altro passo, che svolgerà un ruolo cruciale nella deduzione della necessità della quantizzazione. Egli infatti ammette che le energie η_i dei singoli oscillatori siano variabili casuali (random variables) di tipo iid, ovvero indipendenti ed identicamente distribuite. Dunque ammette che la funzione W abbia la forma

$$W(\eta_1, \dots \eta_n) = w(\eta_1) \dots w(\eta_n) \, , \tag{5.3.1}$$

cioè una forma fattorizzata (energie indipendenti), e con fattori definiti da funzioni (di una sola variabile) tutte uguali (energie identicamente distribuite).

A dire il vero, Poincaré non dice esplicitamente che sta introducendo una ipotesi, e dà invece una argomentazione che tiene quasi una intera pagina (pag. 633), e che noi non riusciamo a seguire completamente. Provvisoriamente diciamo che egli introduce questa ulteriore proprietà come una ipotesi, e su questo punto ritorneremo più avanti.

Le quantità che interessano sono dunque i valori medi X, Y, rispettivamente della energia specifica per atomo, e dell'energia specifica per oscillatore. Più precisamente Poincaré vuole calcolare tali valori medi in un insieme di tipo microcanonico, come diremmo oggi, o più precisamente quando si sappia che il sistema totale ha una energia compresa nello "strato" $E_{tot}, E_{tot} + \delta E_{tot}$, che denoteremo con Δ_{Etot}.

I valori medi di interesse X, Y sono dunque definiti dalle formule

$$M \cdot X = \frac{1}{p} \int_{\Delta_{Etot}} \sum \xi_i \; W \; \mathrm{d}^p \xi \, \mathrm{d}^n \eta$$

[17] Poincaré parla, nell'ultima pagina del suo lavoro, di *"équations a auts brusques"*.

$$M \cdot Y = \frac{1}{n} \int_{\Delta_{Etot}} \sum \eta_j \; W \; \mathrm{d}^p \xi \, \mathrm{d}^n \eta$$

dove

$$M = \int_{\Delta_{Etot}} W \mathrm{d}^p \xi \, \mathrm{d}^n \eta$$

e si è denotato

$$\mathrm{d}^n \eta = \mathrm{d}\eta_1 \ldots \mathrm{d}\eta_n \; , \qquad \mathrm{d}^p \xi = \mathrm{d}\xi_1 \ldots \mathrm{d}\xi_p \; .$$

Osservando che

$$\int_{\sum \xi_i = x} \mathrm{d}\xi_1 \ldots \mathrm{d}\xi_p = \frac{x^{p-1}}{(p-1)!} \; ,$$

è facile vedere che questi tre integrali possono essere espressi in termini della funzione

$$\widetilde{W}(E) = \int_{\sum \eta_j < E} W(\eta_1, \ldots \eta_n) \mathrm{d}^n \eta \; . \tag{5.3.2}$$

Evidentemente, per E fissato questa definisce la probabilità che l'energia totale $\sum_j \eta_j$ degli oscillatori sia minore di E. Dunque, come funzione di E essa rappresenta, nel senso consueto della teoria delle probabilità, la *distribuzione di probabilità* della somma delle energie degli oscillatori, la quale può assumere tutti i valori nell'intervallo $(0, E_{tot})$. Ovviamente, per ogni data energia E del sistema degli oscillatori, la parte restante $E_{tot} - E$ è l'energia degli atomi.

Si verifica facilmente che, in termini della $\widetilde{W}(E)$, gli integrali assumono la forma

$$M = \frac{1}{(p-1)!} \int_0^{E_{tot}} (E_{tot} - E)^{p-1} \; \widetilde{W}(E) \, \mathrm{d}E$$

$$M \cdot X = \frac{1}{p!} \int_0^{E_{tot}} (E_{tot} - E)^p \; \widetilde{W}(E) \mathrm{d}E$$

$$M \cdot Y = \frac{1}{n(p-1)!} \int_0^{E_{tot}} E (E_{tot} - E)^{p-1} \; \widetilde{W}(E) \, \mathrm{d}E \; .$$

Applicazione dei metodi delle grandi deviazioni, e il primo metodo di Poincaré

Dobbiamo dunque valutare gli integrali scritti sopra, nel limite di grande n (numero di oscillatori) per un fissato valore del rapporto p/n (che Poincaré denota con k). A tal fine, Poincaré vuole applicare il classico metodo asintotico di Laplace, il quale richiede un opportuno comportamento asintotico (rispetto ad n) della funzione $\widetilde{W}(E)$, la distribuzione di rpobabillità di $\sum \eta_j$, somma delle energie degli n oscillatori. È questo il problema studiato nella teoria delle grandi

deviazioni, i cui risultati furono stabiliti[18] da Cramér nel 1938 nel caso iid, e poi nella sua forma generale da Gartner nel 1977.

Questo problema viene affrontato da Poincaré con quelli che egli chiamò primo e secondo metodo. Nel secondo, come vedremo, egli anticipa il risultato di Cramér. Il primo metodo, invece, consiste semplicemente nel proporre in forma di *Ansatz*, sulla base dello studio di alcuni casi particolari, una forma asintotica analoga a quella che sará in seguito dimostrata da Gartner.

Ricordiamo che Gartner sostanzialmente richiede soltanto che esista la trasformata di Laplace di $\widetilde{W}(E)$ e che il suo logaritmo cresca linearmente con n, cioè si abbia

$$\log \int_0^\infty e^{-\beta E}\, \widetilde{W}(E)\, \mathrm{d}E \simeq n \ ,$$

condizione ovviamente soddisfatta nel caso particolare iid, che è quello che fu considerato da Cramér, e in effetti anche da Poincaré.

In queste ipotesi si dimostra che esiste una funzione $s = s(E/n)$, convessa, tale che si abbia

$$\widetilde{W}(E) \simeq \exp\Big(ns(E/n) + O(\sqrt{n})\Big)\,, \qquad n \gg 1 \ . \tag{5.3.3}$$

Questa formula, a meno del termine di errore $O(\sqrt{n})$, è quella che viene proposta da Poincaré, con diverse notazioni, nel paragrafo 5 del lavoro.

Nel paragrafo 6 egli calcola anche la forma che prende la funzione $\widetilde{W}(E)$ se vale l'ipotesi di quantizzazione di Planck, mostrando in particolare che allora vale la relazione (5.3.3). Tra l'altro, nel compiere questo calcolo, egli in particolare riottiene la formula combinatoria che Planck aveva utilizzato nel suo lavoro del 14 dicembre 1900. Inoltre, egli manipola la "funzione" delta esattamente con il procedimento e anche le parole che verranno usati in seguito da Dirac (come sostanzialmente aveva già fatto Einstein all'inizio del suo lavoro del 1907 sui calori specifici).

Utilizzando la formula di Gartner (5.3.3), gli integrali prendono una forma tipica che, ad esempio, nel caso di M è data da

$$M = e^{p-1} \int_0^{E_{tot}} \mathrm{d}E \exp\left[n\left(s(\frac{E}{n}) + \frac{p-1}{n} \log \frac{E - E_{tot}}{p-1} + O(1/\sqrt{n}) \right) \right] \tag{5.3.4}$$

Quindi, come si vede, nell'argomento dell'esponenziale si ha una funzione che dipende dalle quantità specifiche E/n, p/n, moltiplicata per il fattore n, che si deve pensare molto grande. Pertanto è chiaro che il valore dell'integrale è determinato dal valore di E (che denoteremo con U) in cui questa funzione assume valore massimo, in cui quindi la derivata si annulla. È questa l'idea base del metodo di Laplace per la stima asintotica degli integrali. Situazioni analoghe si presentano nei due integrali che definiscono MX ed MY, con la sola differenza

[18] H. Cramér, Colloque Consacré à la thèorie des probabilités, **3**, 2 (1938), J. Gärtner, Th. Prob. Appl. **22**, 24 (1977). Si veda anche Ellis, libro.

che vi compaiono due ulteriori fattori. In tal modo, mettendo in rilievo tali fattori, è evidente che possiamo scrivere, seguendo Poincaré,

$$M \cdot X = \left(\frac{E_{tot} - U}{p} + O(1/\sqrt{n}) \right) \frac{1}{(p-1)!} \int_0^{E_{tot}} (E_{tot} - E)^{p-1} \, \widetilde{W}(E) \mathrm{d}E$$
$$M \cdot Y = \left(\frac{U}{n} + O(1/\sqrt{n}) \right) \frac{1}{(p-1)!} \int_0^{E_{tot}} (E_{tot} - E)^{p-1} \, \widetilde{W}(E) \mathrm{d}E \ .$$

Pertanto, trascurando l'errore (piccolissimo nel limite di n molto grande) e semplificando per M in entrambi i membri, si ottiene

$$X = \frac{E_{tot} - U}{p} \qquad Y = \frac{U}{n} \ .$$

D'altra parte, come abbiamo detto, il valore di U è quello per cui la derivata dell'argomento dell'esponenziale che compare nell'integrale (5.3.4) per M si annulla, cioè il valore per cui si ha

$$s'(\frac{U}{n}) = \frac{p}{E_{tot} - U}$$

(dove l'apice denota derivazione rispetto all'argomento e abbiamo sostituito $p-1$ con p), ovvero

$$s'(Y) = \frac{1}{X} \ .$$

Questo risolve il problema di Poincaré in quanto mostra che il rapporto Y/X non dipende da p e da n (e quindi non dipende da $k \equiv p/n$), nel limite in cui valga la stima, in cui cioè si possano trascurare gli errori, che per le quantità specifiche sono di ordine $1/\sqrt{n}$. Abbiamo quindi ottenuto che è possibile avere una termodinamica statistica anche nel caso in cui la misura di partenza non sia quella di Lebesgue.

Da queste formule si ottiene anche immediatamente la formula per l'entropia del sistema di oscillatori, poiché risulta subito che la *funzione di grande deviazione* $s(u)$ (dove $u = U/n \equiv Y$) è l'entropia specifica di tale sistema. Infatti, avendo identificato il valor medio X dell'energia specifica del gas di atomi con la temperatura termodinamica assoluta T, e denotando $S = ns$, si ottiene

$$\frac{1}{T} \equiv \frac{1}{X} = \frac{\mathrm{d}s}{\mathrm{d}u} = \frac{\mathrm{d}S}{\mathrm{d}U} \ ,$$

cioè che a volume costante vale $\mathrm{d}U = T\mathrm{d}S$, e questo identifica $S = ns$ come l'entropia termodinamica del sistema di oscillatori.

D'altra parte, nel teorema di grande deviazione di Gartner si ha anche un ulteriore risultato, ovvero che la funzione di grande deviazione $s(u)$ è legata alla

$\widetilde{W}(E)$ dalla relazione seguente: se $f(\beta)$ è l'energia libera specifica legata alla $s(u)$ dalla consueta relazione termodinamica

$$\beta f(\beta) = \inf_{u>0} (\beta u - s(u)) \ ,$$

allora si ha

$$-\beta f(\beta) = \frac{1}{n} \log \int_0^\infty e^{-\beta E} \ \widetilde{W}(E) \mathrm{d}E \ . \tag{5.3.5}$$

ovvero, in termini dell'energia libera $F = nf$, si ha

$$-\beta F(\beta) = \log \int_0^\infty e^{-\beta E} \ \widetilde{W}(E) \mathrm{d}E \ . \tag{5.3.6}$$

Equivalentemente, ricordando la relazione termodinamica $F = U - TS$ tra energia interna U ed energia libera F, si ha per l'energia media U del sistema di oscillatori l'espressione meccanico–statistica

$$U(\beta) = -\partial_\beta \log \int_0^\infty e^{-\beta E} \ \widetilde{W}(E) \mathrm{d}E \ , \tag{5.3.7}$$

Si è ottenuta dunque la meccanica statistica del sistema degli oscillatori in interazione con un gas di atomi, definita mediante formule che generalizzano quella familiare del formalismo canonico, a cui essa si riduce nel caso particolare $W = 1$.

Ma il punto rilevante ai fini del problema della quantizzazione è che la formula (5.3.7) rende possibile risalire dai dati sperimentali alla distribuzione di energia $\widetilde{W}(E)$. Infatti i dati sperimentali forniscono l'energia interna U del sistema di oscillatori, o equivalentemente la loro energia libera F, in funzione di β. Allora la (5.3.6) fornisce la trasformata di Laplace di $\widetilde{W}(E)$, e infine, per l'unicità dell'antitrasfrormata, anche la $\widetilde{W}(E)$ stessa. Questa è proprio la situazione che avevamo considerato all'inizio del capitolo, con la sola rilevante differenza qui ci si riferisce al sistema totale degli oscillatori, e non ad un singolo oscillatore.

Ora, il procedimento per la determinazione di $S(U)$ a partire da $\widetilde{W}(E)$, non era disponibile a Poincaré, perché sarà scoperto da Cramer (almeno nel caso iid) solamente nel 1938, e dunque egli non era in grado di provare la relazione (5.3.7) in tutta generalità. Quindi, mentra era in grado di mostrare che è possibile una termodinamica, non era ancora in grado di discutere la necessità della quantizzazione. A questo scopo egli introdusse il suo secondo metodo (paragrafo 7 e seguenti).

Il secondo metodo di Poincaré: necessità della quantizzazione

Nel secondo metodo, Poincaré, invece di cercare di dare direttamente un'espressione generale per $\widetilde{W}(E)$, cerca di riscrivere gli integrali che definiscono M, MX e MY, in una forma che gli permetta comunque di applicare il metodo di Laplace,

ovvero in una forma in cui compaia l'esponenziale di una funzione moltiplicata per n. Egli comincia notando che precisamente una forma di tale tipo viene assunta dalla trasformata di Laplace di $\widetilde{W}(E)$, se si fa uso dell'ipotesi iid per le energie degli oscillatori. Infatti, indicando con $Z(\beta)$ la trasformata di Laplace di $\widetilde{W}(E)$, si ha

$$Z(\beta) = \int e^{-\beta E}\widetilde{W}(E)\mathrm{d}E = \int e^{-\beta\sum\eta_i} w(\eta_1)\dots w(\eta_n)\mathrm{d}^n\eta\,,$$

ovvero

$$Z(\beta) = \mathscr{Z}^n(\beta) = \exp(n\chi(\beta))$$

dove, come nel secondo paragrafo, con $\mathscr{Z}(\beta)$ abbiamo indicato la trasformata di Laplace della distribuzione di probabilità $w(\eta)$ del singolo oscillatore, e abbiamo introdotto la notazione

$$\chi = \log\mathscr{Z} \quad \text{ovvero} \quad \mathscr{Z} = e^{\chi}\ .$$

È questo il punto dove tecnicamente entra in modo fondamentale l'ipotesi iid, per cui tutti i risultati dei calcoli vengono espressi in termini della funzione $\mathscr{Z}$, cioè in termini della funzione di partizione del singolo oscillatore. Questo non è possibile nei casi non iid, in cui le quantità relative ai singoli oscillatori non sono accessibili macroscopicamente.

Usando il teorema di inversione della trasformata di Laplace, la funzione $\widetilde{W}(E)$ si esprime dunque come integrale al modo sequente

$$\widetilde{W}(E) = \int_\gamma \mathrm{d}\beta' \exp(\beta' E + n\chi(\beta'))\,,$$

dove γ è un opportuno cammino nel piano β' complesso (in genere si prende una retta parallela all'asse immaginario con parte reale sufficientemente grande). Si ottiene allora il risultato cercato perché, ad esempio ancora per M, si trova l'espressione

$$M = e^{p-1}\int_0^{E_{tot}} \mathrm{d}E \int \mathrm{d}\beta' \exp\left[n\left(\beta'\frac{E}{n} + \chi(\beta') + \frac{p-1}{n}\log\frac{E-E_{tot}}{p-1})\right)\right]\,,$$

dove la costante e^{p-1} si ottiene introducendo l'espressione di Stirling per il fattoriale $(p-1)!$ nell'integrale definente M.

A questo punto, come nel caso del primo metodo, abbiamo di nuovo che nell'argomento dell'esponenziale si ha una funzione che dipende dalle quantitá specifiche E/n, p/n, β, moltiplicata per il fattore n, che si deve pensare molto grande. Pertanto possiamo ancora applicare il metodo asintotico di Laplace, ed il valore dell'integrale è determimato dai valori di E e di β' (che denoteremo con U e con β) in cui questa funzione assume valore massimo. Una piccola precisazione matematica: possiamo sempre deformare il cammino γ, senza che

cambi l'integrale che definisce l'antitrasformata, in modo tale che il punto di massimo (U, β) appartenga al cammino di integrazione. Allora, a meno di errori di ordine $O(1/\sqrt{n})$, come nel caso del primo metodo si otterrà

$$X = \frac{E_{tot} - U}{p} \qquad Y = \frac{U}{n} \, .$$

Notiamo ora che, a differenza del caso precedente, la funzione che compare sotto il segno di integrale dipende da due variabili, per cui il punto di massimo sarà determinato imponendo che si annulli il *gradiente* della funzione, cioè che siano verificate le due relazioni

$$\chi'(\beta) + \frac{U}{n} = 0 \quad , \quad \beta - \frac{p-1}{E_{tot} - U} = 0 \, ,$$

dove abbiamo indicato con χ' la derivata di χ rispetto al suo argomento. In termini di X e di Y troviamo allora le relazioni seguenti

$$X = \frac{1}{\beta} \quad , \quad Y = -\chi'(\beta) \, .$$

Ricordando la definizione di χ, la seconda può essere messa nella forma

$$Y = -\frac{\partial}{\partial \beta} \log \mathscr{Z} \, ,$$

che dimostra appunto la validita della formula (5.2.2) che avevamo ammesso al paragrafo 5.2. Si può dunque ripetere il ragionamento là esposto: i dati sperimentali definiscono univocamente $Y(\beta)$, e quindi univocamente $\mathscr{Z}$. Ma poi, per l'unicità dell'antitrasformata di Laplace, resta definita univocamente anche la distribuzione di probabilitá $\widetilde{w}(\eta)$. Dunque, resta dimostrato che l'unico modo per riprodurre la legge empirica di Planck, è mediante l'introduzione di una distribuzione $\widetilde{w}(\eta))$ concentrata sui livelli discreti. In realtà Poincaré fa di più: sfruttando le proprietà della trasformata di Fourier, egli riesce a mostrare che la funzione $\widetilde{w}(\eta)$ non può essere continua, ma deve presentare delle discontinuità corrispondenti a dei livelli, nella sola ipotesi che la densità di energia del corpo nero (integrata su tutte le frequenze) risulti finita, senza bisogno di richiedere che sia strettamente quella di Planck.[19]

In conclusione, rimane provato che si ha una termodinamica statistica, sotto ipotesi generalissime per la funzione $w(\eta)$. Queste termodinamiche statistiche generalizzano quella consueta, che potremmo chiamare *gibbsiana*, che viene riottenuta nel caso particolare $w(\eta) = 1$. [20] Infine, essendosi ridotto al caso in un singolo oscillatore (qui discusso nella prima sezione), ne conclude che, nell'ipotesi iid, la legge di Planck comporta la necessitá della quantizzazione.

[19] Come fu fatto osservare da Lorentz, quest'ultima osservazione (relativa però al caso di un singolo oscillatore) era stata anticipata da Ehrenfest in un suo lavoro del 1911.

[20] Prima aveva detto *non sappiamo ancora se la legge di ripartizione dell'energia sia indipendene dal rapporto $k = p/n$ comunque si prenda la funzione w* (la densità di probabilità della singola energia).

5.4 Osservazioni sulla necessità della quantizzazione

Mostriamo ora come possano ritenersi fondati i dubbi che Poincaré stesso aveva espresso sulla necessità della quantizzazione. Il punto è che tale necessità è stata dimostrata da Poincaré nell'ipotesi che le energie dei singoli oscillatori siano variabili casuali indipendenti ed identicamente distribuite (iid). Della plausibilità di tale ipotesi discuteremo più sotto.

Energia totale vs energia del singolo oscillatore

Ora vogliamo invece discutere che cosa possiamo dire a proposito di tale necessità nel caso generale, in cui si lasci cadere l'ipotesi iid, e si assuma invece che valga il principio di grandi deviazioni al modo di Gartner, il quale assicura comunque, come abbiamo visto, l'esistenza di una termodinamica statistica (indipendenza dal rapporto p/n).

Assumiamo dunque che l'energia media U del sistema di n oscillatori in interazione con un gas di atomi segua la legge di Planck. Dal teorema di grandi deviazioni nella forma generale (5.3.3) di Gartner si dedurrebbe ancora la quantizzazione del sistema totale se non esistesse il termine di correzione dell'ordine $\sqrt{n}$. Ma nel teorema quel termine di correzione esiste, e poiché la quantizzazione coinvolge una quantità finita (l'azione h di Planck, indipendente da n), tale quantizzazione non risulta efficace nel caso macroscopico di n dell'ordine del numero di Avogadro. Infatti la discretizzazione riguarda "salti" minuscoli rispetto all'incertezza entro cui è garantito il risultato. Dunque nel caso generale non sembra si deduca che la distribuzione di energia del sistema di oscillatori sia quantizzata. Si potrebbe soltanto concludere che è quantizzata la distribuzione di energia di un oscillatore che potremmo chiamare *virtuale*, di cui nessuno immaginerebbe di parlare come se fosse qualcosa di reale. Un oscillatore virtuale la cui energia ad ogni istante sia definita come l'energia specifica per oscillatore U/n, ovvero come l'energia totale del sistema macroscopico di oscillatori, divisa per il numero n di oscillatori.

Potrebbe tuttavia darsi che per qualche motivo ci trovassimo in un caso eccezionale in cui sia nulla l'incertezza prevista in generale dal teorema di grandi deviazioni, e che in tal caso si possa dedurre la necessità per il singolo oscillatore. Ad esempio, è noto che la somma di n variabili *iid* è gaussiana, asintoticamente per grandi n. Ora, questo evidentemente non implica che siano distribuite gaussianamente le singole variabili casuali. Tuttavia, esiste un profondo teorema in virtù del quale si può concludere che le singole variabili sono gaussianamente distribuite se si ammette che la loro somma sia *esattamente* distribuita in maniera gaussiana per ogni n. Perché queste considerazioni siano significative per il nostro caso occorrerebbe dunque che il teorema di grandi deviazioni potesse dimostrarsi senza il termine di incertezza, e inoltre che valesse un teorema del tipo di quello appena citato.

L'esempio di Einstein. A proposito del passaggio dall'energia totale U del sistema di oscillatori all'energia u di un singolo oscillatore, è significativo l'esempio che Einstein riporta almeno due volte nei suoi lavori.[21] Consideriamo il diamante a 73 K, ponendo con Nernst la frequenza ν di oscillazione degli atomi di carbonio pari a $27.3 \cdot 10^{12}$ Hz. Allora, se valutiamo l'energia media u di un singolo oscillatore secondo la formula di Planck (praticamente coincidente con quella di Wien per quei valori di ν e T) si ha

$$\frac{u}{h\nu} = e^{-h\nu/k_B T} \simeq e^{-18.6} \simeq 10^{-8} .$$

A questo proposito Einstein commenta: "*L'energia media dell'oscillatore diventa così una frazione piccolissima, circa* 10^{-8}*, del quanto di energia* $h\nu$*. A ogni istante deve dunque oscillare un solo atomo su* 10^8*, mentre gli altri sono in quiete assoluta. Pur convinti dell'inadeguatezza della nostra meccanica per fenomeni siffatti, un risultato di questo genere appare estremamente singolare.*" Da qualche altra parte lo definisce invece "sconcertante".

Veniamo ora alle incertezze, ovvero alle deviazioni standard (stimate con le formule di Einstein per le fluttuazioni di energia, illustrate nel capitolo a lui dedicato). Se consideriamo una mole di oscillatori, si avrà una energia totale E con valor medio U dato da

$$U \simeq 10^{15} h\nu ,$$

mentre l'incertezza (*standard deviation*) σ_E sarà data da

$$\sigma_E \simeq \sqrt{h\nu U} \simeq 10^7 h\nu \gg h\nu ,$$

cioè l'incertezza è enormemente più grande del salto quantico. È proprio questa la situazione che avevamo incontrato sopra, in relazione all'errore previsto nel teorema di grandi deviazioni.

Invece, se si considera l'oscillatore singolo (virtuale), si avrà una incertezza

$$\sqrt{h\nu u} \simeq 10^{-4} h\nu ,$$

molto più piccola del salto quantico.

Sulla plausibilità dell'ipotesi iid per le energie dei singoli oscillatori

Abbiamo visto come l'ipotesi iid svolga una ruolo centrale per dedurre la necessitá della quantizzazione, e abbiamo ricordato come Poincaré la discute lungo una intera pagina del suo lavoro, e come anche Lorentz avesse interpretato tale discussione come una dimostrazione della proprietà iid. In effetti l'argomentazione di Poincaré non risulta del tutto chiara (almeno a noi) e, anche se con estrema titubanza, ci permettiamo di proporre che tale argomentazione non sia una dimostrazione.

In effetti, sembra che l'unica proprietà degli oscillatori invocata da Poincaré nella sua argomentazione sia il fatto che essi sono supposti essere identici. La situazione sarebbe dunque simile a quella che si ha nel caso di n monete; se esse sono state prodotte tutte con lo stesso materiale e in maniera uniforme, sembrerebbe che si dovesse attribuire a ciascuna di esse la medesima probabilità a priori

[21] Si veda la relazione alla conferenza Solvay 1911, lavoro 10, pag. 246 dell'edizione italiana delle opere, e anche il lavoro 6 (1906), pag. 185.

di condurre, in un lancio, al risultato testa oppure croce. Qui tuttavia le cose sono diverse, a motivo del ruolo della dinamica. Si ricordi che Poincaré aveva convenuto, con Einstein e seguendo Boltzmann, che la probabilità di trovarsi in un insieme dello spazio delle fasi debba essere definita come frazione del tempo di soggiorno, per tempi sufficientemente lunghi. Ora, i risultati della dinamica mostrano che, anche nel caso di oscillatori identici, se la frequenza degli oscillatori è abbastanza grande, allora la densità $w(\eta)$ dell'energia di ogni oscillatore si stabilizza presto su una funzione che dipende fortissimamente dal valore iniziale dell'energia dell'oscillatore considerato, essendo concentrata attorno al valore iniziale. Solo nel caso di frequenze basse quelle funzioni sono tutte uguali, e in accordo con la legge di Maxwell–Boltzmann.

Questo punto, in effetti, era proprio il cavallo di battaglia di Jeans, che lo aveva ampiamente uillustrato alla conferenza Solvay, cui aveva partecipato su indicazione di Rayleigh.[22] Tuttavia, dopo la pubblicazione dell'articolo di Poincaré, probabilmente a causa della sua autorità, egli abbandonò completamente la sua tesi, giungendo fino al punto, come abbiamo ricordato piú sopra, di fare una vera e propria ritrattazione pubblica.

La proprietà sopra citata della distribuzione di probabilità dell'energia di un oscillatore di alta frequenza, ovvero che essa ha una forma strettamentre dipendente dall'energia iniziale, è stata messa in luce in diversi lavori, ad esempio in un lavoro del 2003.[23] Proprio come nel caso di Poincaré, si considerano degli oscillatori di fissata frequenza che collidono singolarmente con atomi (punti materiali), i cui dati iniziali vengono estratti da un campione a una certa temperatura T. Le collisioni si producono attraverso un fissato potenziale. Nella figura 1 del lavoro citato sono riportate le curve che danno in funzione del tempo i valori assunti da un oscillatore: a sinistra un oscillatore di alta frequenza ($\omega = 15$, nelle unità scelte) e a destra uno di bassa frequenza ($\omega = 3$). In effetti, la figura di sinistra è in qualche modo fuorviante, perché la scala delle ordinate è molto amplificata (di un fattore mille) rispetto a quella della figura di destra. Sulla stessa scala di quella di destra, l'energia dell'oscillatore di alta frequenza apparirebbe essere costante, indipendente dal tempo. Con l'ingrandimento usato nella figura, si vede che essa in effetti cambia, ma seguendo un moto a salti, della forma che è caratteristica dei cosiddetti *voli di Lévy*. Nella figura successiva viene poi mostrato che le proprietà statistiche sono completamente diverse nei due casi ("code lunghe nel caso delle alte frequenze").

In conclusione, noi diremmo che le energie degli oscillatori di alta frequenza, a differenza di quelle degli oscillatori di bassa frequenza, anche se possono essere considerate variabili casuali indipendenti, tuttavia non sono identicamente distribuite. Questo fatto potrebbe invalidare la conclusione tratta da Poincaré sulla necessità della quantizzazione. Tale necessità è infatti dimostrata solo per un oscillatore virtuale la cui energia sia definita come l'energia totale del sistema di oscillatori divisa per il numero n di oscillatori. Parafrasando quello che Einstein disse nel suo contributo alla conferenza Solvay del 1911 e continuò a ripetere lungo tutta la sua vita, potremmo dire che se, per semplicità di descrizione, trat-

[22] Si veda la sua lettera riportata all'inizio degli atti della conferenza.

[23] A. Carati, L. Galgani, B. Pozzi, *Phys. Rev. Lett.* **90**, 010601 (2003).

tiamo tutti gli oscillatori come se fossero iid, allora "tutto va come se" l'energia degli oscillatori fosse quantizzata.[24]

Da ultimo, facciamo presente come il problema della necessità della quantizzazione prenderebbe un aspetto del tutto nuovo se si considerassero modelli che includono gli effetti del campo elettromagnetico, come discuteremo nella seconda parte di queste note. In tal caso si dovrebbe tenere conto della forza di reazione di radiazione, che comporta effetti qualitativamente nuovi. In particolare potrebbero presentarsi fenomeni del tipo di quelli che si incontrano nel caso dello strano attrattore di Lorenz, scoperto nel 1963, in cui si ha a che fare con una misura invariante non assolutamente continua rispetto alla misura di Lebesgue.

5.5 Primi tentativi di sostituire la quantizzazione con l'esistenza di una soglia di caoticità: Planck 1911, Einstein–Stern 1913, e l'energia di punto zero

Planck 1911: la seconda "teoria" e l'energia di punto zero

Il primo tentativo di eliminare la quantizzazione o meglio, di sostituirla con l'esistenza di un fenomeno qualitativo di tipo dinamico, fu compiuto da Planck nel 1911, con quella che è comunemente nota come la sua *seconda teoria.* Per qualche ragione che non ci ha rivelato, egli fu indotto a descrivere il processo di interazione fra un risonatore e il campo e.m. nel modo seguente. Un risonatore, partendo da una condizione iniziale di energia nulla, a causa dell'interazione con il campo assorbe energia con un processo continuo. Quando poi l'energia assorbita è giunta al valore $h\nu$ il processo diventa dinamicamente così complesso (caotico o complicato, diremmo oggi[25]) che una descrizione puramente deterministica diventa praticamente impossibile e si richiede una trattazione statistica. Allora egli ammette che, giunto all'energia $h\nu$, il risonatore abbia una certa probabilità di ripiombare immediatamente allo stato di energia nulla, e la probabilità complementare di continuare ad assorbire energia con continuità. Se l'oscillatore continua ad assorbire, il processo probabilistico si ripete quando esso ha assorbito energia totale $2h\nu$ e così via. Sulla base di tali considerazioni egli definisce un processo stocastico, e la probabilità di salto viene fissata in modo tale che l'energia media dell'oscillatore sia proprio quella di Planck.

Questa concezione viene poi rielaborata nella seconda edizione del *Wärmestrahlung* del 1912, da cui è tratta l'unica traduzione inglese esistente. I livelli

[24] Una vecchia citazione del periodo attorno a quello della prima conferenza Solvay (dove si ha una analoga ben nota citazione) si trova anche nel celebre articolo del 1909 *Lo stato attuale della teoria della radiazione*, lavoro 9 dell'edizione italiana, pag. 214, dove dice che "*le variazioni spaziali della radiazione e della pressione di radiazione avvengono* ***come se*** *la radiazione consistesse di quanti* ...".

[25] Stiamo citando a mente le parole dell'articolo di Planck.

dell'oscillatore risultano avere i valori

$$E_n = nh\nu + \frac{1}{2}h\nu\,, \quad n = 0, 1, 2, \ldots \tag{5.5.1}$$

che saranno poi ritrovati da Heisenberg nel 1925 e da Schrödinger nel 1926. Dunque segue che l'energia media specifica u ha l'espressione

$$u = \frac{h\nu}{e^{\beta h\nu} - 1} + \frac{1}{2}h\nu\,. \tag{5.5.2}$$

Il termine $h\nu/2$, che è il solo che resta allo zero assuluto, viene detto *energia di punto zero*, (in inglese, *zero–point energy*) perché *Nullspunkt* è il termine tedesco per la temperatura dello zero assoluto.

Nella seconda edizione del *Wärmestrahlung*, a pag 142 della traduzione inglese, Planck commenta: "*That the oscillators are said to perform vibrations even at the temperature zero, the mean energy of which is so large as $h\nu/2$ and hence may become quite large for rapid vibrations, may at first appear strange. It seems to me, however, that certain facts point to the existence, inside the atoms, of vibrations independent of the temperature, and supplied with appreciable energy ...*" Un effetto attribuito in qualche modo all'energia di punto zero è quello detto di Debye–Waller: da misure di dispersione da raggi Röntgen, sembra potersi dedurre che gli atomi o molecole o ioni costituenti un cristallo abbiano una certa velocità non nulla anche allo zero assoluto.

Può sembrare strano attribuire significato ad un valore definito di energia, quando si sa che l'energia potenziale è definita a meno di ua arbitraria costante additiva. Dal punto di vista della teoria quantistica la cosa si risolve nel modo seguente. Sappiamo che l'energia di punto zero deve essere identificata con l'energia del livello fondamentale (*ground state*) E_0. Ma questo stato ha distribuzioni di probabilità della posizione e del momento che sono sparpagliate, e danno allora all'energia un contributo che è quello che si avrebbe classicamente se l'oscillatore si trovasse ad una energia $h\nu/2$ *sopra il minimo dell'energia potenziale*.

Ai fini di quanto segue, ribadiamo l'idea centrale introdotta da Planck, ovvero che in corrispondenza dell'energia di punto zero si incontrano situazioni di moto complicato, diciamo di moto caotico, come in presenza di una energia di soglia: moti ordinati sotto soglia, disordinati sopra soglia. Questa idea verrà infatti ripresa nel 1916 da Nernst (a parte la scelta di porre la soglia all'energia $h\nu$ invece che ad $h\nu/2$).

Einstein e Hopf, 1911; Einstein e Stern, 1913

Il lavoro di Einstein e Hopf del 1911 è concettualmente molto interessante, ma fino ad oggi nessuno è stato in grado di elaborarlo in modo da farlo diventare una vera teoria. Forse gli autori che modernamente vi sono andati vicino sono quelli che si riconoscono nella cosiddetta elettrodinamica stocastica di Boyer. L'idea

centrale di Einstein e Hopf è di studiare il moto di un sistema di molecole biatomiche (sostanzialmente, ciascuna molecola può essere pensata come costituita da un baricentro e da una molla lineare) che interagiscono mutuamente sia per collisioni, sia con forze di dipolo elettrico mediate dal campo elettromagnetico. Gli autori impongono che si abbia una situazione di equilibrio statistico. Ammettendo che l'energia cinetica media dei baricentri sia identificabile come di consueto con la temperatura, essi vogliono studiare che cosa impone la condizione di equilibrio, sulla energia media U delle molle. Quest'ultima energia risulta determinata da una equazione differenziale ordinaria del primo ordine nella variabile $x = h\nu/k_B T$. Naturalmente, per risolvere l'equazione occorre assegnare un dato iniziale o al contorno, ed essi impongono quello che sembrerebbe spontaneo, ovvero

$$U(T) \to 0 \quad \text{per} \quad T \to 0\,.$$

In tal modo ottengono il risultato di equipartizione, energia $U = Nk_B T$.

Ma dopo la pubblicazione della seconda teoria di Planck, Einstein e Stern hannno una ispirazione. Risolvere la medesima equazione con la condizione al contorno

$$U(T) \to \frac{N}{2} h\nu \quad \text{per} \quad T \to 0\,.$$

Con un calcolo immediato ottengono allora che l'energia media specifica è data dalla formula di Planck, con energia di punto zero $h\nu/2$, ovvero la formula (5.5.2) di Planck.

5.6 La risposta di Nernst a Poincaré: 1916

Una risposta a Poincaré venne data da Nernst in un lungo articolo del 1916. L'idea centrale di Nernst è basata sulla concezione, che gli provenne dalla sua scoperta del terso principio della termodinamica che egli fece nel 1906, che a bassissime temperature l'energia abbia dinamicamente un carattere ordinato (in tal modo precorrendo i risultati della teoria dei sistemi dinamici degli anni 1960), e quindi interpretando l'energia di punto zero di Planck come una energia ordinata. Si tratta dunque di una ricomprensione dell'dea già avanzata da Boiltzmann, che solo una parte dell'energia meccanica contribuisca all'energia interna termodinamica, ovvero solo quella concretamente scambiabile, avendo dinamicamente un carattere disordinato. Procedendo lungo tale intuizione, egli pervenne a una interantissima generalizzazione dell'energia di punto zero, in cui la formula di Planck risulta essere compatibile con il principio di equipartizione, un risultato che a noi pare estremamente significativa.

5.6.1 Il terzo principio della termodinamica (Nernst, 1906), e la concezione dell'energia di punto zero come energia ordinata

Nernst formulò il terzo principio, nel 1906, in una maniera interessantissima completamente fenomenologica e termodinamica, in una maniera qualitativa-

mente ben diversa da quella riportata in una parte preponderante dei manuali, in cui la formulazione viene data in termini di meccanica statistica anziché in maniera fenomenologica. Per i nostri scopi, la formulazione fenomenologica è rilevante, perché mette in luce il carattere ordinato dei moti in prossimità dello zero assoluto.

Anzitutto bisogna avere chiaro che, a differenza della energia interna U, l'energia libera F risulta essere quella parte di energia interna che è atta a compiere lavoro macroscopico. Per questo in tedesco essa viene denotata con la lettere A (*Arbeit* significa lavoro) e non con la lettera F. Anzi, Nernst formula il secondo principio in analogia con il primo. Come il primo principio afferma che esiste la funzione di stato U, l' energia interna al cui incremento contribuiscono sia il lavoro sia il calore, così il secondo afferma che esiste una funzione di stato $F(T) \equiv A(T)$, detta energia libera, che rappresenta la massima quantità di lavoro che il sistema può produrre in un processo isotermo.

Egli aveva osservato, dai grafici delle quantità U ed F in funzione di T (osservate in processi coinvolgenti delle pile, a temperature ambiente o comunque non molto basse), che tali due energie diventano uguali al di sotto di una temperatura caratteristica T^c (detta da lui "*temperatura di degenerazione*"). Egli allora formulò il terzo principio asserendo che tale fenomeno è universale, valendo per ogni sistema, al disotto di una corrispondente temperatura di degenerazione. Dunque, al di sotto della temperatura di degenerazione tutta l'energia interna è disponibile per compiere lavoro macroscopico, e quindi ha carattere "*ordinato*": pertanto l'energia interna ha carattere ordinato non solo in prossimità dello zero assoluto, ma anche a tutte le temperature sotto la temperatura di degenerazione (dell'ordine di 300 K per le pile). In termini forse più familiari a molti lettori, ricordando che l'energia libera si esprime nella forma $F = U - TS$, questo vuol dire che sotto la temperatura di degenerazione, essendo $F = U$, si ha $S = 0$. L'unico punto che potrebbe non essere evidente a un lettore, è che Nernst non parla mai della energia, o energia libera o entropia possedute da un sistema, ma delle corrispondenti quantità ΔU, ΔF o ΔS (che egli denota semplicemente con U, F, S) relative ad una certa trasformazione, ad esempio una reazione chimica.

In conclusione, l'aspetto dinamico caratterizzante i moti allo zero assoluto (o meglio, sotto la temperatura di degenerazione) è il loro carattere di moti prevalentemente ordinati. Si noti bene che fino a non molto tempo fa gran parte della comunità scientifica riteneva, come puro pregiudizio di cui sarebbe interessante studiare l'origine, che nei sistemi classici costituiti da un numero macroscopico di costituenti sarebbero sempre dominanti, a tutte le temperature, i moti caotici. Non meraviglia allora che, dopo la scoperta fatta da Fermi nel suo lavoro con Pasta ed Ulam del 1954, che mostrava l'esistenza di moti ordinati in un modello di solido (monodimensionale, con 32 o 64 particelle, a bassa energia), la quasi totalità della comunità abbia creduto che il fenomeno dovesse scomparire per modelli macroscopici. E questo, nonostante il fatto che nello stesso anno 1954 del lavoro FPU il grande Kolmogorov avesse pubblicato un sottilissimo lavoro matematico in cui mostrava che i rapporti tra moti ordinati e moti caotici dovessere essere

ricompresi in un modo fino ad allora insospettato.

Nernst 1916, risposta a Poincaré: legge di Planck compatibile con la legge di Maxwell–Boltzmann (ovvero, con l'equipartizione)

In un lungo articolo, di difficilissima lettura e di grandiose prospettive,[26] dal titolo *Su un tentativo di ritornare dalla meccanica quantistica alla concezione di variazioni continue di energia*,[27] [28] Nernst fornisce (come dice nelle prime righe) una risposta al lavoro nel quale Poincaré aveva mostrato, nel 1912, che per ottenere la legge di Planck si deve rinunciare "*all'applicabilità del più importante strumento logico, le equazioni differenziali*". Come vedremo, la risposta di Nernst è basata sull'idea centrale già concepita da Boltzmann (si ricordi il caso delle sfere rigide perfettamente lisce, in cui l'energia meccanica di rotazione, pur essendo in generale presente, non contribuisce all'energia interna termodinamica). L'idea è che l'energia termodinamica non coincida con l'energia meccanica del sistema considerato, ma solo con quella frazione di energia che è "disordinata", e quindi può effettivamente essere scambiata in un processo termico. In effetti, come egli stesso dice in due punti del suo lavoro, egli fornisce solo una idea, in qualche modo "*indovina una formula*". Qui cerchiamo di illustrare questa sua divinazione, che corrisponde a rinunciare ad esprimere l'energia termodinamica come valor medio dell'energia meccanica rispetto ad una misura di probabilità invariante.

In breve, egli propone anzitutto la seguente

Congettura dinamica: Per un sistema di oscillatori di frequenza ν, l'energia

$$\varepsilon = h\nu$$

costituisce una soglia di caoticità: al di sotto si hanno moti prevalentemente ordinati (*geordnete*), mentre al di sopra si hanno moti prevalentemente disordinati (*ungeordnete*).[29]

Poi propone il seguente

Schema teorico: Gli oscillatori sono disrtibuiti secondo Maxwell–Boltzmann (e quindi con equipartizione), e dunque è proprio mediante la distribuzione di MB che deve essere calcolata l'energia scambiabile, concepita come l'energia che si ottiene da $k_B T$ sottraendo l'energia ordinata.

[26] Si tenga presente che, come seguito del lavoro qui discusso, nel 1936 Netnst scrisse un lungo articolo sull'energia di punto zero in relazione alla cosmologia, in cui tra l'altro propose che dovesse esistere una radiazione di fondo ad una temperatura che stimava dell'ordine di 2.7 K.

[27] W. Nernst, traduzione inglese ...

[28] Si veda anche l'applicaazione alla cosmologia, in cui Nernst predisse l'esistenza di una radiazione di fondo a circa 3 K. W. Nernst, Ann. d Phys. 32, 44 (1938), Zeits. f. Phys. 606, 633 (1917). Pecker et al, Research in Astronomy and Astrophysics 15 N.4 (2015), H. Kragh. J. Overduin, The weight of the vacuum, Springer.

[29] Questa congettura é conforme alla concezione principale della teoria delle perturbazioni, secondo la quale un oscillatore armonico (i cui moti sono completamente di tipo ordinato), sotto perturbazione mantiene moti prevalentemente ordinati fino ad una soglia di energia E_c crescente linearmente con la sua frequenza ν. Infatti, la quantità rilevante per la teoria delle perturbazioni è l'azione E/ν, che è un "invariante adiabatico", come era ben noto fin dai tempi di Boltzmann.

Seguendo questo schema Nernst ricava la formula di Planck in ambito classico mediante due procedimenti. Il secondo ci risulta del tutto incomprensibile, per la nostra incapacità di seguire le profondissime intuizioni di Nernst. Invece il primo precedimento era stato già compreso, o forse interpretato, molti anni fa,[30] nel modo che ora riportiamo.

Ricordiamo che la densità di probabilità $p_\beta(E)$ dell'energia di un oscillatore secondo MB è data da

$$p_\beta(E) = \frac{1}{Z(\beta)} e^{-\beta E} , \qquad Z(\beta) = \int_0^\infty e^{-\beta E} \mathrm{d}E ,$$

ovvero, poiché si calcola immediatamente $Z(\beta) = 1/\beta$,

$$p_\beta(E) = \beta e^{-\beta E} .$$

Si introducono ora le quantità

$$n_0(\beta) = \int_0^\varepsilon p_\beta(E)\mathrm{d}E , \quad n_1(\beta) = \int_\varepsilon^\infty p_\beta(E)\mathrm{d}E ,$$

che ovviamente danno la frazione di oscillatori sotto soglia e sopra soglia. Si definiscono analogamente le quantità

$$E_0(\beta) = \frac{1}{n_0} \int_0^\varepsilon E\, p_\beta(E)\mathrm{d}E , \qquad E_1(\beta) = \frac{1}{n_1} \int_\varepsilon^\infty E\, p_\beta(E)\mathrm{d}E ,$$

che danno l'energia media del singolo oscillatore sotto soglia e sopra soglia.[31] Si hanno evidentemente le relazioni

$$n_0 + n_1 = 1 , \quad n_0 E_0 + n_1 E_1 = k_B T , \qquad 0 < E_0 < \varepsilon , \quad \varepsilon < E_1 < \infty \quad .$$

Si calcola immediatamente che vale

$$n_1 = e^{-\beta\varepsilon}$$

e quindi si ottiene anche

$$n_0 = 1 - n_1 = 1 - e^{-\beta\varepsilon} .$$

Analogamente, il calcolo dell'integrale che definisce E_1 si compie eseguendo una traslazione di ε nell'energia, riconducendosi in tal modo al classico calcolo che fornisce l'equipartizione per MB. Si ottiene in tal modo

$$E_1 = k_B T + \varepsilon .$$

[30] Si veda L. Galgani ..., L. Galgani, G. Benettin ... , L. Galgani, in ....

[31] In termini probabistici, si tratta di *energie condizionate*.

Possiamo allora ottenere anche l'espressione di E_0. Infatti, dalle due relazioni $n_0E_0 + n_1E_1 = k_BT$, $E_1 = k_BT + \varepsilon$ si ottiene

$$E_0 = \frac{1}{n_0}k_BT - \frac{n_1}{n_0}(k_BT + \varepsilon) ,$$

ovvero, usando $1 - n_1 = n_0$,

$$E_0 = k_BT - \frac{n_1}{n_0}\varepsilon = k_BT - \frac{n_1}{1-n_1}\varepsilon = k_BT - \frac{e^{-\beta\varepsilon}}{1-e^{-\beta\varepsilon}}\varepsilon .$$

Infine, moltiplicando numeratore e denominatore ell'ultimo termine per $\exp(\beta\varepsilon)$ si ottiene la formula divinata da Nernst

$$E_0 = k_BT - \frac{\varepsilon}{e^{\beta\varepsilon} - 1} , \tag{5.6.1}$$

nella quale, quasi magicamente, egli vide presentarsi in maniera inaspettata la formula di Planck,

$$u_{Pl} = \frac{\varepsilon}{e^{\beta\varepsilon} - 1} , \qquad \varepsilon = h\nu , \tag{5.6.2}$$

in un ambito in cui faceva uso della distribuzione di MB. Infatti la (5.6.1) non è altro che

$$E_0 = k_BT - u_{Pl} . \tag{5.6.3}$$

Equivalentemente questa relazione si esprime nella forma

$$k_BT = u_{Pl} + E_0 , \tag{5.6.4}$$

da cui appare che l'energia media k_BT di un oscilatore si decompone nella somma della energia u_{Pl} di Planck e dell'energia ordinata E_0.

Interpretazione dell'energia di Planck copme energia termica (ovvero scanbiabile termicamente)

Si ricordi ora che si ha, come sappiamo, $k_BT = n_0E_0 + n_1E_1$, e dunque, usando $n_0 = 1 - n_1$, si ha anche

$$k_BT = n_1(E_1 - E_0) + E_0 . \tag{5.6.5}$$

In conclusione, per confronto con la (5.6.4), abbiamo anche

$$u_{Pl} = n_1(E_1 - E_0) . \tag{5.6.6}$$

Le relazioni (5.6.4) e (5.6.6) si prestano a una interessante interpretazione. Per semplicità di descrizione pensiamo la frazione n_1 di oscillatori sopra soglia come disposti su un livello di energia E_1, e analogamente la frazione n_0 sotto soglia, come disposti su livello E_0. Per scambi termici di energia, gli oscillatori sotto soglia non contano, e contano solo quelli sopra soglia, che scambiano energia saltando sul livello inferiore. Quindi l'energia scambiabile è $n_1(E_1 - E_0) = u_{Pl}$, cioè proprio l'energia di Planck.

Analogia con la la formula quantistica coinvolgente l'energia di punto zero

La formula (5.6.4) costituisce l'analogo della formula quantistica in cui l'energia media $\langle E \rangle$ di un oscillatore si esprime nella forma

$$\langle E \rangle = u_{Pl} + \frac{1}{2} h\nu \,,$$

in cui figura la "energia di punto zero" $\frac{1}{2}h\nu$. La differenza naturalmente è che l'energia di punto zero è una costante, mentre l'analogo termine E_0 in Nernst dipende dalla temperatura (oltre che dalla frequenza). Non solo. Ma data la relazione $E_0 = k_B T - u_{Pl}$ si ha anche $E_0 < k_T$ sicché sia ha anche $E_0 \to 0$ per $\to 0$.

Invece le cose vanno meglio per alte (e anche non troppo alte) temperature. Infatti sappiamo che per alte temperature si ha (da uno sviluppo al secondo ordine in $h\nu/k_B T$)

$$u_{Pl} \simeq k_B T - \frac{1}{2} h\nu \,,$$

ma in effetti questa formula asintotica già si applica per temperature tali che $k_B T \simeq 2h\,\nu$. Dunque per tali temperature dalla (5.6.3) si ha

$$E_0 \simeq \frac{1}{2} h\nu \,.$$

Pertanto, per tali temperature l'energia ordinata di Nernst coincide con l'energia di punto zero prevista dalla meccanica quantistica, trovata da Heisenberg nel 1925 e da Schrödinger nel 1926.

NOTA PERE GLI AUTORI. Vedere se si possono applicare qui i risulati ottenuti sul LiF, da cui appare che la relazione tra energia specifica e temperatura non e' quella di Gibbs, e anzi si trova una energia si punto zero non nulla. Cosa comporta questa nuova identificazione microscopica della temperatura per il ragionamnento di Nernst ?

NB Nella seconda deduzione di Nernst, che non riusciamo a capire, si ha una zpe indipendente dalla temperatura. Vedere l'articolo del 1936 (circa) dove Nernst applica l'energia di punto zero all'universo e trova che esso ha una temperatura di 2.7 K NB. L'interpretazione del procedimento di Nernst potrebbe essere il seguente. Egli dice che per glli oscillatori ci sono due forme (arten) di energia, e formula un principio di minimo, tipo minimo di energia libera, per determinare la suddivisione dell'energia tra queste due forme, a fissata temperatura. Si puo' pendare che una forma e' quella che cossriponde a stare sui tori, l'altra analoga per la regione tra i tori.

Relazione con il procedimento di quantizzazione

Ben conosciamo il modo, illustrato la prima volta da Einstein nel suo lavoro del 1907 sui calori specifici e riportato in tutti i manuali, con cui si ottiene la formula di Planck u per l'energia media di un oscillatore armonico con energia quantizzata $E_n = n\varepsilon$, $(n = 0, 1, \ldots)$, dove $\varepsilon = h\nu$. Si tratta della relazione

$$u_{Pl} = \sum_{n=0}^{\infty} p_n n\varepsilon \ ,$$

dove

$$p_n = \frac{e^{-\beta n\varepsilon}}{Z_q(\beta)} \ , \quad Z_q(\beta) = \sum_{n=0}^{\infty} e^{-\beta n\varepsilon} \ .$$

Equivalentemente si ha

$$u_{Pl} = -\partial_\beta \log Z_q(\beta) \ . \tag{5.6.7}$$

Quindi, come era evidente, il procedimento di Nernst, che non invoca nessuna quantizzazione ma legge l'energia u_{Pl} di Planck come energia scambiabile, risulta equivalente a quella familiare che vede u_{Pl} come valor medio di energia meccanica, che però é quantizzata.

Tuttavia, si verifica immediatamentte come sia possibile ritrovare questa equivalenza in una maniera del tutto spontanea, con un procedimento che, idealmente, sarebbe stato possibile compiere anche senza avere mai sentito parlare di quantizzazione. Consideriamo infatti la formula di Nernst data dalla seconda delle (5.6.4), in cui u appare come differenza di due termini, $u_{Pl} = k_B T - E_0$. Ora, ben sappiamo che il primo termine, quello di MB, si esprime come

$$k_B T = -\partial_\beta \log Z \ , \quad \text{dove} \quad Z = \int_0^\infty e^{-\beta E} \mathrm{d}E \ ,$$

mentre il secondo, E_0, si esprime, come subito si vede dalla sua definizione, in una forma analoga, data precisamente da

$$E_0 = -\partial_\beta \log Z_0 \ , \quad \text{dove} \quad Z_0 = \int_0^\varepsilon e^{-\beta E} \mathrm{d}E \ .$$

Dunque si ha evidentemente

$$u_{Pl} = -\partial_\beta \log Z_{\text{eff}} \ ,$$

dove

$$Z_{\text{eff}} = \frac{\int_0^\infty e^{-\beta E} \mathrm{d}E}{\int_0^\varepsilon e^{-\beta E} \mathrm{d}E}$$

La struttura stessa di questa formula invita a decomporre il numeratore nella somma di una serie di integrali:

$$\int_0^\infty e^{-\beta E} \mathrm{d}E = \sum_{n=0}^{\infty} \int_{n\varepsilon}^{(n+1)\varepsilon} e^{-\beta E} \mathrm{d}E \ .$$

Ma con il cambiamento di variabile $E = E' + n\varepsilon$ ognuno di tali integrali prende la forma (denotando ancora la variabile di integrazione E' con E)

$$e^{-\beta n\varepsilon} \int_0^{\varepsilon} e^{-\beta E} \mathrm{d}E \ ,$$

sicché tutti i termini contengono come fattore un medesimo ntegrale, che si cancella esattamente con il denominatore. Si resta in tal modo con la serie definente Z_q, ovvero si ha

$$Z_{\text{eff}} = \sum_{n=0}^{\infty} e^{-\beta n\varepsilon} \equiv Z_q$$

e dunque l'energia scambiabile di Nernst si scrive nella forma quantistica, ovvero come energia media del sistema quantizzato. Questo sembra realizzare il desiderio espresso da Einstein già alla prima conferenza Solvay del 1911, ovvero (sostanzialmente nelle sue stesse parole): si dovrebbe poter dimostrare come, ai fini della meccanica statistica, cioè della deduzione meccanicistica della termodinamica, "tutto va come se" i sistemi meccanici fossero quantizzati. Tuttavia (sono ancora le sue parole), la quantizzazione non sarebbe necessaria, essendo solo un metodo comodo per compiere correttamente i calcoli.

Relazione con la formula dinamica "alla Kubo" per la misurazione del calore specifico

Abbiamo dunque mostrato che nella (5.6.4) l'energia interna di Planck appare come la differenza di due termini, di cui il primo è il valore classico di equilibrio, mentre il secondo appare come una correzione che in qualche modo tiene conto del fatto che siamo in presenza di moti non completamente caotici. Se dunque ricaviamo la corrispondente espressione per il calore specifico C_V, semplicemente derivando rispetto a T, troviamo ancora due analoghi termini, di cui il primo è il clasico calore specifico di equilibrio, e il secondo una correzione dovuta al fatto che siamo in presenza di moti non completamente caotici.

Ora, quasi incredibilmente questo è proprio quello che viene dimostrato nella formula alla Kubo che verrà discussa nel prossimo capitolo. Come vedremo, una rilevante differenza è che la correzione data da Kubo ha una forma ancora alquanto generale anziché la forma particolare indovinata da Nernst per l'oscillatore armonico, che per l'energia scambiabile dà esattamente la formula di Planck. Comunque, resta il fatto che nel prosedimento di Kubo il calore specifico risulta essere la differenza di due termini, di cui il primo è esattamente il calore specifico classico (letto come varianza, alla maniera di Boltzmann–Einstein), mentre il secondo termine costituisce una correzione dovuta al fatto che il sistema non è completamente caotico.

Capitolo 6

I tempi di rilassamento in termodinamica e in meccanica statistica. Da Boltzmann e Gibbs a Kubo

Nei capitoli precedenti abbiamo illustrato come Boltzmann proponesse di sfuggire alle conseguenze del principio di equipartizione mediante l'ipotesi che fossero necessari tempi molto lunghi per ottenere la termalizzazione di tutti i gradi di libertà. Però questa idea non fu da lui mai formalizzata. Fu solo dopo gli anni '50 del secolo scorso, a seguito dei lavori di Green, Kubo ed altri che è stato possibile mettere in luce il ruolo che la dinamica gioca nel determinare il valore delle grandezze termodinamiche. In questo capitolo mostreremo come la dinamica possa influire sui valori misurati del calore specifico, cioè come avvenga, se la dinamica non è sufficientemente caotica, che i valori ottenuti possano essere in disaccordo con il principio di equipartizione.

6.1 Introduzione

Il problema dei tempi di rilassamento in termodinamica e in meccanica statistica è un oggetto in qualche modo misterioso, che si aggira dappertutto senza essere stato davvero risolto. Il problema si pone già al livello della definizione di uno stato termodinamico. Ad esempio, in un sistema fluido a fissato numero di moli uno stato termodinamico è individuato tipicamente da temperatura T e volume V, e dunque l'insieme degli stati è un certo dominio in un piano riferito a coordinate (T, V). Naturalmente, quando si dice che il fluido si trova in uno di questi stati si ammette che il sistema sia "in equilibrio". Poi si compie un certo "processo" che conduce ad un altro punto (T', V'), e si deve attendere un tempo sufficiente perché il sistema sia ancora in equilibrio e si abbia quindi un nuovo

stato termodinamico individuato dal punto (T', V'). Se ad esempio cambio il volume di un pistone in cui è contenuto il fluido, devo attendere che il fluido ritorni "in quiete", il che può richiedere "poco tempo", oppure "molto". Situazioni analoghe si presentano nei processi di misurazione. Ad esempio, nel caso del calore specifico a volume costante si fa passare il sistema da T a T' mettendolo a contatto con un termostato a temperatura T', e durante tale processo si compie anche una misura del calore scambiato con il termostato (ad esempio si misura quanto ghiaccio si scioglie se il termostato è a zero gradi centigradi, e il sistema a temperatura superiore). Ci sono naturalmente diversi metodi empirici per determinare se è stato raggiunto uno stato di equilibrio o meno. In quesro capitolo mostreremo che, almeno per il processo di misurazione del calore specifico, esiste una prescrizione teorica che permette di giudicare se l'equilibrio termodinamico sia stato raggìunto o meno. Al tempo stesso questo criterio permette di assegnare al calore specifico un valore anche per sistemi, diciamo, ordinati, diverso da quello ricavato secondo le usuali formule della meccanica statistica.

Nernst e le misure dei calori specifici dei solidi a basse temperature

Proprio quello della misurazione del calore specifico è, a questo proposito, un esempio significativo, quasi paradigmatico, che ha coinvolto nientemeno che il grande Walther Nernst. Nel 1911 egli diede inizio a una campagna sistematica di studi sperimentali sul calore specifico dei solidi per verificare che esso si annullasse per T tendente a zero. Questo infatti era previsto dal terzo principio della termodinamica, che egli stesso aveva formulato nel 1906, ed era suggerito dal lavoro di Einstein del 1907. Tuttavia, nel compiere le misure, proprio quando i valori misurati cominciavano a diventare sensibilmente inferiori a quello previsto dal principio di equipartizione (legge di Dulong e Petit) egli trovò che i tempi necessari perché la misura si stabilizzasse crescevano fortemente. A un certo punto i tempi divenivano talmente lunghi che egli giunse a dubitare che addirittura "esistesse" un valore per il calore specifico a basse temperature. Poi, inaspettatamente, egli trovò che, diminuendo ulteriormente la temperatura, i "tempi di rilassamento" diminuivano, sicché per il calore specifico restavano ben definiti dei valori, che effettivamente diminuivano al decrescere della temperatura.[1]

Tre esempi significativi di situazioni di metaequilibrio

In generale può avvenire che il calore specifico C_V abbia un valore misurato ben preciso, che tuttavia non si accorda affatto con quello predetto dalla meccanica statistica di equilibrio, né classica né quantistica. Un caso tipico è quello discusso da Landau e Lifshitz, che riguarda "the partial equilibrium of a mixture of several substances which interact chemically", e viene da loro discusso nel mo-

[1] Una indicazione di come potrebbe spiegarsi questo fenomeno a livello dinamico potrebbe trovarsi nel lavoro L. Berchialla, L. Galgani, A. Giorgilli, *Localization of energy in FPU chains*, Discr. Cont. Dyn. Systems -A 11, 855-866 (2004).

do seguente. *"Owing to the relative slowness of chemical processes, the equilibrium connected with the motion of molecules (cioè dei baricentri delle molecole) occurs, in general, considerably sooner than the equilibrium connected with the interchange of atoms between the molecules (processo che coinvolge una redistribuzione degli elettroni), i.e., connected with the composition of the mixture".* In un lavoro recente[2] si trova il seguente esempio. "For example. in a system composed of a mole of graphite C and a mole of O_2, at room temperature combustion does not occur (questo è il processo che richiede un tempo lunghissimo, se non si fa intervenire un catalizzatore, qui una fiamma), and the measured heat capacity at constant volume is simply the sum of the specific heats of the two components, i.e., $1R + (5/2)R = (7/2)R$, whereas at equilibrium the heat capacity is that of one mole of carbon dioxide CO_2, namely, $(5/2)R$".

Si ha poi un caso estremamente significativo, in cui si ha discrepanza con le previsioni della meccanica statistica *quantistica di equilibrio*. Anche in questo caso i valori di equilibrio potrebbero essere misurati in tempi brevi in presenza di opportuni catalizzatori, ma in situazioni ordinarie il raggiungimento dell'equilibrio richiede tempi incredibilmente lunghi. L'esempio è quello dell'orto e del para idrogeno, che coinvolge la interazione (non efficace) tra nuclei ed elettroni; si veda il bellissimo libro di Wannier, sec.11-4, pag. 218.[3] Si hanno ancora sostanzialmente due specie chimiche diverse che di fatto (a meno di essere in presenza di opportuni catalizzatori) non interagiscono, sicché il calore specifico misurato risulta essere una somma pesata dei calori specifici delle due specie considerate, previsti dalla meccanica statistica quantistica di equilibrio. Ma questa somma pesata ha valori diversi da quelli previsti, per il sistema totale, dalla meccanica statistica quantistica di equilibrio. Questo fatto sembra suggerire che i metodi della MQ statistica *di equilibrio* siano adeguati per descrivere i fenomeni empiricamente osservati a livello atomico, ma potrebbero non esserlo se sono coinvolti gradi di libertà sia atomici che nucleari, perché essi 'non comunicano" entro tempi "umani".

Un altro un caso significativo di metastabilità è quello dei vetri. Essi sono considerati dei fluidi, ma non appaiono "scorrere", fluire, su scale di tempi umani perché a temperature ordinarie hanno una viscosità con un valore che è qualcosa come 10^{14} volte superiore a quello che presentano diciamo a 1000 K. A temperature ordinarie il rilassamento all'equilibrio coinvolge tempi di tipo almeno geologico.

Infine, come abbiamo già ricordato, molti dei "grandi" avevano indicato la rilevanza, a livello qualitativo, dei tempi di rilassamento. Tra questi, anzitutto Boltzmann stesso,[4] poi Rayleigh[5] e anche Poincaré (con il suo articolo del 1906 citato nl capitolo su FPU).

[2] A. Carati. A. Maiocchi, L. Galgani, *Statistical thermodynamics for metaequilibrium or metastable states*.

[3] G.H. Wannier, *Statistical Physics*, Dover (New York, 1966).

[4] NOTA PER GLI AUTORI. Raccogliere le citazioni da Boltzmann (articolo con Barbara),

[5] Nella lettera inviata a Nernst in occasione della prima conferenza Solvay.

Scopo di questo capitolo

Problemi che coinvolgono tempi di rilassamento all'equilibrio che possono essere lunghissimi, si presentano dunque in molte situazioni. Fu Boltzmann stesso per primo a suggerire che il tener conto di tale fatto (cioè che i tempi di rilassamento possono essere lunghissimi) possa spiegare il fallimento del principio di equipartizione. Egli considera un sistema che si trova in un certo stato termodinamico e possiede l'energia prevista dal principio di equipartizione. Poi viene diminuita la sua temperatura, sicché il nuovo stato termodinamico richiederebbe una certa ben definita diminuzione di energia, ma invece può avvenire che sia lunghissimo il tempo richiesto perché questo processo effettivamente abbia luogo. Il suo esempio preferito è quello di un sistema di atomi che siano sfere perfettamente lisce, perché in tal caso le energie cinetiche di rotazione dei singoli atomi sono tutte costanti del moto, e non si alterano durante le collisioni. Quindi, se si compie un processo di diminuzione di temperatura che coinvolga i baricentri, l'energia di rotazione resterà eternamente uguale a quella iniziale. Una leggera scabrezza permetterà invece degli scambi di energia tra baricentri e rotazioni durante le collisioni, ma il raggiungimento dell'equilibrio corrispondente alla nuova temperatura dei baricentri può richiedere tempi enormi. Esempi analoghi vennero discussi anche da Poincaré, nel suo articolo del 1906 in cui parlava di rilassamenti parziali dopo tempi grandi del primo ordine, del secondo ordine, ... In ogni caso, quello di Boltzmann è un esempio tipico nell'ambito della fisica atomica, in cui falliscono i metodi della meccanica statistica classica di equilibrio (in relazione all'equipartizione dell'energia), sicché potrebbero risultare rilevanti i metodi di non equilibro. Invece, risultano adeguati i metodi di equilibrio della meccanica statistica quantistica. Questa sembra essere la situazione che si presenta in generale per la meccanica statistica classica e quella quantistica in fisica atomica: in molti esempi in cui in ambito classico fallisce la meccanica statistica di equilibrio, e sembrerrebbe richiesta una trattazione mediante metodi di nonequilibrio, risulta invece in ambito quantistico la meccanica statistica di equilibrio.

A questo punto, se facciamo riferimento alla descrizione classica, il problema si complica moltissimo, perché dobbiamo distinguere tra energia che un sistema dovrebbe scambiare (con un processo che potrebbe richiedere tempi lunghissimi) ed energia che effettivamente scambia in tempi "umani", mentre d'altra parte si vuole stabilire con metodi statistici una teoria termodinamica, la quale richiede misurazioni che in qualche modo si stabilizzino. Come si deve allora formalizzare questo fatto? Vedremo nel presente capitolo che, nell'ambito di questo problema davvero delicato, una risposta incredibilmente limpida e concettualmente semplice viene fornita da una formula che descrive come, in virtù della dinamica, il risultato di una misurazione dipende dal tempo di osservazione. Tale formula, è stata stabilita recentemente nel caso del calore specifico, seguendo un procedimento ispirato ad un metodo generale introdotto nel 1957 da Kubo nell'ambito della cosiddetta *teoria della risposta lineare*, e sembrerebbe costituire un paradigma per la descrizione matematica di ogni tipo di misurazione. La

dimostrazione di questa formula costituisce il tema centrale del presente capitolo. La discussione richiederà di premettere alcune informazioni sul cosiddetto insieme (o *enesemble*) di Gibbs.

6.2 Da Boltzmann a Gibbs. La formula di Einstein–Boltzmann per il calore specifico

Abbiamo visto in un precedente capitolo come la statistica di Boltzmann si riferisse a un sistema totale costituito da N sottosistemi identici, come tale statistica riguardasse la distribuzione dei sottosistemi nello spazio μ (lo spazio delle fasi di un generico sottosistema), ovvero la successione $\{n_k\}$ dei numeri di occupazione delle celle dello spazio μ (o eventualmente – passando al continuo – la densità di particelle in ogni punto dello spazio μ). Abbiamo poi visto come venisse attribuita una probabilità ad ognuna di tali distribuzioni, inferendola da una probabilità a priori definita nello spazio Γ (lo spazio delle fasi del sistema totale), tale probabilità essendo proporzionale alla misura di Lebesgue della corrispondente regione dello spazio Γ. In tal modo, quando sia fissata l'energia totale del sistema la distribuzione di massima probabilità (in effetti l'unica avente probabilità sostanzialmente non nulla) risultava essere la distribuzione di Maxwell–Boltzmann per i numeri di occupazione delle cellette, che nel caso continuo è proporzionale ad $\exp(-\beta H)$ dove H è la Hamiltoniana di un sottosistema.

Poi, improvvisamente nel 1902 piombò sulla comunità scientifica come un fulmine a ciel sereno l'opera principale di Gibbs, il primo vero scienziato (chimico–fisico e matematico) del nuovo mondo. Si tratta del libro (del 1902) *Elementary principles in Statistical Mechanics*, che venne subito elogiato con grande enfasi da Poincaré che era allora (al volgere del secolo) uno dei massimi (se non il massimo) scienziato (matematico e fisico–matematico) del vecchio mondo.

Con Gibbs, dello spazio μ si perde in linea di principio ogni traccia (salvo che esso ricompaia in seguito per comodità di calcolo), e tutta la trattazione fa riferimento allo spazio Γ. Si considera un sistema globale, composto da un numero N (da pensarsi enorme) di sottosistemi, diciamo particelle, completamente arbitrari (in particolare, non necessariamente identici), e con una arbitraria hamiltoniana $H(z)$ per il sistema totale. Qui $z = (q, p)$ denota l'insieme delle coordinate canoniche del sistema totale (ad esempio $q = (x_1, x_2, \ldots, x_N)$ se il sistema è costituito di N particelle libere di muoversi su una retta, e x_j sono le corrispondenti coordinate cartesiane).

Le misure di probabilità nello spazio Γ, intese come "stati"

Il grande salto concettuale rispetto a Boltzmann consiste nella assoluta distinzione ed indipendenza mutua tra dinamica e probabilità (o, meglio, come vedremo, tra dinamica e probabilità dei dati iniziali). La dinamica è semplicemente quella hamiltoniana, deterministica, e reversibile sotto inversione temporale (su questo

secondo aspetto ritorneremo più avanti). Stiamo considerando il caso tipico di una hamiltoniana indipendente dal tempo, sicché le equazioni di Hamilton definiscono un flusso $z \to \Phi^t z$ (soluzione dell'equazioni di moto al tempo t con dato iniziale z) che è un gruppo rispetto a t (sostanzialmente, con la proprietà $\Phi^s\big(\Phi^t z\big) = \Phi^{t+s} z$), e per ogni tempo t la trasformazione

$$z \to \Phi^t z$$

è canonica. Dunque, in particolare, ogni dominio iniziale D viene trasformato in un dominio $\Phi^t D$ di forma in generale diversa, ma avente il medesimo volume di D (cioè la medesima misura di Lebesgue). Ovvero, nel cambiamento di variabili $y = \Phi^t z$ si ha

$$y = \Phi^t z \qquad \text{comporta} \qquad \mathrm{d}y = \mathrm{d}z \ .$$

É questo il cosiddetto *primo teorema di Liouville*.

Questa è la dinamica: ogni punto iniziale z produce una sua orbita $\{\Phi^t z\}$, e corrispondentemente ogni variabile dinamica, od osservabile, F (una funzione $F(z)$ su Γ, a valori reali) cambia valore, nel modo dettato dalla dinamica. Se al tempo iniziale il sistema si trova in z, essa vale $F(z)$, mentre al tempo t essa assume il valore $F(\Phi^t z)$. La probabilità in principio entra solo come ausilio di fronte all'ignoranza dei dati iniziali (fatto ovvio nel caso di un sistema macroscopico). Se inizialmente il sistema si trova con probabilità $1/3$ nel punto z e con probabilità $2/3$ nel punto z', allora al tempo t si troverà con probabilità $1/3$ nel punto $\Phi^t z$ e con probabilità $2/3$ nel punto $\Phi^t z'$: la probabilità viene dunque trasportata lungo il flusso.

Più in generale Gibbs introduce una distribuzione di probabilità nello spazio Γ, come una generalizzazione dello stato microscopico del sistema considerato (ricordiamolo: stiamo parlando di un sistema "totale" composto di N sottosistemi arbitrari, ad esempio N particelle). Gibbs considera all'istante iniziale non un singolo sistema "totale", ma, come egli dice, un insieme (o un *ensemble*) di sistemi totali, ciascuno dei quali evolve come se gli altri non esistessero, mediante le leggi dell'unica dinamica che è definita, nello spazio Γ, dalla assegnata hamiltoniana. Si può addirittura passare ad un insieme continuo di sistemi,[6] definito da una densità $\rho(z)$. Un fissato insieme, che permette il calcolo dei valori medi delle variabili dinamiche di interesse, costituisce quindi un "stato", come generalizzazione dello stato microscopico "esatto" che si ottiene nel caso particolare in cui si abbia un solo ben determinato punto iniziale z^*, ovvero nel caso in cui la densita ρ si riduce a una delta di Dirac, $\rho_0(z) = \delta(z - z^*)$.

Consideriamo dunque uno "stato" iniziale nel caso continuo, definito mediante una densità di probabilità iniziale ρ_0, Ciò vuol dire che la probabilità iniziale $P_0(D)$ che il sistema si trovi in un arbitrario dominio D è data da

$$P_0(D) = \int_D \rho_0(z)\mathrm{d}z = \int_\Gamma \rho_0(z)\chi_D(z)\mathrm{d}z$$

[6]Tuttavia, in un suo libro di introduzione alla teoria delle probabilità Varhadan dice che il passaggio al continuo "is a mess".

dove $\chi_D(z)$ è la funzione caratteristica del dominio D. Dunque la probabilità che il sistema si trovi inizialmente nel dominio D viene calcolata come valor medio (rispetto alla densità ρ_0) della funzione (o variabile dinamica o osservabile) $\chi_D(z)$. Analogamente, il valor medio $\langle F\rangle(0)$ di ogni variabile dinamica $F(z)$ sarà definito dalla formula

$$\langle F\rangle(0) = \int_\Gamma \rho_0(z)F(z)\mathrm{d}z \ . \tag{6.2.1}$$

In analogia con l'esempio relativo al "caso discreto" considerato sopra (in cui si avevano due soli dati iniziali), è allora naturale ammettere che il valor medio della variabile dinamica F al tempo t, sia dato da

$$\langle F\rangle(t) = \int_\Gamma \rho_0(z)F(\Phi^t z)\mathrm{d}z \ . \tag{6.2.2}$$

Questa formula può essere riscritta in un modo equivalente come

$$\langle F\rangle(t) = \int_\Gamma \rho_0(z)F_t(z)\mathrm{d}z \ . \tag{6.2.3}$$

dove abbiamo introdotto la funzione $F_t(z)$, "evoluta di F al tempo t", definita da

$$F_t(z) = F(\Phi^t z) \ . \tag{6.2.4}$$

Nota. Ci atteniamo qui a una esposizione di tipo tradizionale, in cui le misure di probabilità sono lisce (*smooth*), ad esempio infinitamente differenziabili. Dopo la scoperta di Kolmogorov del 1954, è tuttavia chiaro che si debbano considerare anche casi molto più complicati, che potrebbeero forse permettere di ottenere risultati significativi per sistemi dinamici classici a basse temperature, il cui prototipo è il modello di Fermi–Pasta–Ulam. Si veda la rassegna A. Carati, L. Galgani. A Maiocchi, F. Gangemi, R. Gangemi, *The FPU problem as a statistical mechanical counterpart of the KAM problem, and its relevance for the foundations of physics*, Regular and Chaotic Dynamics (Mosca, 2018), in corso di stampa.

Il procedimento duale: l'equazione di Liouville

Il modo seguito qui sopra per definire il valor medio $\langle F\rangle(t)$ è analogo a quello che verrà utilizzato in MQ da Heisenberg per definire il valor medio al tempo t di una osservabile, utilizzando lo "stato" (ovvero la misura di probabilità) al tempo 0 e facendo evolvere le osservabili. Vedremo che il metodo da Gibbs è l'equivalente di quello di Schrödinger, in cui evolve lo stato mentre restano fisse le osservabili.

Questo secondo procedimento, in cui si fa evolvere lo stato lasciando fisse le osservabili, è in effetti quello che venne introdotto per primo in ambito classico. La relativa formula si ottiene dalla (6.2.2) introducendo il cambiamento di variabile

$$y = \Phi^t z \ .$$

Ricordando che la trasformazione è canonica sicché $dz = dy$, e chiamando ancora z la nuova variabile y (che, nell'espressione del valor medio come integrale, è una variabile di integrazione o, come si dice, una variabile "muta", il cui nome è arbitrario), la (6.2.2) prende allora la forma

$$\langle F \rangle(t) = \int_\Gamma \rho_t(z) F(z) dz \ . \tag{6.2.5}$$

dove entra la densità di probabilità $\rho_t(z)$ definita da

$$\rho_t(z) = \rho_0(\Phi^{-t} z) \ , \tag{6.2.6}$$

che evidentemente potremo chiamare l' evoluta al tempo t di ρ_0. Per calcolare il valor medio $\langle F \rangle(t)$ di una osservabile F al tempo t si hanno dunque due procedimenti "duali", basati rispettivamente sulle formule (6.2.3) e (6.2.4) se si fanno evolvere le osservabili, "alla Heisenberg", e sulle formule (6.2.5) e (6.2.6) se si fanno evolvere gli stati, "alla Schrödinger".

Dalla relazione (6.2.6) risulta immediatamente[7] che la funzione $\rho(t,z) \equiv \rho_t(z)$ soddisfa l'equazione differenziale

$$\partial_t \rho + \{\rho, H\} = 0 \ , \tag{6.2.7}$$

dove $\{\rho, H\}$ è la parentesi di Poisson di ρ con H. È questo il cosiddetto *secondo teorema di Liouville*, e l'equazione (6.2.7) viene detta *equazione di Liouville*.

Gli stati di equilibrio, e lo stato di Gibbs (o insieme canonico)

Tra tutti gli stati. i più semplici sono quelli che non cambiano nel tempo, ovvero tali che $\rho_t(z) = \rho_0(z)$, e vengono detti stati stazionari o stati di equilibrio. Dalla definizione $\partial_t \rho = 0$, in virtù dell'equazione di Liouville segue subito che gli stati di equilibrio sono caratterizzati dal fatto di essere funzioni che sono costanti del moto non dipendenti esplicitamente dal tempo, ovvero caratterizzate dalla proprietà $\{\rho, H\} = 0$. In particolare, sono stati stazionari tutte le densità $\rho_0(z)$ che dipendono dal punto z solo attraverso la hamiltoniana H.

Tra tutti gli stati di equilibrio che sono funzioni dell'hamiltoniana, Gibbs privilegia la famiglia la cui densità è formalmente l'analogo della distribuzione di Maxwell–Boltzmann trasportata nello spazio Γ, ovvero

$$\rho_\beta(z) \stackrel{\text{def}}{=} \frac{e^{-\beta H(z)}}{Z(\beta)} \ , \qquad Z(\beta) = \int_\Gamma e^{-\beta H(z)} dz \ . \tag{6.2.8}$$

La corrsipondente distribuzione di probabilità vine comunemente chiamata con il nome di **insieme canonico o di Gibbs**. Naturalmente, analogamente a quanto

[7] Infatti la condizione che il valore di ρ venga trasportato inalterato lungo l'orbita soluzione delle equazioni di Hamilton, esprime il fatto che la funzione $\rho(t,z) \equiv \rho_t(z)$ è una costante del moto dipendente dal tempo, ossia che è nulla la derivata (rispetto al tempo) della funzione composta $\rho(t, z(t))$.

avviene con la distribuzione di Maxwell–Boltzmann, per il valor medio $U \stackrel{\text{def}}{=} \langle H \rangle$ dell'energia totale si trova ancora

$$U \equiv \langle H \rangle = -\partial_\beta \log Z \ ,$$

e si mostra ancora che si ha

$$\beta = \frac{1}{k_B T} \ , \qquad Z = e^{-\beta F} \ .$$

Ci ripromettiamo di dare, in una prossima edizione edlle note, almeno una sommaria indicazione di come l'insieme di Gibbs viene giustificato. Una proprietà profonda dell'insieme canonico di Gibbs, si potrebbe dire una proprietà caratteristica, è la seguente. Se si hanno due sistemi definiti nei rispettivi spazi delle fasi Γ, Γ', con hamiltoniane H, H', che si trovano separatamente in stati di Gibbs alla medesima temperatura (inversa) β, allora risulta che lo stato totale, definito nello spazio delle fasi prodotto cartesiano dei due spazi singoli, con hamiltoniana totale $H^{tot} = H + H'$, è anch'esso uno stato di equilibrio alla medesima temperatura. È questa una proprietà immediatamente evidente, conseguenza della struttura della densità di Gibbs:

$$\rho_\beta(z) = \frac{e^{-\beta H(z)}}{Z(\beta)} \frac{e^{-\beta H'(z')}}{Z'(\beta)} = \frac{e^{-\beta H^{tot}(z,z')}}{Z^{tot}(\beta)} \ , \tag{6.2.9}$$

dove

$$H^{tot}(z,z') = H(z) + H'(z') \ , \qquad Z^{tot}(\beta) \stackrel{\text{def}}{=} Z(\beta) Z'(\beta) \ . \tag{6.2.10}$$

Gibbs e MB

Una banale generalizzazione di queste due ultime relazioni ci permette di mostrare come una parte considerevole della teoria di MB venga riottenuta a partire dall'insieme canonico di Gibbs. Per confrontarci con lo schema di Boltzmann, consideriamo il caso in cui il sistema totale è composto di N sottosistemi *identici*, cioè tutti con il medesimo spazio delle fasi, ciascuno con le proprie coordinate canoniche z_i ma con la medesima hamiltoniana $H(z_i)$. Allora l'hamiltoniana totale è data da

$$H^{tot}(z_1, z_2 \ldots, z_N) = \sum_i H(z_i)$$

e l'insieme di Gibbs per il sistema totale assume una forma fattorizzata analoga a quella data sopra nel caso di due sottosistemi, ora però con le hamiltoniane tutte identiche, e anche le funzioni di partizione tutte identiche:

$$\rho_\beta(z) = \frac{e^{-\beta H^{tot}(z_1,\ldots,z_N)}}{Z(\beta)} = \frac{1}{(\mathscr{Z}(\beta))^N} \prod_{i=1}^N e^{-\beta H(z_i)} \ . \tag{6.2.11}$$

Notazioni per il sistema totale e i sistemi parziali. Le notazioni usate qui sopra potrebbero ingenerare qualche confusione. Quando abbiamo definito l'insieme di Gibbs mediante la formula (6.2.8), denotavamo con H l'hamiltoniana del sistema totale, costituito di N sottosistemi arbitrari. Qui invece ci stiamo riferendo a un sistema totale costituito di N sottosistemi *identici*, e quindi l'insieme di Gibbs ha la forma (6.2.8), in cui però deve apparire la hamiltoniana totale $H^{tot} = \sum_i H(z_i)$. La confusione potrebbe essere generata dal fatto che qui, per conformarci alle notazione che avevamo utilizzato nel discutere la distribuzione di MB, stiamo denotando con H la hamiltoniana di ogni singolo sottosistema, e non la hamiltoniana totale, che ora denotiamo con H^{tot}. Analogamente (a differenza della notazione usata nella formula (6.2.8), qui denotiamo con Z^{tot} la funzione di partizione del sistema totale e con Z quella di ogni singolo sottosistema (uguale a quello di tutti gli altri).

In particolare, data l'analogia formale con la distribuzione di MB, avviene ancora che il valor medio U dell'energia totale è dato dalla formula

$$U = -\partial_\beta \log Z \ .$$

Tuttavia, in virtù della forma fattorizzata della densità di Gibbs (per cui $Z = \mathscr{Z}^N$), questa prende poi la forma

$$U = -N\partial_\beta \log \mathscr{Z} \ ,$$

che coincide con la formula che si deduce dalla distribuzione di MB nel caso continuo.

Gibbs vs MB, alla luce della teoria delle grandi deviazioni. Einstein e la statistica del singolo sottosistema

NOTA DIDATTICA. Questo paragrafo deve essere riscritto, illustrando il significato della misura di Gibbs, e dando una idea della sua deduzione.

Vi è tuttavia un punto rilevante in cui la statistica di MB differisce da quella di Gibbs. Si tratta di un punto essenziale per comprendere l'argomento cruciale che oppose per tutta la sua vita Einstein a sostanzialmente tutta la comunità scientifica, per quanto riguarda l'interpretazione della MQ.

Il problema è come si debba intendere la ρ di Gibbs, se in senso forte o in senso debole. Ovvero, se essa possa essere utilizzata per definire i valori medi $\langle F \rangle$ di *tutte* le variabili dinamiche o osservabili $F(z)$ (senso forte), oppure soltanto per i valori medi di una classe ristretta di osservabili (senso debole). La classe ristretta è in effetti quella delle cosiddette funzioni di tipo somma, che sono definite come somma di un gran numero di variabili ciascuna delle quali coinvolge un singolo sottosistema.

Nel primo caso (quando assumiamo la validità dell'insieme di Gibbs, in senso forte) ci troviamo davanti a un postulato, e questo lo facciamo per decreto, come corrispondente alla scelta più semplice possibile. Nel secondo caso stiamo facendo delle ipotesi molto più generali, che sono di tipo alquanto comuni nell'ambito

dei fondamenti della teoria delle probabilità, e si trova che allora la validità del procedimento di Gibbs in senso debole (ovvero solo per valutare i valori medi delle osservabili di tipo somma) discende come un teorema, che viene dimostrato nell'ambito della teoria delle grandi deviazioni.

Ma quale è infine la differenza cruciale tra le due possibili utilizzazioni dell'insieme di Gibbs? Questo possiamo capirlo considerando l'esempio centrale che si presenta nel contesto del metodo di Boltzmann, che consiste nello studiare la successione $\{n_k\}$ di numeri di occupazione delle cellette nello spazio μ. Infatti, in tal caso ciascun numero di occupazione n_k è proprio una osservabile di tipo somma, in quanto definito dalla relazione

$$n_k = \sum_i \chi_k^i \,,$$

dove χ_k^i è la osservabile (definita sullo spazio Γ) che vale 1 se il punto i si trova nella cella k, e 0 altrimenti. Dunque, il virtù del teorema di grandi deviazioni, la formula di MB è giustificata per i numeri di occupazione, e si trova

$$n_k = N \frac{e^{-\beta\epsilon_k}}{\mathscr{Z}(\beta)} \,, \qquad \mathscr{Z}(\beta) = \sum_k e^{-\beta\epsilon_k} \,.$$

Come già abbiamo spiegato nel capitolo su Boltzmann, c'è invece la domanda differente di calcolare le probabilità p_k per l'occupazione delle celle da parte di un *singolo* sottosistema, e la domanda analoga nel caso continuo. Quello che si deduce a colpo dall'insieme di Gibbs, se lo si intende in senso forte, come significativo anche per le osservabili relative a un singolo sottosistema, è che il singolo sottosistema sia distribuito con MB in modo indipendente dagli altri sottosistemi. Naturalmente questo sembra lecito, ed è possibile che sia logicamente coerente. Ma deve essere chiaro che in ogni caso si tratta di un assioma, un assioma indipendente da tutti gli assiomi tradizionali della teoria delle probabilità. Il punto che noi facciamo, seguendo Einstein[8] (che purtroppo non aveva a disposizione i risultati della teoria delle grandi deviazioni), è che questa scelta non è affatto necessaria per la interpretazione degli esperimenti cruciali della fisica atomica che hanno condotto alla MQ.

La formula di Einstein–Boltzmann. Calore specifico e fluttuazioni di energia (all'equilibrio)

Cominciamo a stabilire la formula di Einstein–Boltzmann, e poi la commenteremo cercando di spiegare il significato che le attribuiva Einstein. Si osserva che nell'insieme canonico di Gibbs la energia totale H^{tot} è una variabile definita su tutto lo spazio delle fasi, che assume tutti i valori nel codominio che le compete

[8] Non è un caso che, nei suoi lavori del 1903–1904 sui fondamenti della meccanica statistica, Einstein si muovesse nell'ambito dello spazio Γ, e non in quello dello spazio μ.

(ad esempio nel codominio $(0,\infty)$). Fissata la temperatura inversa β, resta corrispondentemente definita una distribuzione di probabilità per l'energia totale, e in particolare resta definito un valor medio $U = U(\beta)$, che abbiamo visto come calcolare. Come di consueto in tutte le trattazioni probabilistiche, una seconda informazione (alquanto rilevante) sulla distribuzione di probabilità dell'energia è data dalla sua varianza o scarto quadratico medio σ_E^2, definiti da

$$\sigma_E^2 \stackrel{\text{def}}{=} \langle (H-U)^2 \rangle = \langle H^2 \rangle - U^2 \ . \tag{6.2.12}$$

Ora, analogamente alla formula $U = -\partial_\beta \log Z(\beta)$, si ha anche[9,10]

$$\sigma_E^2 = -\partial_\beta U \ , \tag{6.2.13}$$

o equivalentemente (ricordando $\partial_\beta = -k_B T^2 \partial_T$)

$$k_B T^2 C_V = \sigma_E^2 \ . \tag{6.2.14}$$

Capacità termica vs calore specifico. Qui si incontra il fatto, sostanzialmente irrilevante per la nostra discussione, che *il calore specifico per definizione è semplicemente la capacità termica di una mole*. Dunque, volendo fare riferimento al calore specifico anziché alla capacità termica, lasceremo sottinteso d'ora in poi che stiamo considerando il caso in cui il sistema studiato consiste di una mole.

La (6.2.14) è la formula che Einstein dimostrò nel 1904 nel suo secondo lavoro sulla termodinamica statistica, e che egli in seguito attribuì a Boltzmann. Da un punto di vista cinematico essa potrebbe apparire come una specie di curiosità. Come abbiamo già detto nel Capitolo 4 per Einstein aveva una profonda valenza dinamica. Egli ha in mente che lo sparpagliamento dei valori dell'energia del sistema sia dovuto al fatto che il sistema non è isolato, bensì in contatto dinamico con un termostato. È isolato il sistema totale, mentre l'energia del sistema di interesse fluttua al variare del tempo. Ispirandosi ai procedimenti introdotti da Kubo nel 1957, questa profonda intuizione è stata dimostrata assolutamente corretta, e lo illustreremo nel prossimo paragrafo. Si trova che queste "fluttuazioni dinamiche" coincidono con quelle date dalla formula di Einstein–Boltzmann, ma solo se si ha un opportuno rilassamento all'equilibrio. Inoltre sono possibili casi di rilassamento parziale, in cui il calore specifico si stabilizza su un valore inferiore a quello previsto all'equilibrio, dato dalla formula di Einstein–Boltzmann.

[9] Basta ricordare la definizione

$$U \equiv \langle H \rangle = \frac{\int H e^{-\beta H} \mathrm{d}z}{\int e^{-\beta H} \mathrm{d}z} \ ,$$

e usare la regola di derivazione nella forma $(f/g)' = f'/g - f g'/g^2$.

[10] La formula $U = -\partial_\beta Z(\beta)$ e la (6.2.13) rientrano in una categoria di formule ben familiari nella teoria delle probabilità. In particolare, in virtù della positivitá della varianza σ^2, la (6.2.13) implica che U è una funzione decrescente di β.

Qui sotto mostreremo come la relazione (6.2.14) tra calore specifico e varianza dell'energia viene dimostrata assumendo che calorimtro e sistema di interesse siano entrambi descritti all'equilibrio dalla misura di probabilità di Gibbs. Una trattazione analoga, ma un poco piú generale, è stata compiuta nel lavoro A. Carati, A. Maiocchi, L. Galgani, *Statistical thermodynamics for metaequilibrium or metastable states*, Meccanica (2016); (doi: 10.1007/s11012-016-0490-3), in cui si lascia del tutto imprecisata la misura di probabilità del sistema di interesse, e soltanto per il calorimetro si richiede che la misura sia quella di Gibbs. Nel lavoro di Einstein del 1911 alla conferenza Solvay è invece dato un argomento indicante che la relazione in questione ha validità del tutto generale, come una legge di natura, per sistemi macroscopici.

6.3 Il processo di misurazione del calore specifico: equilibrio e metaequilibrio

Mostriamo dunque come una generalizzazione dinamica della formula di Einstein–Boltzmann venga ottenuta con un procedimento alla Kubo, in un modello microscopico che imita un concreto processo di misurazione del calore specifico.

Il calorimetro ad acqua

È utile avere in mente come si compie la misurazione del calore specifico nel più semplice dei metodi classici, quello del calorimetro ad acqua.[11] Si porta il sistema di interesse (ad esempio un pezzo di piombo) ad una certa temperatura iniziale T_{in}, e poi lo si immerge in una quantità nota di acqua contenuta in un vaso Dewar, che si trova ad una temperatura (iniziale) inferiore T'_{in}. Allora l'acqua si scalda e dopo un certo tempo la sua temperatura si stabilizza a un valore finale T'_{fin}. A questo punto il processo di misurazione è terminato e, per definizione di equilibrio, si presume che la temperatura finale T_{fin} del sistema di interesse sia uguale a quella dell'acqua, $T_{fin} = T'_{fin}$. È noto il calore specifico dell'acqua, ovvero il calore necessario per innalzare di un grado centigrado la temperatura di una mole.[12] Allora, avendo misurato l'incremento di temperatura dell'acqua (della quale è noto il peso, compreso il contributo equivalente del calorimetro), si conosce il calore trasferito dal piombo all'acqua. Infine, dividendo per la differenza $T_{in} - T_{fin}$ tra le temperature iniziale e finale del piombo e per il suo peso, si determina il calore specifico del piombo.

Dunque le quantità che concretamente si misurano sono i due incrementi di temperatura, $T_{fin} - T_{in}$ del corpo di interesse, e $T'_{fin} - T'_{in}$ dell'apparato di misura (il calorimetro ad acqua), dove $T_{fin} = T'_{fin}$ (equilibrio finale). Consideriamo

[11] Per un esempio analogo, in cui ri riportano misure concretamente eseguite per cristalli ionici con un calorimentro a ghiaccio, si veda T. Douglas, J.C. Dever, *J. Am. Chem. Soc.* **73**, 1236 (1951).

[12] Questo venne misurato in termini di energia meccanica da Joule attorno al 1842.

il caso in cui si misuri la capacità termica di una mole del corpo di interesse, e si abbiano $\mathcal{N}'$ moli di acqua (il passaggio al caso generale è poi banale). Se C_V e C'_V sono i corrispondenti calori specifici a volume costante (capacità termiche di una mole), poiché è la medesima (in modulo) la quantità di calore scambiata, se denotiamo con Q il calore assorbito dal calorimetro avremo

$$Q = \mathcal{N}' C'_V \left(T'_{fin} - T'_{in}\right) = -C_V \left(T_{fin} - T_{in}\right) . \tag{6.3.1}$$

Per definizione il calore specifico dell'acqua è conosciuto, e dunque il calore specifico del corpo di interesse è allora data da

$$C_V = -\mathcal{N}' C'_V \frac{T'_{fin} - T'_{in}}{T_{fin} - T_{in}} . \tag{6.3.2}$$

Questo per quanto riguarda l'esperimento. Esiste tuttavia anche un procedimento teorico, che chiameremo procedimento "alla Kubo", che permette di ottenere una espressione teorica del calore specifico modellizzando in maniera dinamica il processo di misurazione sopra descritto. Si tratta di un procedimento ispirato alla teoria della risposta lineare introdotta da Kubo nel 1957, in cui ci si riduce ad un calcolo compiuto all'equilibrio, ovvero con i due sistemi alla medesima temperatura, e si considerano le *fluttuazioni* dell'energia di uno dei sistemi, uguali (a meno del segno) a quelle dell'altro sistema.

Un aspetto particolarmente rilevante della formula alla Kubo per il calore specifico è che essa fornisce uno strumento teorico per determinare in maniera dinamica il tempo in cui il processo di misurazione possa considerarsi terminato. Tale strumento, come vedremo, è la cosiddetta "funzione di autocorrelezione temporale" dell'energia scambiata.

6.4 Realizzazione matematica: il modello microscopico

Abbiamo dunque a che fare con due corpi, il sistema su cui si compie la misura (che chiamiamo anche sistema di interesse) e il calorimetro. Più in generale, consideriamo il caso di un sistema costituito da due corpi, un corpo S (il sistema di interesse) con hamiltoniana H, definita sullo spazio delle fasi $\mathcal{M}$ (le cui coordinate denoteremo globalmente con x), e un corpo S' (il calorimetro), con hamiltoniana H' definita sullo spazio delle fasi $\mathcal{M}'$, con coordinate x'. Il sistema composto $\mathsf{S}^{tot} = \mathsf{S} + \mathsf{S}'$ sarà allora descritto nello spazio delle fasi $\mathcal{M}^{tot} \stackrel{\text{def}}{=} \mathcal{M} \times \mathcal{M}'$ le cui coordinate denoteremo con $z = (x, x')$. Più avanti, l'hamiltoniana H del sistema di interesse sarà denotata semplicemente con E.

Supponiamo inizialmente i due corpi isolati termicamente. Questo implica che le energie dei due corpi sono entrambe costanti nel tempo, per cui il sistema è definito dall'hamiltoniana

$$H^{tot}(z) = H(x) + H'(x') , \quad z = (x, x') .$$

Ora, se tra i corpi realizzo un contatto termico, ciò significa che le energie dei due corpi possono cambiare, e quindi vi deve essere una certa hamiltoniana di interazione H^{int}, sicché l'hamiltoniana totale sarà

$$H^{tot}(z) = H(x) + H'(x') + H^{int}(z)\,, \quad z = (x, x')\,.$$

Dunque le due hamiltoniane H ed H' non sono più costanti del moto, cioè le energie dei due sistemi S ed S' possano variare, e si realizza in tal modo la possibilità di scambio di calore tra i due sistemi. Parliamo di calore, perché penseremo che gli scambi avvengano senza che si producano variazioni di volume dei sue sistemi, cioè senza che si compia lavoro macroscopico. Ad esempio, i due sistemi siano contenuti in due cilindri, a contatto attraverso una base comune, fissa.

Si ha dunque una hamiltoniana di interazione. Si deve tuttavia pensare che essa sia "piccola", cioè abbia valori trascurabili rispetto a quelli dei due sistemi. Infatti normalmente si assume che l'interazione termica dei corpi avvenga attraverso le loro superfici di contatto, mediante l'interazione a breve *range* delle particelle che costituiscono i corpi. Questo significa che se ad esempio il primo corpo è costituito da N particelle, avrò $H \sim N$, mentre $H_{int} \sim N^{2/3}$ (il contributo è solo superficiale), e dunque il contributo di H_{int} all'energia totale può essere trascurato (per corpi macroscopici).[13]

Dunque, in presenza dell'hamiltoniana di interazione possono aversi scambi di energia tra i due sottosistemi. In questo modo, però, il secondo principio della termodinamica non può essere verificato strettamente se si ammette la reversibilità microscopica, che definiremo subito sotto. Infatti, in corrispondenza di un certo dato iniziale $z \stackrel{\text{def}}{=} (x, x') \in \mathscr{M}^{tot}$ e del suo evoluto $\Phi^t z$ al tempo t, consideriamo l'incremento di energia (inteso in senso algebrico, cioè un incremento che può essere positivo o negativo) del sistema di interesse

$$\Delta E(z) \stackrel{\text{def}}{=} E(\Phi^t z) - E(z) \equiv H(\Phi^t z) - H(z)\,. \tag{6.4.1}$$

Lo scriveremo anche nella forma

$$\Delta E(z) \stackrel{\text{def}}{=} E_t(z) - E(z)\,,$$

se si usa la notazione già precedentemente introdotta $F_t(z) \stackrel{\text{def}}{=} F(\Phi^t z)$ per la funzione $F_t(z)$, evoluta al tempo t di una generica funzione $F(z)$.[14]

Ammettiamo ora che per un "dato iniziale" z l'incremento di energia ΔE relativo a un dato tempo t sia ad esempio negativo. Avviene allora che, per i sistemi che presentano la proprietà di reversibilità temporale, esiste il movimento "inverso" in cui l'incremento di energia (valutato per il medesimo intervallo di tempo) è esattamente uguale in modulo ed opposto di segno. Questa proprietà di reversibilità microscopica venne usata sistematicamente nell'ambito della meccanica statistica da Onsager nel 1931, nella sua trattazione della termodìnamica di nonequilibrio.

[13] Tipicamente, in opportune unità si ha $H \simeq 10^{23}$, $H_{int} \simeq 10^{16}$. e quindi $H_{int}/H \simeq 10^{-6}$.

[14] Si noti che ΔE dipende parametricamente dal tempo t di evoluzione. Questa dipendenza parametrica verrà in generale sottintesa, cioè non verrà denotata esplicitamente.

La reversibilità temporale

Ricordiamo che la proprietà di reversibilità temporale viene formulata in termini della *trasformazione di parità*, $\mathscr{P} : \mathscr{M}^{tot} \to \mathscr{M}^{tot}$, definita come quella che inverte il segno di tutte le velocità delle particelle costituenti il sistema totale. Si pensi a un sistema costituito da N particelle, e siano $q \equiv (q_1.q_2,\ldots,q_{3N})$ le corrispondenti coordinate cartesiane, e $p \equiv (p_1.p_2,\ldots,p_{3N})$ i corrispondenti momenti. Allora, se $z = (q,p)$, si definisce[15]

$$\mathscr{P} z \stackrel{\text{def}}{=} (q,-p) \quad \text{e dunque} \quad \mathscr{P}^2 = \mathscr{I} \ ,$$

dove $\mathscr{I}$ è la trasformazione identità. Allora un sistema dinamico, con una corrispondente evoluzione temporale (o "flusso") Φ^t, si dice reversibile se, per ogni dato iniziale z e per ogni tempo t, si ha $\Phi^t \mathscr{P} \Phi^t z = \mathscr{P} z$, cioè

$$\Phi^t \mathscr{P} \Phi^t = \mathscr{P} \quad \text{ovvero} \quad (\mathscr{P}\Phi^t)^2 = \mathscr{I} \ .$$

In altri termini, prendo il punto "finale" $\Phi^t z$ (evoluto al tempo t del dato iniziale z) e inverto le velocità; faccio poi evolvere ancora per il tempo t, e devo ottenere un punto che coincide con quello iniziale z a meno dell'inversione delle velocità.

È ben noto, e si dimostra immediatamente, che nei sistemi hamiltoniani la dinamica è reversibile se le hamiltoniane sono pari nelle velocità delle particelle, cioè se vale[16]

$$H(\mathscr{P} z) = H(z) \ .$$

Consideriamo hamiltoniane H ed H' (del sistema di interesse e del calorimetro) che siano pari nei momenti, sicché sarà pari nei momenti anche l'hamiltoniana totale.[17] .

Come banalissimo esercizio si vede allora immediatamente che si ha[18]

$$\Delta E(\mathscr{P}\Phi^t z) = -\Delta E(z) \ .$$

Si capisce così come, a prima vista, la reversibilità microscopica sembrerebbe costituire un ostacolo insormontabile se si vuol dedurre il secondo principio della termodinamica. È questo il cosiddetto *Paradosso di Loschmidt* (un allievo di

[15]La trasformazione $\mathscr{P} : (q,p) \to (Q,P) = (q,-p)$ è una trasformazione canonica, avente come generatrice la funzione $S = -\mathbf{q}\cdot\mathbf{Q}$. Evidentemente si ha $\mathscr{P}^{-1} = \mathscr{P}$, ovvero $\mathscr{P}^2 = \mathscr{I}$.

[16]Basta osservare che nelle equazioni in forma di Hamilton il cambiamento di segno nelle p, a secondo membro viene compensato dalla variazione di segno nella derivata rispetto al tempo nel primo membro se, insieme alla trasformazione nelle variabili p, si compie anche la trasformazione $t \to -t$.

[17]Dunque consideriamo una situazione in cui i sistemi non sono immersi in un campo magnetico esterno.

[18]Infatti

$$\Delta E(\mathscr{P}\Phi^t z) = E(\Phi^t \mathscr{P} \Phi^t z) - E(\mathscr{P}\Phi^t z) = E(z) - E(\Phi^t z) = -\Delta E(z) \ .$$

Planck), che venne formulato verso la fine dell'800 e, insieme con il *paradosso della ricorrenza di Zermelo*, veniva avanzato come obiezione apparentemente insormontabile contro le teorie cinetiche, le quali pretendevano di dedurre la termodinamica dalla dinamica microscopica reversibile. A obiezioni di questo tipo Boltzmann rispondeva con varu argomenti che sarebbe interessante analizzare.[19] La trattazione data in questo capitolo fornisce una risposta moderna. La sua caratteristica consiste, piuttosto ironicamente, nel dedurre il secondo principio proprio facendo uso essenziale della proprietà di reversibilità microscopica. Si potrebbe addirittura congetturare che non si possa dedurre il secondo principio da una dinamica microscopica che non possieda la proprietà di reversibilità temporale.[20, 21]

In ogni caso è dunque chiaro che la soluzione al paradosso di Loschmidt può ottenersi solo con metodi statistici. Si rinuncia cioè all'idea di valutare l'incremento di energia relativo a ogni singolo dato iniziale, e ci si mette invece al livello in cui si intende per incremento di energia il valore medio degli incrementi relativi ai singoli dati iniziali. In altri termini, il secondo principio verrà dedotto, ma solo in senso debole, cioè in senso statistico. È questo un punto che richiederebbe un'analisi piú approfondita, mostrando come il valor medio sia una buona stima per il valore "tipico", relativo a un dato iniziale "tipico".

Definizione meccanico–statistica del calore assorbito dal calorimetro

Dobbiamo quindi introdurre una densità di probabilità $\rho_0(z)$ dei dati iniziali nello spazio delle fasi totale. Denotando con $\langle \cdot \rangle$ il corrispondente valore medio, definiamo dunque il calore Q assorbito dal calorimentro come[22]

$$Q \stackrel{\text{def}}{=} -\langle \Delta E \rangle = -\int \Delta E(z)\, \rho_0(z) \mathrm{d}z \ . \tag{6.4.2}$$

Questa definizione è veramente naturale. Si pensi ad esempio ad un numero finito di punti iniziali z_i, ciascuno munito di una sua probabilità. Ognuno di essi "produce" una corrispondente energia assorbita dal calorimetro, $-\Delta E(z_i)$, e noi semplicemente ne prendiamo il valore medio, inteso semplicemente come media aritmetica. Ovviamente, come più volte sottolineato, tale variazione di energia viene interpretato come *calore* Q assorbito dal calorimetro (in senso algebrico, ovvero assorbito o ceduto), perché non si ha lavoro macroscopico (essendo tenute fisse le pareti).

[19] Si veda la sezione II, pag. 113 del lavoro n. 39 pag. 116 delle opere di Boltzmann, Vol. 2., o anche le *Lectures on gas theory*, parte finale.

[20] Si noti che invece Clausius si proponeva di dedurre la termodinamica ammettendo la irreversibilità già a livello microscopico. Su questo punto mette l'accento Tatiana Ehrenfest–Afanassjeva, nella sua prefazione al libro P. Ehrenfest, T. Ehrenfest–Afanassjeva, *The conceptual foundations of the statistical approach in mechanics*, Dover (New, York, 1990).

[21] Stranamente, sembra che questo fosse anche l'atteggiamento di Einstein, nel suo primo lavoro del 1903 sui fondamenti della meccanica statistica (NOTA PER GLI AUTORI: Controllare.

[22] Si ricordi che ΔE è l'increnento di energia del sistema di interesse, opposto in segno all'iincremento di energia del calorimetro. Per questo nella definizione appare il segno —.

6.5 La formula di Kubo per il calore specifico, e deduzione del secondo principio della termodinamica

Scelta della misura di probabilità iniziale

Si noti che, dalla definizione data, non risulta affatto ovvio che il calore assorbito dal calorimetro abbia un segno definito, perché nella formula la variazione di energia $\Delta E(z)$ dipendente dai dati iniziali appare alla prima potenza, mentre i dati iniziali sono suddivisi in coppie che producono il medesimo incremento in modulo, ma di segno opposto. D'altra parte appare evidente che il calore assorbito deve dipendere dalla scelta della misura di probabilità dei dati iniziali, e ad esempio esso deve essere nullo in una situazione di equilibrio, in cui i due dati iniziali della medesima coppia (che portano a incrementi uguali ma di segno opposto) devono avere lo stesso peso (la stessa densità di probabilità).

Si deve dunque stabilire quale sia la corretta misura di probabilità con cui pesare i dati iniziali in una situazione di non equilibrio[23].

È ben naturale scegliere la misura richiedendo che, prima che i due corpi siano posti in contatto termico, i dati iniziali siano distribuiti secondo una misura fattorizzata, perché i due sistemi sono indipendenti. La scelta più naturale é quella di prendere come fattori due misure di Gibbs, a temperatura inversa β per il corpo di interesse ed a temperatura inversa β' per il calorimetro, ovvero

$$\rho_0(z) = \frac{e^{-\beta H(x)}}{Z(\beta)} \frac{e^{-\beta' H'(x')}}{Z'(\beta')} , \tag{6.5.1}$$

dove Z e Z' sono le corrispondenti funzioni di partizione (fattori costanti, rispetto a z). Se ora mettiamo istantaneamente in contatto i due corpi (attraverso una base dei due cilindri in cui sono contenuti), la hamiltoniana di interazione farà variare nel tempo le loro energie. L'evoluzione in presenza della hamiltoniana di interazione, avrà l'effetto che ora variano nel tempo i valori medi delle variabili dinamiche $F(z)$.

Naturalmente, nel caso particolare in cui le due temperature siano uguali, $\beta = \beta'$, la densità di probabilità corrisponde a una misura di equilibrio per il sistema totale, e i valori medi non cambieranno col tempo. Denotando la densità di equilibrio con $\rho^{eq}(z)$ e i corrispondenti valori medi con $\langle\cdot\rangle^{eq}$, si avrà allora

$$\rho^{eq}(z) = \frac{e^{-\beta\left(H(x)+H'(x')\right)}}{Z(\beta)Z'(\beta)} , \tag{6.5.2}$$

$$\langle F\rangle^{eq} = \int F(z)\rho^{eq}(z)\mathrm{d}z ,$$

$$\langle \Delta F\rangle^{eq} = 0 .$$

[23] O meglio, gli insiemi di dati iniziali, perché nel continuo ogni dato iniziale ha misura nulla.

La forma differenziale del calore assorbito dal calorimetro

Veniamo ora al caso di una misura di probabilità iniziale di non equilibrio, $\beta \neq \beta'$. Naturalmeente, nel caso concreto che stiamo discutendo le due temperature inverse β e β' saranno quelle corrispondenti alle temperature iniziali del sistema di interesse e del calorimetro:

$$\beta \equiv \beta_{in} = 1/k_B T_{in} \ , \quad \beta' \equiv \beta'_{in} = 1/k_B T'_{in} \ .$$

Rimandando la discussione[24] del caso generale, consideriamo il caso di piccola differenza di temperatura iniziale, $T_{in} = T'_{in} + \mathrm{d}T_{in}$, sicché stiamo denotando

$$\mathrm{d}T_{in} = T_{in} - T'_{in} \ .$$

Corrispondentemente si avrà $\beta = \beta' + \mathrm{d}\beta$.

Valutiamo dunque la forma differenziale δQ del calore, che si ottiene dalla (6.4.2) per il calore Q assorbito dal calorimentro, sviluppandola al primo ordine in $\mathrm{d}\beta$ attorno a $\beta = \beta'$. Ricordando $\langle \Delta E \rangle^{eq} = 0$ si ha

$$\delta Q = -\,\mathrm{d}\beta \int \Delta E(z) \frac{\partial \rho_0}{\partial \beta}|_{\beta=\beta'}(z)\mathrm{d}z \ , \tag{6.5.3}$$

ovvero, [25] in virtù della dipendenza esponenziale di ρ_0 da β,

$$\delta Q = \mathrm{d}\beta \int \Delta E(z) H(z) \rho^{eq}(z)\mathrm{d}z \ . \tag{6.5.4}$$

Un astuto cambiamento di variabile di integrazione

Operiamo ora, nell'integrale (6.5.4), un astuto cambiamento di variabile suggerito dalla reversibilità temporale del flusso, ovvero operiamo la sostituzione

$$z = \mathscr{P} \Phi^t y \ .$$

Utilizziamo ora le seguenti proprietà

- la trasformazione $y \to z$ è canonica (come composizione di due trasformazioni canoniche, l'evoluzione temporale a tempo fissato, e la parità $\mathscr{P}$), sicché
$$\mathrm{d}z = \mathrm{d}y \ ;$$
- si ha
$$\Delta E(\Phi^t y) = -\Delta E(y)$$
come abbiamo visto poco sopra

[24] NOTA PER GLI AUTORI. PARTE NON ANCORA SCRITTA.

[25] Si osservi che sembrerebbe esserci un altro termine, che si ottiene derivando la funzione di partizione $Z(\beta)$. Ma tale termine è proporzionale a $\langle \Delta E \rangle^{eq}$, che è nullo in virtù della proprietà delle misure di equilibrio: Si ha $\langle E_t \rangle^{eq} = \langle E \rangle^{eq}$, ovvero $\langle \Delta E \rangle^{eq} = 0$

- la misura di equilibrio è invariante sotto evoluzione temporale, ovvero si ha
$$\rho^{eq}(\mathscr{P}\Phi^t y) = \rho^{eq}(\Phi^t y) = \rho^{eq}(y)$$
(la prima uguaglianza segue dalla assunta invarianza delle hamiltoniane parziali rispetto alla trasformazione di parità)

Si ha allora per δQ una seconda espressione, data da

$$\delta Q = -\,\mathrm{d}\beta \int \Delta E(y) H(\Phi^t y) \rho^{eq}(y) \mathrm{d}y \ . \tag{6.5.5}$$

Dunque, prendendo per δQ la semisomma delle due espressioni trovate (e naturalmente usando il fatto che la variabile "muta" y può essere denotata con z), si trova la formula
$$\delta Q = -\mathrm{d}\beta \frac{1}{2} \int (\Delta E(z))^2 \rho^{eq}(z) \mathrm{d}z \ ,$$
che scriviamo anche
$$\delta Q = -\frac{1}{2} \langle (\Delta E(z))^2 \rangle^{eq} \mathrm{d}\beta \ . \tag{6.5.6}$$
Equivalentemente ricordando $\mathrm{d}\beta = -\,\mathrm{d}T_{in}/k_B T^2$ (e sottintendendo la variabile corrente z in $\Delta E(z)$), si ha

$$\delta Q = \frac{1}{2} \frac{\langle (\Delta E)^2 \rangle^{eq}}{k_B T^2} \mathrm{d}T_{in} \ . \tag{6.5.7}$$

Deduzione microscopica del secondo principio (in virtù della reversibilità temporale)

La relazione (6.5.7) mostra che il calore assorbito dal calorimetro ha manifestamente un segno ben determinato (a differenza della espressione corrispondente alla originaria definizione, che coinvolgeva l'incremento $\Delta E(z)$ alla prima potenza, e non al quadrato). In particolare essa mostra anche che il calore assorbito dal calorimentro ha proprio il segno di $\mathrm{d}T_{in} = T_{in} - T'_{in}$, e dunque il calore fluisce spontaneamente (se non si agisce con altri mezzi sul sistema totale) dal corpo più caldo al corpo più freddo, come afferma il secondo principio della termodinamica nella forma di Clausius.[26] Dunque abbiamo dedotto il secondo principio (irreversibililtà macroscoppica) *in vitrù della reversibilità microspopica*, e non, come spesso viene detto. *nonostante* la reversibilità microscopica (al punto che si potrebbe ritenere che non sia possibile dedurre il secondo principio nel caso

[26]Se inizialmente il sistema di interesse è a temperatura superiore a quella del calorimetro, allora il calorimetro assorbe calore. In effetti, secondo Fermi si veda (E. Fermi, Termodinamica), questa costituisce una forma debole del secondo principio, nel senso che il sistema presenta una irreversibiltá macroscopica.

di presenza di attrito[27]). La deducibilità dalla irreversibilità macroscopica della reversibiltà microscopica fu messa in luce da Onsager nei suoi lavori attorno all'anno 1931 sulla termodinamica di non equilibrio.

Questa dimostrazione può essere estesa al caso in cui sia finita, anziché infinitesima, la differenza $T_{in} - T'_{in}$ delle temperature iniziali.

La formula "alla" Kubo per il calore specifico (caso del termosstato)

La relazione (6.5.7) fornisce anche direttamente la formula analitica per il calore specifico, almeno nel caso in cui il calorimetro sia un ternostato, cioè sia tanto grande da garantire cke la sua temperatura non cambi a seguito dell'assorbimento del calore ceduto dal sistema studiato, ovvero si abbia $T'_{fin} = T'_{in}$. Infatti, ricordando che in ogni caso si ha anche $T_{fin} = T'_{fin}$ (all'equilibrio le temperature finali del calorimento e del sistema di interesse sono uguali), nel caso del termostato si ha in definitiva

$$T_{in} - T'_{in} = T_{in} - T'_{fin} = T_{in} - T_{fin} \ ,$$

e quindi $\mathrm{d}T_{in}$ viene a rappresentare (a meno del segno) anche l'incremento della temperatura che si produce nel sistema di interesse a seguito del processo di misurazione. Per definizione di calore specifico abbiamo dunque, dalla (6.5.7),

$$C_V = \frac{1}{2}\,\frac{\langle(\Delta E)^2\rangle^{eq}}{k_B T^2} \qquad \text{caso del} \quad \text{termostato} \ . \tag{6.5.8}$$

Complementi: l'espressione del calore specifico nel caso di un calorimetro generico.

È poi un semplice esercizio ricavare la formula del calore specifico nel caso generale, di un calorimentro con capacità termica $\mathcal{N}'C'_V$ finita. Il problema che si presenta è che la formula (6.3.1) contiene differenze tra temperature finali ed iniziali del sistema di interesse e del calorimentro, ovvero $T_{fin} - T_{in}$ e $T'_{fin} - T'_{in}$, mentre la formula teorica (6.5.7) da noi ottenuta per δQ fa riferimento alla differenza $\mathrm{d}T_{in} = T_{in} - T'_{in}$ tra le temperature *iniziali* del sistema di interesse e del calorimetro.

Usiamo allora il fatto che, in virtù di $T'_{fin} = T_{fin}$, si ha

$$\left(T'_{fin} - T'_{in}\right) - \left(T_{fin} - T_{in}\right) = T_{in_} T'_{in} \equiv \mathrm{d}T_{in} \ . \tag{6.5.9}$$

Riscriviamo poi la (6.3.1), con δQ invece di Q, mettendo in evidenza le differenze di temperatura del calorimetro e del sistema di interesse: ovvero

$$T'_{fin} - T'_{in} = -\frac{\delta Q}{\mathcal{N}'C'_V} \ , \qquad T_{fin} - T_{in} = \frac{\delta Q}{C_V} \ . \tag{6.5.10}$$

[27]NOTA PER GLI AUTORI. Illustrare l'esempio della particella in campo magnetico assegnato

Sottraiamo poi queste due relazioni e denotiamo

$$\delta Q = \alpha \mathrm{d}T_{in}$$

avendo posto

$$\alpha = \frac{1}{2}\,\frac{\langle(\Delta E)^2\rangle^{eq}}{k_B T^2}\,. \tag{6.5.11}$$

Cancellando il fattore comune $\mathrm{d}T_{in}$, si ha allora

$$\alpha\left(\frac{1}{C_V}+\frac{1}{\mathscr{N}' C'_V}\right)=1\,,$$

e infine

$$C_V = \mathscr{N}' C'_V\,\frac{\alpha}{\mathscr{N}' C'_V - \alpha}\,, \tag{6.5.12}$$

che evidentemente nel caso del termostato $\mathscr{N}' \gg 1$ si riduce alla formula (6.5.8) precedentemente trovata.

6.6 Calore specifico e dinamica. La formula "alla Kubo" e le correlazioni temporali. Generalizzazione della formula di Einstein–Boltzmann

Consideriamo ora il caso del termostato. che è quello cui faceva riferimento Einstein quando introdusse la sua formula esprimente il calore specifico in termini delle fluttuazioni di energia (la sua varianza) valutate nell'insieme canonico. La via per la rielaborazione della formula per il calore specifico, che ci permetterà di constatare se il processo di misurazione è terminato, si apre quando si scrive in maniera più esplicita l'espressione che abbiamo appena trovato, mettendola nella forma

$$k_B T^2\, C_V = \frac{1}{2}\left\langle\left(E_t - E\right)^2(z)\right\rangle^{eq}\,, \tag{6.6.1}$$

in cui figurano esplicitamente sia la quantità E, sia la sua evoluta temporale E_t, sicché nel problema entra ora esplicitamente, in maniera formale, la dinamica.

Sviluppiamo ora il quadrato. Osserviamo che, come consueta proprietà per i valori medi all'equilibrio, si ha $\left\langle E_t^2(z)\right\rangle^{eq} = \left\langle E^2(z)\right\rangle^{eq}$, e aggiungiamo e sottraiamo il termine U^2 dove $U \equiv \langle E(z)\rangle^{eq}$. Otteniamo allora

$$k_B T^2 C_V = \sigma_E^2 - \mathscr{C}_E(t)\,. \tag{6.6.2}$$

dove σ_E^2 è la varianza di E all'equilibrio, mentre

$$\mathscr{C}_E(t) \stackrel{\text{def}}{=} \left\langle E_t E\right\rangle^{eq} - U^2\,, \quad (U = \left\langle E\right\rangle^{eq})\,, \tag{6.6.3}$$

è la cosiddetta *autocorrelazione temporale* dell'energia E, dove la media è ancora calcolata all'equilibrio.

Per comprendere il significato di questa definizione, bisogna ricordare che per ogni coppia di variabili aleatorie o casuali (*random variables*) f, g su uno spazio di probabilità, se si denota il valor medio con $\langle \cdot \rangle$, la corrispondente correlazione $\mathscr{C}_{f,g}$ è definita da

$$\mathscr{C}_{f,g} = \langle f\, g \rangle - \langle f \rangle \langle g \rangle ,$$

ovvero come la differenza tra valor medio del prodotto, e prodotto dei valori medi. La differenza da zero della correlazione è in qualche modo una misura della loro indipendenza, perché la correlazione di due variabili indipendenti è nulla. Qui le due variabili sono l'energia E_t al tempo t e quella, E, al tempo iniziale;[28] quindi si parla di **autocorrelazione temporale** dell'energia del sistema di interesse. Si intuisce immediatamente che in un moto sufficientemente caotico l'autocorrelazione temporale di ogni variabile dinamica tende a zero. Ed in effetti, i sistemi dinamici detti di tipo *mixing* (ovvero, mescolanti) vengono definiti proprio come quelli che hanno tale proprietà (le correlazioni temporali di ogni coppia di funzioni tende a zero per $t \to +\infty$).[29]

La formula (6.6.2), che generalizza quella di Einstein–Boltzmann, è davvero alquanto rilevante, perché mostra che il calore specifico misurato si riduce a quello di equilibrio se il sistema è abbastanza caotico da far sì che l'autocorrelazione temporale dell'energia del sistema su cui si esegue la misura si annulli per tempi sufficientemente lunghi. Il tempo dopo il quale la correlazione è praticamente nulla viene allora identificato con il tempo di rilassamento, o il tempo caratteristico per il compiersi della misura.

Ma potrebbe presentarsi anche un fenomeno diverso, ovvero che al crescere del tempo l'autocorrelazione temporale, invece di tendere a zero, tenda a un valore non nullo in maniera tale che fino a tempi molto più lunghi non si manifesti alcuna tendenza a un successivo rilassamento. In tal caso l'osservatore giudicherebbe di trovarsi a tutti gli effetti in una situazione di equilibrio (o di almeno pratico equilibrio), ma il calore specifico presenterebbe un valore evidentemente diverso da quello di equilibrio (e presumibilmente inferiore). Almeno qualitativamente ci si troverebbe allora in una situazione analoga a quella che si presenta all'equilibrio secondo la meccanica statistica quantistica. Si tratta in effetti di una situazione simile a quella che si presenta nei vetri, in cui il sistema si trova in una situazione di apparente equilibrio. Esso si trova a tutti gli effetti in un ben definito stato termodinamico, ma presentando valori delle quantità termodinamiche (calore specifico etc.) diverse dai valori di equilibrio. La possibilità di questa analogia, nell'ambito del problema FPU, fu messa in luce per la prima volta nel 1982 da un gruppo di studiosi attorno a Parisi.[30]

Si noti che, mentre il calore specifico di equilibrio è una proprietà che dipende solo dalla hamiltoniana del sistema di interesse, essendo del tutto indipendente dalla hamiltoniana di interazione con l'apparato di misura (il calorimetro), invece il tempo di misurazione

[28] Si ricordi $\langle E_t \rangle^{eq} = \langle E \rangle^{eq} = U$.

[29] Dunque la sola proprietà di ergodicità, discussa nel corso di Meccanica Analitica 2, non basta per recuperare i valori canonici per tutte le quantità termodinamiche. Sono necessarie proprietà più forti di tipo *mixing*, anch'esse discusse in quel corso.

[30] Fucito et al, 1982

dipende in maniera essenziale dall'hamiltoniana di interazione. In generale si presume che tale tempo sia piccolo o almeno ben definito, e che quindi se ne possa prescindere. Naturalmente, diventa allora un problema matematico della teoria dei sistemi dinamici classici determinare se questo rilassamento avvenga davvero (e, in caso affermativo, dopo quanto tempo) per un dato sistema in interazione con un apparato di misura. Da un punto di vista concettuale, tale punto è stato discusso ad esempio nel lavoro di Birge e Nagel, dal titolo significativo *"Observation of time dependent specific heat"*.[31] Il problema è particolarmente acuto nel caso dei vetri, in cui il rilassamento richiede tempi geologici.

Osservazione. Energia posseduta vs energia scambiata, in termodinamica e in meccanica statistica. Significato del risultato "alla Kubo".

Ritorniamo alle origini. Abbiamo visto come la teoria di Clausius mostra che, almeno per un gas diluito, la temperatura è proporzionale al valor medio di una quantità meccanica del sistema, ovvero la somma delle energie cinetiche dei baricentri. Dunque un incremento di temperatura corrisponde a un incremento di energia cinetica dei baricentri. E poiché in meccanica esiste anche l'energia del sistema, si potrebbe congetturare che l'energia meccanica (media) del sistema possa essere identificata con l'energia termodinamica U del sistema, che è la quantità fenomenologica definita in virtù del primo principio.

Ma questa affermazione è fuorviante, perché la termodinamica non parla assolutamente mai dell'energia di un sistema, ma parla invece dell'energia, diciamola ΔU che, durante il passaggio (o la trasformazione, come si dice) da uno stato a un altro, il sistema considerato scambia con un altro sistema. Ad esempio, nelle trasformazioni di stato, tipicamente nel caso liquido–gas, nel passaggio dallo stato liquido allo stato gassoso si ha l'energia ΔU di vaporizzazione (alla temperatura di ebollizione, dipendente parametricamente dalla pressione), e la quantità opposta nella trasformazione inversa di liquefazione. Questa è la quantità che viene misurata, misurando ad esempio quanto ghiaccio viene sciolto nel passaggio di stato.[32] Analogamente si misura l'energia libera ΔF prodotta o assorbita in una trasformazione. La termodinamica non ci dice nulla invece su una eventuale energia U o una energia libera F che il sistema eventualmente "possiede". Tutta la termodinamica viene costruita sulla base di queste quantità scambiate, e possiamo parlare ad esempio di una energia U solo se convenzionalmente assegnamo un valore di U, diciamo U_0, a uno "stato di riferimento", e misuriamo (ad esempio con un calorimetro) il calore scambiato nel passaggio tra i due stati.[33]

[31] N.O. Birge, *Phys. Rev.* B **34**, 1631 (1986); N.O. Birge, S.R. Nagel, *Phys. Rev. Lett* **25**, 2674 (1985). Si veda anche A. Carati, L. Galgani, *Europhys. Lett.* **74**, 528 (2006).

[32] Spesso non si tiene presente che queste energie sono enormi. Per fare evaporare l'acqua occorrono 540 kilocalorie per chilo, ovvero un'energia grossomodo cinque volte superiore a quella richiesta per portate un litro di acqua da 0 a 100 gradi centigradi (a pressione ordinaria). Ci si rende conto facilmente che l'energia di condensazione liberata da una grande nuvola in un temporale è dell'ordine dell'energia liberata dall'esplosione di una bomba atomica.

[33] Questi concetti sono illustrati in maniera limpidissima nel libro di Nernst sul terzo principio, W. Nernst, *The new heat theorem*.

Ad esempio, nella fisica atomica, quando si parla dell'energia di un atomo, si prescinde dall'energia che possiede il suo nucleo, costituito dai suoi protoni e neutroni (i nucleoni). Pare che ai nucleoni si debba attribuire una temperatura enorme, dell'ordine di 10^{14} K, ma questa energia non viene scambiata con i gradi di libertà atomici (energie dei baricentri o energie delle rotazioni e delle oscillazioni molecolari) o con gli elettroni, e quindi viene del tutto trascurata nella fisica atomica. I gradi di libertà atomici e quelli nucleari costituiscono, dal punto di vista termodinamico, due sistemi indipendenti, che non comunicano. Potrebbero forse comunicare su scale di tempi dell'ordine dell'età dell'universo. Analogamente, anche le molecole costituenti i vetri non "termalizzano" completamente con l'ambiente, e restano in uno stato di metaequilibrio fino a scale di tempi di ordine geologico.

Tuttavia, nei manuali di meccanica statistica si valuta sempre l'energia interna U di un sistema come media dell'energia meccanica e quindi equiparando l'energia posseduta con l'energia scambiabile.[34] Questa tradizione in effetti risale a Boltzmann e a Gibbs (anche se Boltzmann manifestò pubblicamente tutti i suoi dubbi sulla sua validità). Essa fallisce gravemente nelle trattazioni classiche della fisica atomica, e funziona sostanzialmente molto bene nelle corrispondenti trattazioni quantistiche. Ma non sempre.[35] [36]

[34] Una notevole eccezione è il manuale di G.H. Wannier, *Statistical Physics*, Dover (New York, 1966). Naturalmente, questo concetto, che si debba sempre fare riferimento all'energia scambiata, è ritenuto praticamente ovvio da uno scienziato del calibro di Nernst, come si vede sia nel suo manuale che già abbiamo citato, sia nel suo fondamentale libro sul terzo principio(W. Nernst, *The new heat theorem*, Dover.)

[35] Si tratta del problema dell'orto-idrogeno e del para-idrogeno.

[36] NOTA PER GLI AUTORI. Nel manosacritto esiste un commento da chiarire.