

Appunti di fisica statistica

Tommaso Miliani

20-11-25

1 Spazio delle fasi, macrostati e microstati

1.1 Macrostat e microstat

A livello macroscopico si può definire un sistema termodinamico di un fluido attraverso poche coordinate termodinamiche, ossia pressione, volume e temperatura (**Macrostat**). Per quanto riguarda lo stato microscopico, si devono considerare i vettori posizione delle particelle e i vettori velocità, inoltre il numero di particelle N è dell'ordine del numero di Avogadro. Posso dunque dire che la descrizione di uno stato microscopico è di natura statistica e dunque per ogni stato corrisponde un valore della probabilità. Dunque la funzione probabilità dipende sia dal tempo che dai vettori posizioni e dai vettori velocità.

$$\rho(\vec{r}_1, \dots, \vec{r}_n, \vec{v}_1, \dots, \vec{v}_n, t) d^3\vec{r}_1 \dots d^3\vec{r}_n d^3\vec{v}_1 \dots d^3\vec{v}_n$$

Questa funzione è dunque una funzione simile alla funzione densità di probabilità: lo stato microscopico dunque è definito da questa funzione e prende il nome di **microstat**. Il differenziale d^3 vuol dire che devo differenziare rispetto alle tre coordinate spaziali, mentre non dipende dal differenziale temporale.

1.2 Spazio delle fasi

Lo **spazio delle fasi** è quello spazio che è definito da tutti i vettori posizione e dunque ha dimensione $6N$. La funzione di probabilità è dunque una funzione che è definita a partire di questo spazio nel seguente modo:

$$\rho : F \times \mathbb{R} \rightarrow \mathbb{R} \quad (1)$$

Ossia la funzione distribuzione di probabilità ρ è definita a partire dallo spazio delle fasi F e dallo spazio cartesiano. La fisica statistica vuole trovare il macrostat a partire dal microstat. Il problema di determinare la distribuzione di probabilità delle variabili microscopiche e associarle alle variabili macroscopiche è un problema totalmente risolto per l'equilibrio termodinamico secondo la **meccanica statistica all'equilibrio**, secondo la quale si ha un modo univoco per determinare la funzione di probabilità; fuori dall'equilibrio non si riesce ancora a determinare la suddetta funzione. Le ipotesi per determinare ρ sono le seguenti:

1. **Equilibrio termodinamico:** per cui ρ non dipende da t (si riducono le $6N + 1$ variabili a $6N$)
2. All'equilibrio il sistema è omogeneo e dunque la densità numerica è omogenea. Questo significa che la probabilità di trovare un atomo non dipende dalla posizione e dunque la funzione ρ ha solamente $3N$ variabili poiché non si considerano le posizioni.
3. Tutte le velocità degli atomi siano indipendenti ed identicamente distribuite, dunque la funzione ρ non è altro che il prodotto della densità di probabilità delle singole variabili (indipendenza statistica) e dunque

$$\rho(\vec{v}_1, \dots, \vec{v}_n) = \prod f(\vec{v}_i)$$

4. Inoltre la funzione $f(\vec{v})$ non dipende dalle tre componenti ma solamente dalla loro somma in quadratura:

$$f(\vec{v}) = f(v_x)f(v_y)f(v_z) \implies f(\vec{v}) = f(v^2)$$

Esiste una scala minima nello spazio delle fasi al di sotto della quale non si può scendere: se esistesse uno spazio delle fasi molto piccolo nel quale le variabili sono distribuite ugualmente, non posso dire che una di queste sia più probabile delle altre. Posso dunque introdurre un cubetto che abbia lato $(\Delta x \Delta v)^{3N}$ all'interno del quale la funzione ρ diventa costante e non riesce a variare. La meccanica quantistica ci pone un limite

per cui l'incertezza non può essere resa piccola a prescindere sia per la posizione che per la velocità. Se esiste un volumetto elementare con dimensione diversa da zero, si può pensare di poter "cubettare" lo spazio delle funzioni N dimensionale e dunque, grazie all'ipotesi che si è fatto, ρ sia costante e molto simile tra i vari volumetti. Questo vuol dire che, dato che sono un insieme discreto e numerabile, esiste allora una collezione discreta di stati microscopici del nostro sistema: invece di considerare la funzione densità di probabilità con $6N$ variabili, si può considerare un insieme discreto di M variabili con M il numero di stati macroscopici, per cui la distribuzione fatta dagli stati p_1, \dots, p_M è discreta. Allora la probabilità dello stato a -esimo non è altro che

$$P_a = \rho(\vec{r}_1, \dots, \vec{r}_n, \vec{v}_1, \dots, \vec{v}_n) \quad \vec{r}_1, \dots, \vec{r}_n, \vec{v}_1, \dots, \vec{v}_n \in (\Delta x \Delta v)^{3N}$$

La procedura di partire da una situazione continua ad una discreta prende il nome di **Coarse graining** e la probabilità da continua a discreta prende il nome di **coarse-grained** simile alle fotografie digitali, anche se a livello pratico il limite di "risoluzione" di questa tecnica lo dà la meccanica quantistica. Ci sono tantissimi microstati che corrispondono allo stesso stato macroscopico, dunque esiste una relazione asimmetrica per cui esistono infiniti valori di velocità per cui

$$p = p(\langle v^2 \rangle) \quad T = T(\langle v^2 \rangle)$$

Il fatto di aver ipotizzato di poter scrivere la probabilità come funzione dello spazio delle fasi è già una approssimazione, dunque se conoscessi con certezza le posizioni e le velocità al variare del tempo, si avrebbe che una certa coordinata dello spazio delle fasi è esattamente uno: non può essere descritta da una funzione ma da una distribuzione: nel caso in cui so risolvere la dinamica del sistema e conosco con infinita precisione i dati iniziali, la funzione ρ non sarebbe una funzione ma una distribuzione. Assumere dunque che la probabilità si possa scrivere come funzione di probabilità è anche questa una approssimazione grossolana.

2 Determinare la quantità di informazione nella distribuzione

Assumendo di aver già fatto coarse graining per la nostra funzione di probabilità, e presa una collezione di valori per la probabilità, il valore della probabilità di uno di questi N stati è massima mentre la probabilità per tutti gli altri stati è nulla: dunque siamo nella situazione di informazione più grande possibile: la distribuzione di probabilità è dunque certa. Si può dunque provare a determinare la quantità di informazione all'interno della nostra distribuzione attraverso la funzione che quantifica la mancanza di informazione:

$$H = - \sum_{a=1}^M p_a \ln p_a = \left\langle \ln \frac{1}{p} \right\rangle \quad (2)$$

Questa funzione si esprime anche in funzione del valore atteso della probabilità: il fatto che ci sia $\frac{1}{p}$ già questo di per sé mi permette di determinare la quantità di "ignoranza" che ho sul sistema. Si può dimostrare che, se questa funzione è una buona approssimazione per la mancanza di informazione, allora deve essere massima quando non si ha tanta informazione e minima quando invece se ha tanta.

$$p_k = 1 \quad p_a = 0 \quad \forall a \neq k$$

Ossia posso tirare fuori dalla somma il caso particolare p_k :

$$H = -p_k \ln p_k \sum_{a \neq k}^M p_a \ln p_a = -1 \ln 1 \sum_{a \neq k}^M p_a \ln p_a = 0$$

Che fa zero per il limite notevole. Dunque

$$H = 0 \iff p_k = 1 \quad p_a = 0 \quad \forall a \neq k$$

Dato che questi numeri p_a sono tutti $0 < p_a < 1$, allora ogni addendo della sommatoria è positivo, dunque la funzione $H \geq 0$, e dato che si è trovato un valore per cui $H = 0$, allora si è trovato un minimo per la funzione H . Bisogna invece dimostrare che la situazione nella quale in cui tutte le probabilità sono uguali allora H è massima. Si deve dunque trovare un estremo della funzione in modo tale che le n variabili si sommino ad 1. Utilizzando i moltiplicatori di Lagrange, posso ottenere

$$F = - \sum_{a=1}^M p_a \ln p_a + \lambda \left(\sum_{a=1}^M p_a - 1 \right)$$

E vado dunque ad annullare il gradiente della funzione:

$$\frac{\partial F}{\partial p_k} = 0 \quad \forall k = 1 \quad \frac{\partial F}{\partial \lambda} = 0$$

Allora posso fare

$$\begin{aligned} \frac{\partial F}{\partial p_k} &= -(\ln p_k + 1) + \lambda \quad \forall k \\ \frac{\partial F}{\partial \lambda} &= \sum_{a=1}^M p_a - 1 \end{aligned}$$

Imponendole uguali a zero:

$$\begin{aligned} \ln p_k &= (-1 + \lambda) \quad \forall k \\ \sum_{a=1}^M p_a &= 1 \end{aligned}$$

A questo punto la prima mi dà informazioni secondo la quale tutte le p_k hanno un dato valore p e la seconda invece mi dice che sono esattamente uguali:

$$p_k = p \quad \forall k \implies p_k = \frac{1}{M}$$

Con le derivate sconde si può ottenere il caso per il quale

$$\frac{\partial^2 H}{\partial p_k^2} = -\frac{1}{p_k}$$

E dunque, dato che le derivate sono massime, si è dimostrato che ha un massimo e dunque si ha poca informazione sul sistema.

$$H_{MAX} = -\sum_{a=1}^M \frac{1}{M} \ln \frac{1}{M} = \frac{1}{M} \sum_{a=1}^M \ln M = \ln M$$

Il valore massimo della funzione è dunque il logaritmo del numero di stati che può assumere il sistema. Mi aspetto anche una situazione intermedia per la quale si ha una situazione di probabilità intermedia: immaginando di avere una biblioteca con un certo numero di libri, questi libri possono essere ordinati sicuramente secondo un certo criterio. Alla fine si arriva a dire che si ha l'informazione massima possibile: i modi in cui posso disporre N libri in M possibili modi è esattamente uguale al numero di permutazioni: $M = N!$. Supponendo di avere libri di tipi diversi (ordinati per categorie ma comunque disordinati nelle loro categorie), allora il numero di combinazioni ordinate possibili M deve ridursi notevolmente per cui

$$H = \ln(N_1! \dots N_k!)$$

Ed è dunque sempre minore di:

$$H_{MAX} = \ln((N_1 + \dots + N_k)!)$$

Dove M è il numero di stati complessivi del sistema e k è il numero di sottostati. Dietro a questa considerazione c'è il teorema di Sheldon.

Teorema 2.1 (Teorema di Sheldon).

La funzione H è l'unica funzione unica a meno di una costante moltiplicativa che ha le seguenti proprietà (in realtà appartiene ad una famiglia di funzioni):

1. Fa zero sulla distribuzione certa
2. Massima sulla distribuzione uniforme
3. È additiva per distribuzioni indipendenti
4. È monotona crescente in M

Dimostrazione. No dimostrazione.

□

Si può dimostrare che è però additiva per la probabilità: per una certa distribuzione di probabilità la probabilità per la distribuzione congiunta H è data da:

$$\begin{aligned} \{p_a^{(1)}\} & \quad a = 1, \dots, M_1 \\ \{p_b^{(2)}\} & \quad a = 1, \dots, M_2 \\ \prod_{ab} &= p_a^{(1)} p_b^{(2)} \end{aligned}$$

$$H_{a+b} = - \sum_{a=1}^{M_1} \sum_{b=1}^{M_2} \prod_{ab} \ln \prod_{ab}$$

Che diventa con semplici passaggi

$$= - \sum_{a=1}^{M_1} \sum_{b=1}^{M_2} p_a^{(1)} p_b^{(2)} \ln p_a^{(1)} - \sum_{a=1}^{M_1} \sum_{b=1}^{M_2} p_a^{(1)} p_b^{(2)} \ln p_b^{(2)}$$

Dato che le distribuzioni sono normalizzate e, scomponendo le sommatorie, si ottiene il risultato

$$- \sum_{a=1}^{M_1} p_a^{(1)} \ln p_a^{(1)} - \sum_{b=1}^{M_2} p_b^{(2)} \ln p_b^{(2)} = H_1 + H_2$$

E dunque è additiva.