

MODULO B : METODI DI SIMULAZIONE IN MECCANICA STATISTICA

Usando i "grandi" calcolatori è possibile risolvere numericamente, e problema a N corpi (classici e quantistici, fermioni a parte) con una precisione controllabile e soddisfacente e con un numero di particelle $N \leq 10^4 - 10^5$ sufficiente nella stragrande maggioranza delle situazioni per rappresentare il limite termodinamico.

I metodi numerici possibili rientrano in due classi:

- A) DINAMICA MOLECOLARE + Consiste nella integrazione numerica delle eq. del moto. L'informazione statistica si ottiene prendendo le medie temporali di opportune variabili dinamiche: è la realizzazione pratica del programma di Boltzmann (1959 - 1964)
- B) IL METODO MONTE CARLO - Con esso si calcolano direttamente le medie sull'ensemble di Gibbs. Ciò si ottiene introducendo un tempo ed un movimento fittizio (ossia una evoluzione fittizia) del sistema che permette di realizzare un campionamento di configurazioni del sistema nello spazio delle configurazioni con probabilità proporzionale a $e^{-\beta V}$. (1952)

Entrambi i metodi richiedono di fissare: α) la legge di evoluzione β) le condizioni al contorno; γ) le condizioni iniziali.

~~Per~~ A seconda della situazione fisica in esame la scelta delle suddette condizioni può essere variata. Vediamo ~~come~~ qual'è però il caso standard (e storicamente venuto prima). Il punto α) lo metteremo in seguito perché è specifico del metodo; vediamo gli altri.

β) CONDIZIONI AL CONFINAMENTO

Consideriamo un sistema di N ($100 \div 10000$) particelle in un volume V fisso, generalmente ^{ov} cubo $\square$ di lato L .

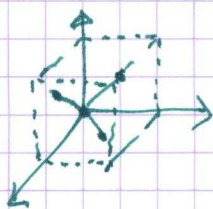
Le condizioni al bordo rappresentano l'azione del recipiente

del sistema. Si tratta di un effetto di superficie che deve essere minimizzato se il nostro scopo è quello di ottenere comportamenti rappresentativi del limite termodinamico. In questa ottica, alle pareti riflettenti (che sembra essere la scelta naturale), si aggiungono le condizioni periodiche: si aggiungono al sistema considerato tutte le sue immagini ottenute traslando il sistema in tutte le direzioni secondo un qualunque multiplo di L .

Stime teoriche e/o ottenute dalla simulazione mostrano che le fluttuazioni termodinamiche intorno alle quantità medie sono di ordine $(1/N)$ con un coefficiente dell'ordine dell'unità. Questo significa che medie ottenute dalla simulazione di $N \sim 1000$ particelle differiscono dal corrispondente limite termodinamico per non più dell'1% (non è vero per fermioni in stato fondamentale...)
T & T_F VEDI RETRO →

1) CONDIZIONI INIZIALI

1- L'informazione statistica che vogliamo ottenere dall'evoluzione (reale DM o Fittoria MC) del sistema deve essere indipendente dalle condizioni iniziali (sistemi caotici e/o ergodici). Tuttavia per definire una evoluzione occorre fissarle. ~~Però~~ Si ha un'ampia arbitrarietà in questa scelta ma, se si vuole che il sistema raggiunga l'equilibrio in un tempo breve conviene scegliere le più regolari possibili. Una buona scelta è horizonare le palle nei nodi di un reticolo di Bravais, e in particolare del reticolo in cui cristallizza la sostanza in esame. Per esempio nel sistema LJ (Argon) il reticolo ~~è~~ naturale è fcc con 4 atomi per cella. ~~Però~~



Se la densità da simulare è n allora
il lato della cella elementare è a allora
si ha:
$$n = \frac{4}{a^3} \Rightarrow a = \sqrt[3]{\frac{4}{n}}$$

• Una volta scelte le condizioni periodiche al bordo il sistema è diventato periodico a tutti gli effetti. Vedremo successivamente come calcolare il valore dell'energia in un tale sistema.

Qui vogliamo discutere alcune limitazioni introdotte dalle PBC.

- In particolare la periodicità del sistema limita i modi ~~per~~ collettivi nel sistema a quei modi compatibili con la scatola.

Come in M.Q. lo spettro dei modi diventa discreto e i vettori $\underline{k}$

permissibili sono solo: $\underline{k} = \frac{2\pi}{L} (n_1, n_2, n_3) \quad n_1, n_2, n_3 = \text{interi.}$

Se quindi si vogliono studiare fluttuazioni di grande lunghezza d'onda si possono incontrare problemi. Vedremo più avanti anche problemi legati alla dinamica delle fluttuazioni.

- Un secondo problema (minore però) è che il sistema non è veramente isotropo (si vede tramite la $g(r)$).

Il lato della scatola di simulazione sarà un multiplo intero di a : $L = ma$
 e il numero totale di pille è quindi: $N = nL^3 = 4 \cdot m^3$

Questo dà luogo ai numeri magici usati in simulazione: 32, 64, 128, 256 etc.

P²: Quanto detto finora è quanto basta per una simulazione con metodo MC. Nella MD invece occorre fissare anche le velocità ~~al posto~~ ^{invece} iniziali perché le eq. da integrare sono del secondo ordine. Anche qui c'è un'ambiguità: una ^{gli} periferici motori addetti in genere si procede come segue. Si campiano le velocità iniziali da una Maxwelliana corrispondente ad una temperatura T_0 (che non ha il significato di temperatura perché in generale il sistema a $t=0$ non è in equilibrio). Durante l'evoluzione si segue l'andamento istantaneo di alcuni ~~dei~~ indicatori statistici. Quando questi ~~si~~ ^{si} stabiliscono in modo stazionario si è arrivati all'equilibrio e si può cominciare ad effettuare l'esperimento di simulazione.

Per completare questi preliminari dobbiamo ancora spiegare come si possano campiare 3N numeri da una Maxwelliana (gaussiana).

Il primo passo per realizzare un campionamento da una distribuzione data consiste nel disporre di un computer ai numeri pseudo casuali distribuiti uniformemente tra $[0, 1]$.

Questo si ottiene da un calcolatore con l'algoritmo seguente che permette di prendere la parte meno significativa del risultato di moltiplicazioni complicate.

1) Consideriamo una macchina che lavora con una parola di 2^{64} bits (n). Scriviamo la formula di ricorrenza:

$$X_{n+1} = X_n \cdot R_0 \pmod{2^{64}} \quad n=0, 1, \dots$$

con R_0 un numero complicato (es. 5^{13}) e x_0 , l'inizio della sequenza è anche lui complicato (es un numero primo 103).

Allora la successione degli x_n è costituita da numeri compresi tra 0 e 2^{64} senza regolarità evidenti e che soddisfanno tutti i test di ~~casualità~~ casualità (media, varianza, momenti superiori etc). Basta dividerli per 2^{64} per ottenere il campione voluto di numeri pseudo-aleatori. Considerati perché data la media due, sicuramente si avranno ricorrenze nella successione, al più tardi dopo 2^{64} esponenti. Tuttavia per tutti gli scopi pratici risultano aleatori.

- 2) Supponiamo ora di dover ricavare dalla successione di numeri pseudo-aleatori generativi dei numeri distribuiti secondo una legge di probabilità $f(x)$. Abbiamo cioè una variabile aleatoria Y distribuita uniformemente e ne vogliamo una X distribuita con $f(x)$. Quello che dobbiamo fare è un cambiamento di variabile nel seguente modo:

Supponiamo che:

$$Y: g(y) = \begin{cases} 1 & y \in (0, 1) \\ 0 & y \notin (0, 1) \end{cases} \Rightarrow \int_0^1 g(y) dy = 1$$

$$X: f(x) : \int_{-\infty}^{\infty} f(x) dx = 1$$

cerchiamo $y = \phi(x)$ che sia invertibile e che:

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_{y(-\infty)}^{y(\infty)} \frac{dy}{|\phi'(x(y))|} f(x(y)) = \int_0^1 g(y) dy$$

L'identificazione naturale è:

$$\begin{cases} y(-\infty) = 0 \\ y(\infty) = 1 \end{cases} \quad g(y) = \frac{f(x(y))}{|\phi'(x(y))|} = 1$$

$$\Rightarrow f(x(y)) = |\phi'(x(y))| \rightarrow \phi(x) = \int_{-\infty}^x f(x) dx = \overset{\substack{\uparrow \\ \text{funzione} \\ \text{cumulativa}}}{F(x)} = y$$

NOTA: Per poter cadere il modello e quindi ottenere la relazione finale è possibile solo grazie al fatto che $f(x) \geq 0$ per definizione e quindi $F(x)$ è monotona non decrescente

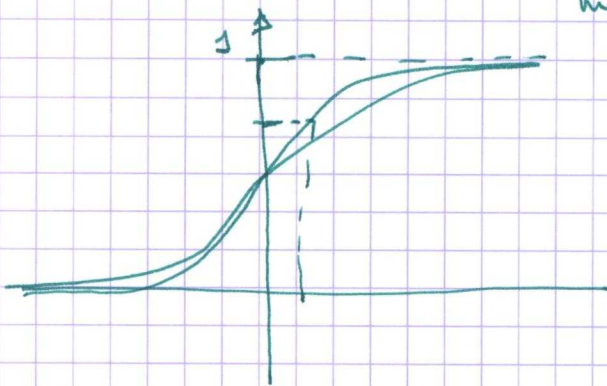
e la sua derivata ha segno definito tra $(0,1)$ / Are note.

Poi

Se ora inverto la relazione $Y=F(x)$ ottengo l'esempio cercato: $X=F^{-1}(Y)$

Nel caso specifico gaussiano la funzione cumulativa è l'error function

$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{-\infty}^x e^{-x'^2} dx'$ e l'inversione può essere fatta graficamente ma non analiticamente.



Quindi campione uniformemente Y
e poi prendo $X=F^{-1}(Y)$

Non è pratico invertire in un computer una funzione implicita come l' $\text{erf}(x)$

Allora in pratica si usano due variabili aleatorie uniformi (U, V)

e si ottengono due variabili gaussiane standardizzate (X, Y)

con il seguente procedimento:

$$f(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} \rightarrow F(x, y) = \frac{1}{2\pi} \int_{-\infty}^x dx' \int_{-\infty}^y dy' e^{-\frac{x'^2+y'^2}{2}}$$

$$\text{in coordinate polari: } \begin{cases} x = R \cos \theta \\ y = R \sin \theta \end{cases} \Rightarrow F(R, \theta) = \frac{1}{2\pi} \int_0^R \int_0^{2\pi} r dr d\theta e^{-r^2/2}$$

questo significa che la densità di probabilità di R e θ è il prodotto di due distribuzioni indep. per θ è uniforme tra $(0, 2\pi)$, per r è quasi gaussiana tra $(0, \infty)$: $f(r, \theta) = \frac{1}{2\pi} r e^{-r^2/2}$

Assoc. V a θ : $\theta = 2\pi V$

Poi: $F(R) = \int_0^R r e^{-r^2/2} dr = U$ e la inverta $\hat{F}(R) = U = 1 - e^{-R^2/2}$

$$\Rightarrow \left\{ \begin{array}{l} R = \{-2 \ln(1-U)\}^{1/2} \\ \theta = 2\pi V \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{l} X = [-2 \ln(1-U)]^{1/2} \cos(2\pi V) \\ Y = [-2 \ln(1-U)]^{1/2} \sin(2\pi V) \end{array} \right. \quad \begin{array}{l} (X, Y) \text{ sono distribuite} \\ \text{secondo gaussiane} \\ \text{standardizzate.} \end{array}$$

Ora le velocità sono estratte da una Maxwelliana:

$$f(v_{i\alpha}) = \frac{1}{\sqrt{2\pi} \, k_B T / m} e^{-\frac{m}{2k_B T} v_{i\alpha}^2} \quad \begin{array}{l} \langle v_{i\alpha} \rangle = 0 \\ \sigma_{i\alpha}^2 = k_B T / m \end{array}$$

$$\Rightarrow v_{i\alpha} = \sqrt{\frac{k_B T}{m}} X$$

NOTA: X, Y non sono statisticamente indipendenti quindi per avere $3N$ velocità indipendenti devo estrarre $6N$ numeri gaussiani.

• CALCOLO ENERGIA E FORZE PER POTENZIALI A CORTO RANGE

Come già detto, l'assunzione di conduttori periodici al bordo, rende il modello da risolvere intrinsecamente periodico nello spazio. Se le interazioni sono di coppia additive la formula che esprime l'energia totale è:

$$U_{\text{tot}} = \frac{1}{2} \sum_{ij} \sum_{\underline{n}} v(|\underline{r}_{ij} + \underline{n}L|) \quad (1)$$

dove $v(|\underline{r}|)$ è l'interazione di coppia (UJ), $\underline{n}$ è un vettore di interi e l'indice indica che $i \neq j$ se $\underline{n} = 0$.

Se il potenziale è a corto range, ossia decade rapidamente con la distanza, è possibile con ottima approssimazione semplificare di molto l'espressione dell'energia totale semplicemente limitando $v(r)$ con un $v_c(r)$:

$$v_c(r) = \begin{cases} v(r) & r \leq r_c \\ 0 & r > r_c \end{cases} \quad v_c(r) = v(r) \Theta(r_c - r)$$

ossia trascurare le interazioni al di là di una certa distanza r_c detto raggio di cut-off.

Se è possibile scegliere $r_c \leq L/2$ allora la somma su $\underline{n}$ in (1) si riduce al solo termine $\underline{n} = 0$ (~~come la somma su $\underline{n}$ nella formula originale~~).

~~quindi:~~ ~~$\sum_{\underline{n} \neq 0}$~~ per il quale $|\underline{r}_{ij} + \underline{n}L| = \text{minimo}$

$$r_{ij}^* = \min_{\underline{n}} |\underline{r}_{ij} + \underline{n}L| \leq r_c$$

~~CUT-OFF SPERICO~~
CUT-OFF SPERICO

L'espressione per l'energia totale diventa:

$$U_c = \frac{1}{2} \sum_{ij} v_c(r_{ij}^*) = \sum_{ij} v_c(r_{ij}^*)$$

Naturalmente l'energia del modello originale sarà:

$$U_{\text{tot}} = U_c + \Delta U$$

Il termine correttivo ΔU può essere stimato assumendo che al di là del cut-off le correzioni sferiche siano trascurabili:

$$\Delta U = \frac{N\rho}{2} \int_{r_c}^{\infty} d\vec{r} \, g(r) v(r) \simeq \frac{N\rho}{2} \int_{r_c}^{\infty} 4\pi r^2 dr v(r)$$

Chiaramente la conversione della minima immagine può essere applicata solo se ΔU è ragionevolmente piccolo rispetto a U_c . Nel caso di potenziali che decadono più lentamente di r^{-3} , ΔU è infinito e non possiamo usare questo metodo.

Il truncamento del potenziale, cambia anche il valore delle altre grandezze termodinamiche. Per esempio la pressione di campo abbiamo visto che è ottenibile da:

$$p^{ex} = - \frac{2\pi}{3} \rho^2 \int_0^{\infty} r^3 g(r) v'(r) dr$$

Allora se vogliamo stimare la p^{ex} del modello originale a partire da quella del sistema con v_c si ha:

$$\begin{aligned} p_c^{ex} &= - \frac{2\pi}{3} \rho^2 \int_0^{\infty} r^3 g(r) dr v_c'(r) = - \frac{2\pi}{3} \rho^2 \int_0^{\infty} r^3 dr g(r) \{ v'(r) \theta(r_c - r) + v(r) \delta(r - r_c) \} \\ &= - \frac{2\pi}{3} \rho^2 \int_0^{r_c} r^3 dr g(r) v'(r) + \frac{2\pi}{3} \rho^2 \int_0^{r_c} r^3 g(r_c) v(r_c) \end{aligned}$$

quindi $p_{MC}^{ex} \simeq p_c^{ex} - \underbrace{\frac{2\pi}{3} \rho^2 r_c^3 g(r_c) v(r_c)}_{(*)} - \frac{2\pi}{3} \rho^2 \int_{r_c}^{\infty} r^3 dr v'(r)$

Il termine $(*)$ deriva dalla discontinuità del potenziale al cut-off che dà origine ad una forza impulsiva.

Anche nel MC, ~~come~~ le forze non giocano nessun ruolo, in DM la traiettoria è generata a partire dalle forze e avere forze impulsive è

piuttosto sgradevole e porta ad errori grandi nell'algoritmo.

Per questo motivo al modello di potenziale precedente si preferisce sostituire il seguente:

$$v_c(r) = [v(r) - v(r_c)] \delta(r_c - r)$$

In questo caso otteniamo:

$$\begin{cases} U_{PBC} = U_c + \left[2\pi \rho \int_0^{r_c} r^2 g(r) dr \right] v(r_c) + 2\pi \rho N \int_{r_c}^{\infty} v(r) r^2 dr \\ P_{PBC}^{ex} = P_c^{ex} - \frac{2\pi}{3} \rho^2 \int_{r_c}^{\infty} r^3 v'(r) dr \end{cases}$$

Quindi il termine impulsivo nella pressione è sparito. Il termine tra parentesi nell'espressione dell'energia esprime il numero medio di coppie in interazione diviso 2 per evitare il conteggio doppio. Se n è il numero medio di vicini ad una palla nel raggio r_c :

$$\frac{Nn}{2} = 2\pi \rho N \int_0^{r_c} r^2 g(r) dr$$

Quel termine riaggiunge quindi quanto tolto avendo shiftato il potenziale.

Con questo secondo modello abbiamo raggiunto lo scopo di eliminare il termine impulsivo nella pressione. Tuttavia le forze derivate dal $v_c(r)$ sono ancora discontinue: $v_c'(r) = v'(r) \delta(r_c - r) + [v(r) - v(r_c)] \delta(r_c - r)$ in quanto $v_c'(r) = v'(r)$ $r < r_c$, $v_c'(r) = 0$ $r > r_c$.

Se si vuole evitare ai problemi ai termini per gli algoritmi di DM si possono scegliere modelli ancora diversi di potenziale che regolamentano la derivata prima. Per esempio:

$$\phi_1(r) = [v(r) - v(r_c) - v'(r_c)(r - r_c)] \delta(r_c - r)$$

$$\phi_1'(r) = [v'(r) - v'(r_c)] \delta(r_c - r) - [v(r) - v(r_c) - v'(r_c)(r - r_c)] \delta(r - r_c)$$

In questo modello si ha:

$$U_\phi = 2\pi\rho \int_0^r r^2 g(r) dr [v(r) - v(r_c) - v'(r_c)(r - r_c)] =$$

$$= 2\pi\rho \int_0^r r^2 g(r) v(r) dr - \cancel{2\pi\rho} \frac{Nn}{2} [v(r_c) - r_c v'(r_c)] - \left(2\pi\rho \int_0^r r^2 g(r) dr \right) v'(r_c)$$

$$\Rightarrow U_{PBC} = U_\phi + \frac{Nn}{2} [v(r_c) - r_c v'(r_c)] + v'(r_c) \left\{ 2\pi\rho \int_0^r r^2 g(r) dr \right\} N$$

$$P_\phi^{ex} = -\frac{2\pi}{3} \rho^2 \int_0^r r^3 g(r) [v'(r) - v'(r_c)] = -\frac{2\pi}{3} \rho^2 \int_0^r r^3 v'(r) g(r) + \frac{2\pi}{3} \rho^2 v'(r_c) \int_0^r r^3 g(r) dr$$

$$\Rightarrow P_{PBC}^{ex} = P_\phi^{ex} - \left(\frac{2\pi}{3} \rho^2 \int_0^r r^3 g(r) dr \right) v'(r_c) - \frac{2\pi}{3} \rho^2 \int_{r_c}^\infty r^3 g(r) v'(r) dr$$

Un altro modello ancora, ~~più simile~~ ~~neglio~~ del precedente perché più simile al modello di Lennard-Jones, si ottiene definendo un intervallo Δ e sostituendo il potenziale shiftato con un polinomio di 3° grado in $(r - \Delta, r_c)$. ~~talché~~ ~~coefficienti~~ ~~del~~ ~~polinomio~~ si possono scegliere richiedendo che: $P_3(r_c) = P_3'(r_c) = 0$

$$P_3(r - \Delta) = v(r - \Delta) - v(r_c)$$

$$P_3'(r - \Delta) = v'(r - \Delta)$$

Il potenziale così definito si scrive:

$$\phi_2(r) = [v(r) - v(r_c)] \theta(r_c - \Delta - r) + P_3(r) \theta(r - r_c + \Delta) \theta(r_c - r)$$

$$\phi_2'(r) = v'(r) \theta(r_c - \Delta - r) - [v(r) - v(r_c)] \delta(r_c - \Delta - r) + P_3'(r) \theta(r - r_c + \Delta) \theta(r_c - r) +$$

$$+ P_3(r) \{ \delta(r - r_c + \Delta) \theta(r_c - r) - \theta(r - r_c + \Delta) \delta(r_c - r) \}$$

e quindi

$$U_{\phi_2} = 2\pi p \left\{ \int_0^{r_c-\Delta} v(r) r^2 g(r) dr - v(r_c) \int_0^{r_c-\Delta} r^2 g(r) dr + \int_{r_c-\Delta}^{r_c} P_3(r) r^2 g(r) dr \right\} =$$

$$= 2\pi p \left\{ \int_0^{r_c-\Delta} v(r) r^2 g(r) dr + \int_{r_c-\Delta}^{r_c} P_3(r) r^2 g(r) dr \right\} - v(r_c) \frac{N n'}{2}$$

con $n' =$ numero medio di vicini in una sfera di raggio $(r_c - \Delta)$

$$p_{\phi_2}^{ex} = -\frac{2\pi p^2}{3} \int_0^{\infty} r^3 \phi_2'(r) dr g(r) =$$

$$= -\frac{2\pi p^2}{3} \left\{ \int_0^{r_c-\Delta} r^3 g(r) v'(r) dr - \cancel{[v(r_c-\Delta) - v(r_c)](r_c-\Delta)^3 g(r_c-\Delta)} + \right.$$

$$\left. + \int_{r_c-\Delta}^{r_c} P_3'(r) r^3 g(r) dr + \cancel{P_3(r_c-\Delta)(r_c-\Delta)^3 g(r_c-\Delta)} - \cancel{P_3(r_c) r_c^3 g(r_c)} \right\} =$$

ma per costruzione:

$$\begin{cases} P_3(r_c-\Delta) = v(r_c-\Delta) - v(r_c) \\ P_3(r_c) = 0 \end{cases}$$

$$p_{\phi_2}^{ex} = -\frac{2\pi p^2}{3} \left\{ \int_0^{r_c-\Delta} v'(r) r^3 g(r) dr + \int_{r_c-\Delta}^{r_c} P_3'(r) r^3 g(r) dr \right\}$$

$$\Rightarrow \begin{cases} U_{pbc} = U_{\phi_2} + v(r_c) \frac{N n'}{2} + 2\pi p \int_{r_c-\Delta}^{r_c} [v(r) - P_3(r)] r^2 g(r) dr + 2\pi p \int_{r_c}^{\infty} r^2 g(r) dr \\ p_{pbc}^{ex} = p_{\phi_2}^{ex} - \frac{2\pi p^2}{3} \int_{r_c-\Delta}^{r_c} [v'(r) - P_3'(r)] r^3 g(r) dr - \frac{2\pi p^2}{3} \int_{r_c}^{\infty} v'(r) r^3 dr \end{cases}$$

Il metodo del cut-off sferico rende il problema calcolabile con un algoritmo di ordine N^2 ($O(N^2)$) ossia per calcolare l'energia totale bisogna fare un doppio loop sulle particelle e quindi un ordine $N(N-1)/2$ operazioni. Anche se la distanza ~~tra~~ di coppia è maggiore di r_c , il lavoro per calcolarla è comunque da fare.

- Un modo per rendere questo algoritmo di ordine $N \ln N$ con $n \ll N$ può essere ottenuto nel caso r_c è considerevolmente minore di $L/2$ con l'aiuto della cosiddetta tavola dei vicini.

Questa consiste in una tavola di memoria ~~in~~ costruita con tutti le coppie ~~del~~ ^a sistema $r_{ij} \leq (r_c + \Delta)$ che viene ~~a~~ ~~meno~~ aggiornata solo ~~ogni~~ ogni 10-50 passi. In effetti il tempo tra due aggiornamenti successivi dipende da Δ e dal coefficiente di diffusione:

se $\langle |\Delta x_i|^2 \rangle \geq \Delta$ anche per una sola i allora aggiorniamo.

~~Per di più se~~ Se $r_c + \Delta \leq L/2$ allora n , il numero di vicini di una pila è $\ll N$ e l'algoritmo è di ordine (Nn) .

Se invece $r_c \leq L/2$ la tavola non aiuta perché comunque per poter usare la tavola deve essere $r_c + \Delta \leq L/2$. ~~quindi~~

- Una convenzione alternativa al cut-off sferico è la cosiddetta CONVENZIONE DELLA MINIMA IMMAGINE ossia considerare per ogni data pila i , tutte le interazioni con le altre pile del sistema periodico all'interno di una scatola cubica di lato L centrata in

$$r_i: \quad r_{ij}^* = \min_j |r_{ij} + nL|$$

In questo caso ~~ha~~ bisogna sempre calcolare $\frac{N(N-1)}{2}$ interazioni e ~~risolvere~~ il sistema risulta meno isotropo.

PROPRIETA' TERMODINAMICHE E $g(r)$

~~temperatura~~

- TEMPERATURA: mentre il metodo MC standard lavora a NVT e quindi T è un vincolo esterno, in DM standard si lavora in NVE . In questo caso la temperatura è una osservabile da calcolare nel corso della simulazione. Associamo alla temperatura termodinamica la media (temporale) della osservabile microscopica:

$$\hat{T}(t) = \frac{2}{3k_B} \hat{K}(t)$$

dove $\hat{K}(t)$ è l'energia cinetica istantanea: $\hat{K}(t) = \frac{1}{2N} \sum_i^N \frac{p_i^2(t)}{m_i}$

quindi: $T = \langle \hat{T}(t) \rangle_t \cong \frac{m}{3N(N+1)} \sum_0^N n \sum_i^N \sum_\alpha^3 v_{i,\alpha}^2(nh)$

h è il periodo di integrazione della soluzione numerica della dinamica e N è il numero di frame cui si media.

L'errore statistico associato a questa media è facile da ottenere sulla base delle teorie degli errori campionari ricordando che le $\{v_{i,\alpha}\}$ sono variabili distribuite gaussianamente a media zero e varianza $k_B T/m$.

Per tali variabili: $\langle v_{i,\alpha}^4 \rangle = 3 \langle v_{i,\alpha}^2 \rangle^2$ $[\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2 = 2\sigma_x^4]$

Inoltre la varianza della media campionaria ottenuta da n osservazioni indipendenti ciascuna di varianza σ^2 è: $\sigma_{\langle x \rangle}^2 = \frac{\sigma_x^2}{n}$

Allora se chiamo ν il numero di frame al di là del quale due misure di $\hat{T}(t)$ sono statisticamente correlate si ottiene:

$$\frac{\delta T}{T} = \sqrt{\frac{2 \nu}{3N(N+1)}}$$

che per $V \approx 30$, $N \approx 900$, $\rho \approx 1000$ da' circa il 3%.

ENERGIA INTERNA DI ECCESSO ESPRESSIONE DI ECCESSO:

Come visto dipende da come viene trattato il potenziale a corto range. Per potenziali a corto range con cut-off sferico e shiftati si ha:

$$\frac{U^{ex}}{N} = \frac{1}{2N} \left\langle \sum_{i \neq j} v_c(r_{ij}^*) \right\rangle + \frac{\langle N \rangle}{2} v(r_c) + 4\pi \rho \int_{r_c}^{\infty} r^2 v(r) dr$$

Derivabile:
$$p^{ex} = -\frac{1}{6V} \left\langle \sum_{i \neq j} r_{ij}^* \frac{\partial v_c}{\partial r_{ij}^*} \right\rangle - \frac{4\pi}{3} \rho^2 \int_{r_c}^{\infty} r^3 v'(r) dr$$

Se anche al potenziale short-range ho interazioni coulombiane allora l'energia coulombiana come visto precedentemente deve essere aggiunta e il contributo alla pressione delle interazioni coulombiane si ottiene direttamente osservando che il viriale coulombiano è:

$$\bar{v}_{\text{coul}}(r_{ij}) = \bar{r}_{ij} \cdot \frac{d}{d\bar{r}_{ij}} \left(\frac{1}{|r_{ij}|} \right) = \bar{r}_{ij} \cdot \frac{\bar{r}_{ij}}{|\bar{r}_{ij}|} \left(-\frac{1}{|\bar{r}_{ij}|^2} \right) = -\frac{1}{|\bar{r}_{ij}|} = -v(r_{ij})$$

quindi
$$p_{\text{coul}}^{ex} = -\frac{1}{6V} \left\langle \sum_{i \neq j} v(r_{ij}) \right\rangle = \frac{1}{3V} U_{\text{coul}}^{ex}$$

CALCOLO DELLA $g(r)$

Il calcolo della $g(r)$ procede essenzialmente dall'interpretazione della $g(r)$ come densità condizionata delle p.p.le a distanza data da una prefissata. Si ha:

$$V^2 P^{(2)}(r', r' + r) = g(r)$$

$$\Rightarrow 4\pi \rho \int_V dr r^2 g(r) = N = \int dr N(r)$$

dove l'ultima uguaglianza definisce l'istogramma $N(r)$ che può

essere ottenuto direttamente durante il calcolo e che fornisce il numero medio di palle a distanza fra r e $r+dr$.
 $\uparrow$ primo dell'istogramma.

Da $N(r)$ si ottiene:

$$g(r) = N(r) / 4\pi r^2$$

In pratica:

$$g(r) \approx \frac{N(r, r+\Delta r)}{\frac{4\pi}{3} [(r+\Delta r)^3 - r^3]}$$

11,05,98 / FINE

|————|

13,05,98/4

DINAMICA MOLECOLARE : ALGORITHMI DI INTEGRAZIONE

E' ovvio che un buon programma di DM ha bisogno di un buon algoritmo d'integrazione. Comunque ~~è~~ è facile riconoscere un cattivo algoritmo ma non altrettanto facile riconoscere un buon algoritmo.

Bisogna considerare i seguenti requisiti:

- ① velocità di calcolo
 - ② memoria richiesta
 - ③ Accuratezza verso fine step.
 - ④ Conservazione dell'energia $\left\{ \begin{array}{l} \text{short time fluctuations} \\ \text{long time drift} \end{array} \right.$
 - ⑤ Verifica delle leggi di conservazione del moto: time-reversibility e conservazione del volume in Γ .
 - ⑥ Facilità di coding
- I punti 1 e 2 non sono molto rilevanti oggi perché 1) il ~~tempo~~ tempo per integrare le eq. $\sim O(N)$ e in genere $\ll$ del calcolo delle forze
 2) la memoria per un calcolo DM è trascurabile anche per sistemi grandi (non troppo!!).

il punto ③ è invece cruciale perché permette di prendere un time step grande e quindi ridurre il numero di calcoli delle forze ($O(N^2)$).

L'algoritmo senza dubbio più usato è il cosiddetto algoritmo di Verlet, che si ottiene da un semplice sviluppo di Taylor della traiettoria intorno al tempo presente:

$$\begin{cases} r(t+h) = r(t) + h \dot{r}(t) + \frac{h^2}{2} \ddot{r}(t) + \frac{h^3}{6} \dddot{r}(t) + O(h^4) \\ r(t-h) = r(t) - h \dot{r}(t) + \frac{h^2}{2} \ddot{r}(t) - \frac{h^3}{6} \dddot{r}(t) + O(h^4) \end{cases}$$

$$\begin{cases} r(t+h) + r(t-h) = 2r(t) + h^2 \ddot{r}(t) + O(h^4) \\ r(t+h) - r(t-h) = 2h \dot{r}(t) + O(h^3) \end{cases}$$

$$\Rightarrow \begin{cases} r(t+h) = 2r(t) - r(t-h) + h^2 \ddot{r}(t) + O(h^4) \\ \dot{r}(t) = \frac{r(t+h) - r(t-h)}{2h} + O(h^3) \end{cases}$$

Considerazioni:

- 1) richiede poco tempo di CPU
- 2) richiede poco storage ($r(t)$, $r(t-h)$, $F(t)$)
- 3) Le traiettorie (posizioni) sono all'ordine h^4 , infatti le velocità non sono usate per avanzare la dinamica, ma solo calcolate per ottenere le medie d'ensemble. Quindi la minore precisione sulle velocità non rovina la dinamica.
- 4) Sono time reversible invariant, ossia se vedo indietro nel tempo ripercorro esattamente (a meno dell'arrotondamento reale) la traiettoria da dove sono venuto.

5) Fornisce posizioni e velocità a tempi diversi (sgradevole).

La proprietà 4 è cruciale per la bontà di un integratore.

Infatti, data la natura ergodica dei sistemi che si studiano non è possibile trovare un integratore perfetto (ovvero che segua la vera traiettoria) anche prendendo espressioni in serie a ordine molto più alto. Infatti a causa dell'instabilità di Lyapunov anche il solo errore di arrotondamento rende dopo breve tempo la traiettoria "vera" e quella numerica completamente scorrelate.

Se però si vuole fare la Mec. Stat. si ha bisogno di campionare la superficie ad energia costante e quindi è cruciale per i nostri scopi la conservazione dell'energia a tempi lunghi.

L'idea intuitiva è che la traiettoria numerica rimane sempre vicina a qualche traiettoria "vera" mai conosciuta e questo dà una spiegazione del perché la DM ~~non~~ dà ottime predizioni anche per le proprietà dinamiche. Tuttavia non esiste nessuna prova matematica ma solo evidenze "sperimentali".

Un altro aspetto legato alla conservazione dell'energia è la conservazione del volume in Γ . Algoritmi non time-reversible invariant non possono conservare il volume in Γ e quindi l'energia.

VELOCITY VERLET

Per risolvere la sgradevolezza 5) possiamo riscrivere Verlet in una versione che fornisce posizioni e velocità a stesso tempo. Questo è il cosiddetto Velocity-Verlet:

$$\begin{cases} r(t+h) = r(t) + h \dot{r}(t) + \frac{h^2}{2} \ddot{r}(t) + O(h^3) \\ \dot{r}(t+h) = \dot{r}(t) + \frac{h}{2} [\ddot{r}(t) + \ddot{r}(t+h)] + O(h^3) \end{cases}$$

Benché appare diverso, il velocity-verlet fornisce la stessa traiettoria numerica del Verlet, ed ha quindi le stesse proprietà. Per vedere ciò basta usare la prima delle equazioni ⁽¹⁾ per ottenere:

$$r(t+2h) = r(t+h) + h \dot{r}(t+h) + \frac{h^2}{2} \ddot{r}(t+h)$$

$$r(t) = r(t+h) - h \dot{r}(t) - \frac{h^2}{2} \ddot{r}(t)$$

Sommando queste equazioni si ha:

$$r(t+2h) + r(t) = 2r(t+h) + h [\dot{r}(t+h) - \dot{r}(t)] + \frac{h^2}{2} [\ddot{r}(t+h) - \ddot{r}(t)]$$

Sostituendo la seconda equazione (2) per $\dot{r}(t+h) - \dot{r}(t)$ si ha:

$$r(t+2h) + r(t) = 2r(t+h) + h^2 \ddot{r}(t+h) \quad \text{ovvero Verlet standard}$$

~~Per argon at~~

- Scelta di h: h deve essere scelto non troppo piccolo a fronte la dinamica computazionale che si richiede ad essere costante ma allo stesso tempo non troppo grande da rendere instabile l'algoritmo. Il test è sempre la conservazione (short- or long time) dell'energia e dell'impulso. Per l'argon al punto triplo: $h \approx 10^{-15} \div 10^{-14}$ sec ~~di più~~ è un buon compromesso.

Con questo passo il sistema raggiunge l'equilibrio in un 500 ÷ 600 passi e la ~~la~~ $C_v(t)$ muore in 300 ÷ 400 passi

~~Per~~

- Per sapere quando si è arrivati all'equilibrio bisogna monitorare le grandezze di interesse e vedere da che tempo in poi fluttuano in maniera stazionaria. Da lì si appare di essere all'equilibrio.

FORMULAZIONE LIOUVILLIANA DEGLI ALGORITMI TIME-REVERSIBLE

Fino qui abbiamo introdotto gli algoritmi dal punto di vista delle matematiche applicate. Partiamo ora dalla meccanica newtoniana in Γ . Sappiamo che l'operatore che genera il moto è l'operatore di Liouville:

$$iL = \dot{r} \frac{\partial}{\partial r} + \dot{p} \frac{\partial}{\partial p} = iL_r + iL_p$$

L'evoluzione del moto di una qualunque funzione di Γ è:

$$f(r(t), p(t)) = e^{iL t} f(r(0), p(0))$$

Consideriamo i due termini di iL (non commutanti) separatamente. Allora poiché dall'eq. di Hamilton:

$$\begin{cases} \dot{r} = \frac{p}{m} & (\text{indipendente da } r) \\ \dot{p} = -F(r) & (\text{indipendente da } p) \end{cases}$$

$$\begin{aligned} \text{Si ha: } e^{iL_p t} f(r(0), p(0)) &= e^{t \dot{r} \frac{\partial}{\partial r}} f(r(0), p(0)) = \\ &= f(r(0) + \dot{r}(0)t, p(0)) \end{aligned}$$

allo stesso modo:

$$e^{iL_r t} f(r(0), p(0)) = f(r(0), p(0) + t \dot{p}(0))$$

Quindi possiamo scrivere $e^{iL t} \approx e^{iL_r t} e^{iL_p t}$ $\textcircled{A}$

ovvero se gli operatori fossero commutativi avremmo risolto il problema dell'evoluzione temporale.

Naturalmente la $\textcircled{A}$ non è mai vera ma possiamo approssimarla con:

$$e^{i\dot{L}t} = \left(e^{i\dot{L}t/P} \right)^P = \lim_{P \rightarrow \infty} \left[e^{i\dot{L}_P \frac{t}{2P}} e^{i\dot{L}_R t/P} e^{i\dot{L}_P \frac{t}{2P}} \right]^P$$

13,5/28/6

questo significa però fare infiniti passi infinitesimi.

Se prendiamo P finito si può mostrare che l'errore è di ordine $e^{(+1/P^2)}$.

Questa è la fattorizzazione di Trotter del dinamilliano!!

Chiamo $t/P = \tau$

$$e^{i\dot{L}\tau} = e^{i\hbar(L_P + L_Q)} = e^{i\hbar L_P} e^{i\hbar L_Q} = \frac{\tau^2}{2} O(\tau^2)$$

$$e^{i\dot{L}t} \approx \left[e^{i\frac{\tau}{2} \dot{L}_P} e^{i\tau \dot{L}_R} e^{i\frac{\tau}{2} \dot{L}_P} \right]^P + O(e^{+1/P^2}) O(\tau^3)$$

Quindi per arrivare al punto t faccio P passi di lunghezza τ e ad ogni passo devo applicare l'operatore:

$$e^{i\frac{\tau}{2} \dot{L}_P} e^{i\tau \dot{L}_R} e^{i\frac{\tau}{2} \dot{L}_P} f(r(0), p(0)) =$$

$$= e^{i\frac{\tau}{2} \dot{L}_P} e^{i\tau \dot{L}_R} f\left(r(0), p(0) + \frac{\tau}{2} \dot{p}(0)\right) =$$

$$= e^{i\frac{\tau}{2} \dot{L}_P} f\left(r(0) + \tau \dot{r}(0), p(0) + \frac{\tau}{2} \dot{p}(0) + \tau \dot{r}(0)\right) =$$

$$= f\left(r(0) + \tau \dot{r}(0) + \frac{\tau}{2} \dot{p}(0), p(0) + \frac{\tau}{2} \dot{p}(0) + \frac{\tau}{2} \dot{p}(0) + \tau \dot{r}(0)\right)$$

Se per esempio applico il simbolo hamiltoniano al vettore $\begin{Bmatrix} r(0) \\ p(0) \end{Bmatrix}$ ho:

$$e^{i\frac{\tau}{2} \dot{L}_P} e^{i\tau \dot{L}_R} e^{i\frac{\tau}{2} \dot{L}_P} \begin{Bmatrix} r(0) \\ p(0) \end{Bmatrix} =$$

$$= e^{i\frac{\tau}{2} \dot{L}_P} e^{i\tau \dot{L}_R} \begin{Bmatrix} r(0) \\ p(0) + \frac{\tau}{2} F(r(0)) \end{Bmatrix} =$$

$$= e^{i \frac{\tau}{2} \Delta p} \left\{ \begin{array}{l} r(0) + \frac{p(0)}{m} \tau \\ p(0) + \frac{\tau}{2} F\left[r(0) + \frac{\tau p(0)}{m}\right] \end{array} \right\} =$$

$$= \left\{ \begin{array}{l} r(0) + \frac{\tau}{m} \left[p(0) + \frac{\tau}{2} F(r(0)) \right] \\ p(0) + \frac{\tau}{2} F(r(0)) + \frac{\tau}{2} F\left[r(0) + \frac{\tau p(0)}{m}\right] \end{array} \right\} = \left\{ \begin{array}{l} r(\tau) \\ p(\tau) \end{array} \right\}$$

$+ \frac{\tau^2}{2m} F(r(0))$

$$\left\{ \begin{array}{l} r(\tau) = r(0) + \frac{\tau}{m} p(0) + \frac{\tau^2}{2m} F(0) \\ p(\tau) = p(0) + \frac{\tau}{2} \{ F(0) + F(\tau) \} \end{array} \right\}$$

Abbiamo quindi ritrovato l'algoritmo Velocity-Verlet. Poiché ogni passo temporale corrisponde all'applicazione di 3 operazioni traslazioni e ogni traslazione ha Jacobiano unitario, allora l'intero passo temporale ha jacobiano unitario. Dunque l'algoritmo di Verlet conserva la misura (è simplettico) come la vera dinamica ~~microscopica~~ continua.

CALCOLO DELLE PROPRIETA' DINAMICHE CON LA DM

Abbiamo visto che la conoscenza delle proprietà dinamiche di un sistema è la conoscenza delle funzioni di correlazione temporali (o spatio-temporali) delle quantità "locali" associate al sistema. Se $A(t)$ è una quantità temporale, omonimo processo stocastico, e se siamo all'equilibrio allora le f.d.c. correlazione temporale del processo può essere calcolata come media sulla traiettoria.

$$C_0(t) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T ds A(s) A(t+s) = \lim_{T \rightarrow \infty} C_T(t)$$

Questa è la maniera di fare le medie con la DM e quindi è l'espressione operativa per il calcolo delle f.d.c. temporali.

Naturalmente in DM si conosce solo una traiettoria e per un tempo finito quindi possiamo calcolare $C_T(t)$ con T finito e usarla come stima di $C_0(t)$. Questa stima sarà affetta da errori statistici (che è importante stimare) e da effetti sistematici che bisogna minimizzare.

❶ Gli effetti sistematici derivano essenzialmente dal numero di N finito e dalle condizioni al bordo. ~~tra~~ Oltre agli effetti già discussi per le proprietà statiche (dipendenza $1/N$ delle proprietà termodinamiche (120) e discretizzazione dello spettro dei modi collettivi) abbiamo un effetto ~~sistematico~~ di bordo sulla dinamica. Infatti dopo un tempo che è dell'ordine di grandezza del tempo che impiega un disturbo (perturbazione) ad attraversare la scatola la dinamica risulta alterata dal bordo. Con condizioni periodiche al bordo, una perturbazione creata al centro della scatola ~~arriva~~ arriva al bordo e interferisce con la sua immagine nella scatola.

ma è la sola possibile per ottenere una stima teorica di queste variabili. Se $A(t)$ è gaussiana:

$$\langle A_1 A_2 A_3 A_4 \rangle = \langle A_1 A_2 \rangle \langle A_3 A_4 \rangle + \langle A_1 A_3 \rangle \langle A_2 A_4 \rangle + \langle A_1 A_4 \rangle \langle A_2 A_3 \rangle$$

e quindi

$$\langle \Delta(t_1) \Delta(t_2) \rangle = \frac{1}{\tau_0^2} \int_0^{\tau_0} ds_1 \int_0^{\tau_0} ds_2 \left[C_{\infty}(s_2 - s_1) C_{\infty}(s_2 - s_1 + t_2 - t_1) + C_{\infty}(s_2 - s_1 + t_2) C_{\infty}(s_2 - s_1 - t_1) \right]$$

In generale $C_{\infty}(t)$ decade a zero dopo qualche tempo caratteristico τ_c . Nelle simulazioni $T \gg \tau_c$ ($T \approx 30 \tau_c$). Inoltre t_1 e t_2 ci interessano per $t_i \sim \tau_c \ll T$. Il contributo dominante dagli integrali viene dalla regione $|s_1 - s_2| \sim \tau_c \ll T$. Facendo un cambiamento di variabili si ha:

$$\begin{aligned} \text{ha: } (\sigma = s_1 - s_2, s_2 = s_2) \\ \langle \Delta(t_1) \Delta(t_2) \rangle &\approx \frac{1}{\tau_0^2} \int_{-\infty}^{\infty} d\sigma \int_0^{\tau_0} ds_2 \left[C_{\infty}(\sigma) C_{\infty}(\sigma + t_2 - t_1) + C_{\infty}(\sigma + t_2) C_{\infty}(\sigma - t_1) \right] \\ &\approx \frac{2}{\tau_0} \int_{-\infty}^{\infty} d\sigma C_{\infty}^2(\sigma) = \frac{4}{\tau_0} \int_0^{\infty} d\sigma C_{\infty}^2(\sigma) \end{aligned}$$

Se definisco:

$$\tau = 2 \int_0^{\infty} d\sigma C_{\infty}^2(\sigma) / C_{\infty}^2(0) \quad \left[\text{esattamente } C_{\infty}(t) = C_{\infty}(0) e^{-t/\tau} \right]$$

si ha:

$$\langle \Delta(t_1) \Delta(t_2) \rangle \approx 2 \frac{\tau}{\tau_0} C_{\infty}^2(0)$$

~~Questo~~ Questa stima ci dice che l'errore da attribuire a $C_{\infty}(t)$ come stima di $C_{\infty}(t)$ è quindi

$$\propto \left(\frac{\tau}{\tau_0} \right)^{1/2}$$

che è difficile da rendere trascurabile perché decresce come

La radice del tempo di integrazione τ delle espressioni del moto.

Se oltre alla quantità di interesse è di angoli p.lle (come la correlazione velocità-velocità) allora si guadagna un ulteriore fattore $N^{-1/2}$ perché sono mediane a tutte le p.lle

$$\langle v(t)v(0) \rangle \approx \frac{1}{N} \sum_i^N \frac{1}{\tau} \int_0^\tau ds \langle v_i(t+s)v_i(s) \rangle$$

COEFFICIENTI DI TRASPORTO

• Diffusione

Per derivare un'espressione della diffusione in termini di medie di ensemble partiamo dall'espressione di Fick e dall'eq. di continuità per la concentrazione di particelle etichettate.

Fick : $\underline{j} = -D \nabla c$

$\uparrow$ corrente di particelle $\nwarrow$ concentrazione di particelle

$\underline{j}$ e c sono campi macroscopici locali e non osservabili microscopici. Stiamo pensando di aver veduto su volumetti piccoli ma macroscopici, e di assumere l'equilibrio locale.

CONTINUITA :

$$\frac{\partial c}{\partial t} + \nabla \cdot \underline{j} = 0$$

non esistono sorgenti di particelle etichettate

$$\Rightarrow \partial_t c + \nabla^2 c = 0$$

c.c. : $c(r, 0) = \delta(r)$

$$c(r, t) = \frac{1}{(4\pi Dt)^{d/2}} e^{-\frac{r^2}{4Dt}}$$

ovvero una distribuzione localizzata a $t=0$ si diffonde e si mescola in gaussiane.

$$\int_V dr c = 1$$

D è legato al movimento random di $c(r, t)$:

$$\langle r^2(t) \rangle = \int \frac{d\bar{r}}{(4\pi Dt)^{d/2}} e^{-\frac{r^2}{4Dt}} r^2 = \sigma^2 = 2Dt d$$

$$\langle |r(t) - r(0)|^2 \rangle = 2dDt$$

ma $r(t) = \int_0^t v(t') dt' + r(0)$

$$\Rightarrow \mathcal{L} D t = \left\langle \int_0^t dt' \int_0^t dt'' v(t') \cdot v(t'') \right\rangle = \int_0^t dt' dt'' \langle v(t') \cdot v(t'') \rangle$$

~~per ipotesi~~

A questo punto devo associare medie ai campi ~~locali~~ e medie microscopiche (di ensemble) e identificare la $\langle v(t'') \cdot v(t') \rangle_c \equiv \langle v(t'') \cdot v(t') \rangle_{ens.}$

Per sistemi all'equilibrio, isotropi ^{omogenei} e ~~invarianti per trasl.~~

$$\begin{aligned} \langle v_\alpha(t'') \cdot v_\beta(t') \rangle &= C_{v_\alpha v_\beta}(t'', t') = C_{v_\alpha v_\beta}(t'' - t') = \langle v_\alpha(t'' - t') \cdot v_\beta(0) \rangle = \\ &= C_{v_\alpha v_\beta}(t'' - t') \delta_{\alpha\beta} \\ &\quad \underline{\underline{\text{diagonale}}} \end{aligned}$$

~~ma non è così~~

Usiamo ora il lemma seguente:

$$\left\{ \begin{array}{l} \text{Data una funzione di due variabili } f(x, y): f(x, y) = f(|x - y|) \\ \text{allora } \int_0^L dx \int_0^L dy f(|x - y|) = 2L \int_0^L d\sigma f(\sigma) \left(1 - \frac{\sigma}{L}\right) \end{array} \right.$$

$$\begin{aligned} \Rightarrow \mathcal{L} D t &= 2t \int_0^t dz C_{vv}(z) \left(1 - \frac{z}{t}\right) = \\ &= 2t \int_0^t dz C_{vv}(z) - 2 \int_0^t dz z C_{vv}(z) = \\ &= 2t A(t) - 2B(t) \end{aligned}$$

Per t molto grandi

$$B(t) \rightarrow B$$

} constant

$$\Rightarrow \mathcal{L}Ddt = \mathcal{L}tA - \mathcal{L}B$$

$$\Rightarrow \boxed{A \longleftrightarrow D}$$

$$t \rightarrow \infty$$

quindi solo per tempi grandi ottengo la diffusione di Fick (Einstein) dallo

~~Radice~~ Quadrato e potenza delle potenze
quadrato

Formule di Green-Rubio:

DIFFUSIONE: $D = \lim_{t \rightarrow \infty} \frac{1}{2t} \langle \hat{v}_x^2(t) \rangle$

coefficiente
di trasporto

Diffusion

mobilität: $\mu = \frac{D}{kT}$

VISCOSITA' DI TAGLIO : Newton:

$$\tau_{xy}(r,t) = \eta [\partial_x u_y(r,t) + \partial_y u_x(r,t)]$$

↑
viscosità di taglio

$$\eta = \frac{1}{V k_B T} \int_0^\infty dt \langle \hat{\sigma}_{xy}(0) \hat{\sigma}_{xy}(t) \rangle_\infty$$

$$\frac{\partial}{\partial x_j} \mathcal{L} = \sum_i^N \left[m_i \hat{v}_i^x \hat{v}_i^y + \frac{1}{2} \sum_{j \neq i} \hat{x}_{ij} \hat{p}^y(r_{ij}) \right]$$

CONDUCIBILITA' TERMICA:

Fourier:

$$\vec{J}^q(r,t) = -\lambda_T \vec{\nabla} T(r,t)$$

↑
condensability
termine

$$\lambda_T = \frac{1}{v k T^2} \int_0^\infty dt \langle \hat{j}_z^q(t) \hat{j}_z^q(0) \rangle_{eq}$$

$$\hat{J}_z^q = \frac{d}{dt} \sum_i^N \frac{\hat{z}_i}{2} \left(\frac{1}{2} m \hat{v}_i^2 + \sum_{j \neq i} v(r_{ij}) \right)$$

CONDUCEBILITA' ELETTRICA : Ohm : $\underline{j}^{el}(r,t) = \sigma \underline{E}(r,t)$

↑
scalare per sistemi
~~isotropi~~ isotropi

$$\sigma = \frac{1}{V k_B T} \int_0^\infty dt \langle \hat{j}_x^{el}(t) \hat{j}_x^{el}(0) \rangle_{eq}$$

$$\hat{j}_x^{el} = \sum_i^N q_i \hat{v}_i$$

↑
carica della particella i-esima.

DINAMICA MOLECOLARE IN ALTRI ENSEMBLES

In natura l'accoppiamento con il mondo esterno avviene attraverso le pareti del recipiente che contiene il sistema in esame, sia che si tratti di scambio di calore, che di calore e massa (non consideriamo qui l'interazione con la radiazione che invece avviene solo in superficie). Abbiamo però visto come in simulazione si sia cercato di eliminare le pareti con l'adozione delle PBC.

Sembra quindi improponibile pensare alla modellizzazione della dinamica di un sistema in accoppiamento con un ~~termostato~~ bagno termico o con un pistone o meglio con un bagno di p.le a potenziale chimico fissato.

In effetti i primi tentativi fatti modellizzando l'accoppiamento naturale hanno messo in evidenza con chiarezza che l'accoppiamento di superficie è da evitare quanto più possibile.

Si è allora passati ad accoppiamenti di volume con bagni in grado di scambiare rispettivamente calore, momento e p.le.

Per far questo si può in generale usare un approccio ~~alla~~

~~stocastico~~ stocastico (Andersen) oppure un approccio dinamico (hanno che per il gran canario).

DM nel CANONICO:

- L'approccio stocastico (Andersen) consiste nel rappresentare l'accoppiamento con i termostati attraverso coefficienti stocastici istantanei che cambiano la velocità di alcune p.le scelte a caso e che agiscono di tanto in tanto. Senza entrare di più nei dettagli del metodo (vedi Frankel Smit) si può dimostrare che l'ensemble risultante da una tale Hamiltoniana è canonico ~~questo~~ cioè quello che si voleva. Tuttavia il metodo risulta inefficiente e quindi ha perso

il carattere dinamico della DM rispetto a MC.

- L'approccio dinamico introdotto da Nozè e riformulato da Hoover invece considera l'ensemble canonico dinamicamente e non ha elementi stocastici ^e ~~che~~ permette di ottenere una traiettoria continua e spazialmente stile per calcolare anche le proprietà dinamiche.

22.3.82
11.12.89
3

1. Un metodo di D.M. per generare l'ensemble canonico.
 2. Natura dell'ensemble
 3. Equivalenza con il microcanonico
-

1. Dinamica (arbitrariamente inventata) non derivabile (almeno direttamente) da una hamiltoniana:

$$\begin{aligned}
 (*) \quad & \frac{dr_i}{dt} = \frac{p_i}{m_i} \\
 & \frac{dp_i}{dt} = -\frac{\partial V_N}{\partial r_i} - \zeta p_i \\
 & \frac{d\zeta}{dt} = \left(\sum_i \frac{p_i^2}{m_i} - 3Nk_B T \right) / Q \\
 & \frac{d\ln S}{dt} = \zeta
 \end{aligned}$$

dove la costante Q fa le funzioni di una "massa" e governa l'intensità dell'accoppiamento. La quantità

$$\begin{aligned}
 \mathcal{H} &\equiv H^T + 3Nk_B T \ln S \\
 &= \sum \frac{p_i^2}{2m} + V_N + Q \zeta^2 / 2 + 3Nk_B T \ln S
 \end{aligned}$$

pur non essendo il hamiltoniano del sistema dinamico (*) è conservata:

22.3.89
11.12.88
X

$$\frac{dH}{dt} = \sum \frac{p_i}{m} \dot{p}_i + \sum \nabla_i V_N \dot{r}_i + Q \dot{\zeta} + 3Nk_B T (\dot{\ln \zeta})$$

$$= \sum_i \left[\dot{r}_i \nabla_i V_N - \zeta \dot{r}_i p_i + \dot{r}_i \nabla_i V_N \right] + \zeta \left[\sum_i \frac{p_i^2}{2m_i} - 3Nk_B T \right] + 3Nk_B T \dot{\zeta} = 0 !$$

Inoltre, notando che s è solo una variabile ausiliaria di ζ (ζ determina s , ma non ne è determinata) e concentrando l'attenzione sulle variabili $\{p, r, \zeta\}$, si dimostra facilmente che l'ensemble generato da queste dinamiche è

$$\rho(p, r, \zeta) \propto \exp(-H^T/k_B T)$$

* Scriviamo l'eq. di Liouville per queste dinamiche hamiltoniane ($x \equiv (r, p, \zeta)$)

$$\frac{\partial \rho}{\partial t} + \text{div } \rho \dot{x} = 0$$

$$\frac{\partial \rho}{\partial t} + \dot{x} \nabla_x \rho + \rho \nabla \dot{x} \equiv \frac{d\rho}{dt} + \rho \nabla \dot{x} = 0$$

$$\text{ma } \rho \nabla \dot{x} = \rho \left\{ \sum_i \left[\frac{\partial}{\partial r_i} \dot{r}_i + \frac{\partial}{\partial p_i} \dot{p}_i \right] + \frac{\partial}{\partial \zeta} \dot{\zeta} \right\}$$

$$= \rho \cdot [-\zeta \cdot 3N] \Rightarrow \frac{d\rho}{dt} = 3N\zeta \rho$$

Ma ~~ma~~ $\frac{dH^T}{dt}$ è uguale a $-3N\zeta k_B T$

$$\frac{dH^T}{dt} = \frac{\partial H^T}{\partial t} + \sum_i \left[\frac{\partial H^T}{\partial r_i} \dot{r}_i + \frac{\partial H^T}{\partial p_i} \dot{p}_i \right] + \frac{\partial H^T}{\partial \zeta} \dot{\zeta}$$

$$= \sum_i \left[\left(-\dot{p}_i + \zeta p_i \right) \dot{r}_i + \frac{p_i}{m_i} \dot{p}_i \right] + Q \dot{\zeta}$$

$$= -\zeta \sum_i \frac{p_i^2}{m_i} + Q \dot{\zeta} \left(\sum_i \frac{p_i^2}{m_i} - 3N k_B T \right) / Q$$

$$= -3N k_B T \dot{\zeta} \quad \text{dato che } \frac{d \langle E \rangle}{dt} = 0 \quad \text{e} \quad \dot{\zeta} = \frac{d(\ln \zeta)}{dt}$$

per cui

$$\frac{dp}{dt} = p \cdot \frac{-1}{k_B T} \frac{dH^T}{dt}$$

Di qui:

$$p = c e^{-H^T/k_B T} \quad \text{c.v.d.}$$

Come al solito c è una normalizzazione:

$$c^{-1} = \int dr dp d\zeta e^{-H^T/k_B T} = c_c^{-1} \cdot c_\zeta^{-1}$$

Si noti che $\{r, p\}$ e ζ sono fra loro INDIPENDENTI:

$$\begin{aligned} p(r, p, \zeta) &= p^c(r, p) \cdot p^1(\zeta) \\ &= c_c e^{-\beta H} \cdot c_\zeta e^{-\beta Q \zeta^2/2} \end{aligned}$$

in cui c_c è l'normalizzazione del primo termine

22.3.89
11.12.88
6

2. E' chiaro che l'ensemble che ha a che fare con la forza del sistema non ha nulla a che fare con la variabile ζ . Perciò l'ensemble è:

$$\rho_{\text{ridotto}}(\{r, p\}) = \int d\zeta \rho(\{r, p\}, \zeta)$$

$$= \rho^c = c_c e^{-\beta H}$$

calcolo banale poiché $\{r, p\}$ e ζ sono variabili indipendenti. ζ è solo un accoppiamento di H con il resto del mondo e Q regola l'intensità dell'accoppiamento:

$Q \rightarrow \infty \Rightarrow \zeta \rightarrow 0$ e quindi se $\zeta(0) = 0$ il sistema ridiventa hamiltoniano e microcanonico;

$Q \rightarrow 0 \Rightarrow$ l'en. canonica è immediatamente riportata al suo valore medio, ma la

distribuzione è molto perturbata. Come che sia l'ensemble è quello canonico e, per le proprietà indipendenti dal tempo (che sono quelle di cui ci occupiamo in ore), il valore di Q non gioca alcun ruolo.

Ora:

$$c_c^{-1} = \int d\Gamma e^{-\beta H}$$

e

$$-\frac{A}{T} \equiv k_B \ln W_c = k_B \ln \int d\Gamma e^{-\beta H}$$

Da cui la formula in termodinamica

$$\dot{\vec{r}} = \frac{\vec{p}}{m} + \gamma(\vec{r} - \vec{r}_0)$$

$$\dot{\vec{p}} = \vec{F} - (\gamma + \delta)\vec{p}$$

$$\dot{\delta} = \frac{\hbar}{Q}(k - k_0)$$

$$\dot{\eta} = \frac{V}{M}(\vec{p}(t) \cdot \vec{p}_0)$$

$$\dot{V} = 3V\gamma$$

$$p(t) \frac{1}{3} \text{Tr}(\sigma(t)) =$$

$$= \frac{1}{3V} \langle \sum_i \left[\frac{p_i p_i}{m} + r_i \cdot F_i \right] \rangle$$

$$\mathcal{H}_c = \underbrace{K_N + V_N}_{\mathcal{H}_T} + \frac{1}{2} M \eta^2 + \frac{1}{2} Q \delta^2 + \underbrace{P_0 V}_{} + \int_0^t dt' \delta(t')$$

$$\beta \frac{d\mathcal{H}_T}{dt} = \sum_x \frac{\partial \mathcal{H}_T}{\partial x} = \frac{\partial \mathcal{H}_T}{\partial t}$$

$$f \propto e^{-\beta \mathcal{H}_T}$$

IL METODO MONTE CARLO IN MECCANICA STATISTICA CLASSICA

Vedremo in seguito che il ^{formalismo} ~~dei~~ ~~metodi~~ ~~per simulare~~ degli integrali funzionali o dei cammini di Feynman per sistemi quantistici a temperatura finita permette di mappare il problema della meccanica statistica d'equilibrio di un sistema quantistico su N-corpi, sul problema di certi sistemi classici equivalenti. E' quindi importante conoscere un metodo numerico che permetta di ottenere medie di Ensemble di sistemi classici a molti gradi di libert . Tale metodo   stato introdotto nel 1953 e va sotto il nome di Metodo Monte CARLO DI METROPOLIS

(Metropolis, Rosenbluth², Teller², J. Chem. Phys. 21, 1087). Prima di descrivere il metodo vediamo di comprendere il problema da risolvere.

Si tratta ^{per esempio} di ottenere stime di integrali di Bose (medie di ensemble) del tipo: (CANONICO).

$$\langle A \rangle = \frac{\int_{\Omega} d\vec{r}_1 \dots d\vec{r}_N e^{-\beta V_N(\vec{r}_1 \dots \vec{r}_N)} A(\vec{r}_1 \dots \vec{r}_N)}{\int_{\Omega} d\vec{r}_1 \dots d\vec{r}_N e^{-\beta V_N(\vec{r}_1 \dots \vec{r}_N)} = Z_N}$$

dove $\frac{e^{-\beta V_N}}{Z_N}$   la distribuzione di Boltzmann normalizzata

$$\int_{\Omega} d\vec{r}_1 \dots d\vec{r}_N \frac{e^{-\beta V_N}}{Z_N} = 1$$

La dimensione dello spazio delle configurazioni   Z^N . E' quindi chiaro che gi  per $N \geq 2$ tali integrali non possono essere ottenuti con i normali metodi di quadratura. In questo caso infatti la complessit  dell'algoritmo di integrazione a griglia regolare   ESPONENZIALE con N nel senso che il numero di punti griglia cresce esponenzialmente con N ($\sim M^{3N}$ ^{numero di punti per dimensione}).

Si ricorre quindi ad un metodo di campionamento stocastico dello spazio delle configurazioni: ~~questo  ~~ il metodo Monte Carlo.

~~Se~~ Se le funzioni integrate fossero regolari si potrebbe pensare di considerare N configurazioni ~~scelte a~~ ~~maniera~~ indipendenti distribuite uniformemente su Ω e trovare la media di A con

$$\langle A \rangle \approx \Omega^2 \left[\frac{1}{N} \sum_i P(i) A(i) \right]$$

Tuttavia sappiamo che la distribuzione di Boltzmann è molto irregolare essendo grande ~~estremo~~ per configurazioni con energie ~~stazionarie~~ vicine all'energia termodinamica di equilibrio e molto piccola (estremamente...)

per configurazioni con energia molto diversa dal valore termodinamico.

Inoltre sappiamo che le fluttuazioni di E/N decrescono come $1/\sqrt{N}$

quindi questa distribuzione diventa sempre più concentrata intorno alla superficie ad $E = \bar{E}(T, V, N)$ al crescere di N , ~~che significa~~ come intorno ~~una~~ ~~regione~~ ~~di~~ ~~dimensione~~ ~~che~~ ~~ad~~ ~~una~~ ~~regione~~ $3N-1$ dimensionale.

Il metodo Monte Carlo con campionamento semplice è quindi ~~estremamente~~ estremamente inefficiente per la Merc. Statistica. Si ricorre quindi al CAMPIONAMENTO DI IMPORTANZA (IMPORTANCE SAMPLING) con l'idea di

campionare ~~equale~~ configurazioni distribuite con la distribuzione di Boltzmann. L'idea dell'importance sampling ~~è da~~ ~~si può~~ spiegare

con il seguente esempio, Supponiamo di voler calcolare l'integrale

midheersvale: $I = \int_a^b dx f(x) \approx \frac{1}{L} \sum_i^L f(x_i) (b-a) \quad \% \text{ duetuo } \Rightarrow$

Posso anche descrivere questo integrale come:

$$\underline{I} = \int_{a_1}^b dx \frac{P(x)}{w(x)} \quad \text{au} \quad w(x) \neq 0 \quad x \in [a_1, b]$$

Assumiamo di sapere che $v(x) = u'(x)$ con $\begin{cases} u(x) \geq 0 \\ u'(x) \geq 0 \end{cases}$

$$\begin{cases} u(a) = 0 \\ u(b) = 1 \end{cases} \rightarrow \text{normalisierte} \\ \text{Dir}(x)$$

altre persone niscune:

Dato una distribuzione $p(x)$ $x \in [a, b]$: $p(x) \geq 0$

$$\int_a^b p(x) dx = 1$$

allora l'integrale $I = \int_a^b p(x) f(x) dx = \langle f \rangle_p$

può essere stutato da:

$$I \approx \frac{1}{M} \sum_{i=1}^M f(x_i) \quad \text{dove } x_i \text{ sono estratte da } p(x)$$

Nel nostro caso $I = \int_a^b f(x) dx$ e vogliamo esprimerlo con probabilità

$$p(x) = \frac{1}{b-a} \Rightarrow I = (b-a) \int_a^b \frac{f(x)}{(b-a)} dx \approx (b-a) \frac{1}{M} \sum_{i=1}^M f(x_i)$$

(82)

$$I = \int_0^1 du \frac{f(x(u))}{w(x(u))} \quad \text{usando la relazione inversa } x=x(u)$$

possiamo ora campionare uniformemente tra 0 e 1 la u e stimare l'integrale come:

$$I_{rw} \approx \frac{1}{L} \sum_{i=1}^L \frac{f(x(u_i))}{w(x(u_i))} \approx \left\langle \frac{f}{w} \right\rangle \quad \text{alla distribuzione uniforme tra } [0,1]$$

La varianza di questa stima ci dà capire cosa ci si può guadagnare a prendere $w(x)$ diversa da $1/(b-a)$.

La varianza di I_{rw} è:

$$\sigma_{I_{rw}}^2 = \frac{1}{L} \sigma_{(f/w)}^2 = \frac{1}{L} \left[\left\langle \left(\frac{f}{w} \right)^2 \right\rangle - \left\langle \frac{f}{w} \right\rangle^2 \right] \quad \text{COMPONENTI INDIPENDENTI}$$

quindi $\sigma_{I_{rw}}^2$ è sempre $O(1/L)$ però è ^{tanto} più piccola quanto più f/w è funzione smooth nell'intervallo $x \in [a,b]$. Al limite se $f/w = \text{costante} \Rightarrow \sigma_{I_{rw}}^2 = 0$.

Se invece usiamo il MC semplice con $w=1$ allora la varianza può diventare molto grande se $f(x)$ è molto irregolare.

Tornando alla meccanica statistica è chiaro che la migliore funzione di $w(x_1, \dots, x_N)$ è proprio la distribuzione di Boltzmann. Qui però il problema è che non sappiamo chi è l'unico integrale [equivalente di u di sopra]. Quindi dobbiamo sviluppare un algoritmo che riesca a campionare la distribuzione di Boltzmann. Se riusciamo a fare ciò allora le medie di fase diventano semplici medie aritmetiche delle osservabili su realizzazioni indipendenti del sistema:

$$\langle A \rangle \approx \frac{1}{N} \sum_{i=1}^N A(i)$$

↑ numero di configurazioni del sistema campionate con $e^{-\beta H(i)}$

ALGORITMO DI METROPOLIS

E' appunto la tecnica che, usando la teoria delle catene di Markov, permette di generare una sequenza di stati microscopici del sistema, e che a tempi lunghi sono distribuiti secondo la distribuzione di Boltzmann.

Per semplificare ~~la trattazione~~, e formalizzare pensiamo lo spazio delle configurazioni suddiviso in un numero grande ma finito di celle (S celle) così che da continuo diventa discreto. Lo stato del sistema è individuato dalla cella in cui cade il punto rappresentativo $i = 1, S$. Usiamo quindi la teoria delle catene di Markov discrete. [Nota: non è solo una semplificazione ma in realtà sul computer due punti che differiscono per meno della precisione della macchina sono ~~gli~~ lo stesso stato. Quindi la discretizzazione è molto nel calcolo numerico].

Per specificare stati successivi della catena introduciamo un tempo fittizio non fisico e diciamo che stati successivi si realizzano a istanti successivi.

DEFINIZIONE 1: PROVE INDIPENDENTI: Successione di prove indipendenti associate alle possibili prove successive sempre la stessa probabilità:

$E_k \rightarrow p_k$ e la probabilità di una certa sequenza è semplicemente il prodotto delle prob. delle singole realizzazioni:

$$P(E_0, \dots, E_n) = p_{j_0} p_{j_1} \dots p_{j_n}$$

2): CATENE DI MARKOV: è la generalizzazione più semplice delle sequenze di prove indipendenti.

Nelle catene di Markov stazionarie la probabilità di essere in certo numero all'istante i dipende solo dal valore assunto dalla catena al posto $i-1$. Non è più sufficiente specificare una probabilità del singolo stato (o ~~realizzazione~~ risultato della singola prova) ma bisogna ~~anche~~ specificare $\forall E_k$ la probabilità condizionata di E_k dato in qualunque E_j alla prova precedente.

$$p_{jk} = P(E_k | E_j)$$

Se chiamiamo a_k la probabilità di E_k a $t=0$, la probabilità di una realizzazione dell'intera catena $\in$ con $1, 2, 3, \dots, n$ passi è:

$$P(E_{j_0} E_{j_1}) = a_{j_0} p_{j_0 j_1}$$

$$P(E_{j_0} E_{j_1} E_{j_2}) = a_{j_0} p_{j_0 j_1} p_{j_1 j_2}$$

$\vdots$

$$(4) P(E_{j_0} \dots E_{j_n}) = a_{j_0} p_{j_0 j_1} \dots p_{j_{n-1} j_n}$$

CATENA DI MARKOV

Le matrici delle probabilità condizionate (matrici di transizione)

1) ha elementi positivi o nulli:

$$p_{jk} \geq 0$$

2) è normalizzata:

$$\sum_k p_{jk} = 1 \quad \forall j$$

$\Rightarrow P$ è una matrice STOCASTICA

DEFINIZIONE: la probabilità di transizione da E_j a E_k in n passi

si indica con $\phi_{jk}^{(n)}$ ed è:

$$\phi_{jk}^{(1)} = p_{jk}$$

$$\phi_{jk}^{(2)} = \sum_v p_{jv} p_{vk} = (P^2)_{jk}$$

$$\vdots$$

$$\phi_{jk}^{(n)} = \sum_v \phi_{jv}^{(n-1)} p_{vk} = (P^n)_{jk}$$

Se a_j è la probabilità dello stato E_j a $t=0$, la probabilità di avere E_k all' n -esima prova ($t=n$) è:

$$a_k^{(n)} = \sum_j a_j \phi_{jk}^{(n)}$$

~~Si aspetta che al crescere di n diminuisca la dipendenza da i .~~ Se $a_j = \delta_{ij} \Rightarrow a_k^{(n)} = (P^n)_{jk} = \phi_{jk}^{(n)}$

Ci si aspetta che al crescere di n diminuisca la dipendenza da i .

Questa aspettativa risulta confermata, sotto certe ipotesi, dal seguente th

che restringe la categoria di catene di Markov che si usano:

TEOREMA: In una catena IRRIDUCIBILE con solo STATI ERGODICI

i limiti:

$$(1) \quad u_k = \lim_{n \rightarrow \infty} p_{jk}^{(n)} \quad \forall k$$

esistono e sono indipendenti da j (stato iniziale). Inoltre:

$$(2) \quad u_k \geq 0$$

$$(3) \quad \sum_k u_k = 1$$

$$(4) \quad u_j = \sum_i u_i p_{ij}$$

ovvero u è autovettore sinistro di P con autovalore 1.

Viceversa una p_{ij} che individui una u_k che soddisfa le (2), (3), (4) e che abbia solo stati ergodici implica la (1)

N.B. L'ergodicità degli stati e l'irriducibilità si riducono rispettivamente, nel caso in cui il numero degli stati sia finito, alle seguenti proprietà:

- ergodici: a) ogni stato ricorre sicuramente e con tempo medio di ricorrenza $< \infty$
- irriducibili: b) $\forall (j, k) \exists n: p_{jk}^{(n)} > 0$ ovvero da ogni stato si può andare in ogni altro in un tempo finito.

Il significato del teorema, che non dimostriamo, può essere apprezzato considerando lo sviluppo della distribuzione iniziale a_j :

$$a_k^{(n)} = \sum_j a_j p_{jk}^{(n)} \xrightarrow{n \rightarrow \infty} \sum_j a_j u_k = u_k$$

Cioè per ogni distribuzione iniziale la distribuzione asintotica degli stati è proprio la u_k .

Poiché inoltre $a_j = u_j$ implica $a_k^{(n)} = u_k \quad \forall n$ (u_k = distribuzione invariante)

si può dire che, dopo un tempo sufficientemente lungo, gli stati saranno distribuiti con la distribuzione invariante u_k che è la distribuzione di equilibrio della catena.

L'algoritmo di Metropolis consiste proprio nel partire da una nota

$u_k = Z^{-1} e^{-\beta V(k)}$ e nel cercare una p_{ij} che abbia u_k come

distribuzione limite: $\lim_{n \rightarrow \infty} p_{ik}^{(n)} = u_k$

Per soddisfare la condizione (4) basta trovare una matrice che soddisfi la condizione di REVERSIBILITÀ MICROSCOPICA ossia:

$$(5) \quad P(E_j, E_k) = u_j p_{jk} = P(E_k, E_j) = u_k p_{kj} \Rightarrow \boxed{u_j p_{jk} = u_k p_{kj}} \\ \text{REVERS. MICROSCOPICA}$$

Infatti $\sum_k u_j p_{jk} = \sum_k u_k p_{kj} \Rightarrow u_j = \sum_k u_k p_{kj} \equiv (4)$ perché $\sum_k p_{jk} = 1$

Se ora gli stati sono ergodici e la p_{jk} è irriducibile (non si può dimostrare).
vale il teorema e la catena di Markov così costruita converge alla
distribuzione canonica nel senso che, asintoticamente, i vari stati
microscopici ricorrono con una frequenza proporzionale al fattore di
Boltzmann.

La condizione di reversibilità microscopica: $u_j p_{jk} = u_k p_{kj} \quad \forall (j, k)$
non determina unicamente le p_{jk} . Innanzitutto non conosciamo
né $u_k = e^{-F(E_k)}/Z$ perché non conosciamo la normalizzazione Z .

Infatti questo è proprio lo scopo del MC. La conoscenza di Z è equivalente
alla conoscenza di $F(T, V)$ ossia del comportamento termodinamico.

Conosciamo invece i rapporti tra probabilità di diversi stati: u_k/u_j

Il Metodo di Metropolis è una particolare scelta per ammorbidire
il BIANCO DETTAGLIATO. In questo metodo la probabilità di transizione
è ~~data~~ separata in una probabilità a "misuri" di cambiare lo stato
stato E_k a partire da un dato stato E_j : $T(j \rightarrow k)$
e una probabilità di accettazione $A(j \rightarrow k)$ con $0 \leq A \leq 1$ tale che

$$P_{jk} = T_{jk} A_{jk} \quad P(j \rightarrow k) = T(j \rightarrow k) A(j \rightarrow k)$$

la caratteristica di $T(j \rightarrow k)$ è che ~~questa~~ dobbiamo essere in grado
di campionare direttamente da questa distribuzione.

Nell'algoritmo di Metropolis generalizzato le mosse di prova sono accettate con probabilità:

$$A(j \rightarrow k) = \min [1, q(j \rightarrow k)]$$

$$q(j \rightarrow k) = \frac{u_k}{u_j} \frac{T(k \rightarrow j)}{T(j \rightarrow k)}$$

SCELTA DI METROPOLIS

E' facile verificare che questa scelta verifica il bilancio dettagliato

Si considerino i vari casi:

a) $j=k \Rightarrow A_{jk}=1 \quad T_{jk}=T_{jj} \quad u_j=u_k \Rightarrow ok$

b) $q_{jk}=1 \quad A_{jk}=1 \quad A_{kj}=1/q_{jk}=1 \Rightarrow u_j T_{jk} A_{jk} = u_k T_{kj} A_{kj}$

$\Rightarrow \frac{u_j T_{jk}}{u_k T_{kj}} = 1$ ma questa è proprio l'ipotesi $q_{jk}=1 \Rightarrow ok$

c) $q_{jk} < 1 \quad \left\{ \begin{array}{l} A_{jk} = q_{jk} \\ A_{kj} = 1 \quad (q_{kj} > 1) \end{array} \right. \quad u_j T_{jk} q_{jk} = \cancel{u_j T_{jk}} \frac{u_k T_{kj}}{\cancel{u_j T_{jk}}} = u_k T_{kj} \quad (ok)$

d) $q_{j \rightarrow k} > 1 \rightarrow A_{jk} = 1 \quad \Rightarrow u_j T_{jk} = u_k T_{kj} q_{kj} = u_k T_{kj} \frac{u_j T_{jk}}{u_k T_{kj}} \quad (ok)$
 $q_{k \rightarrow j} = \frac{1}{q_{jk}} < 1 \Rightarrow A_{kj} = q_{kj}$

ATTENZIONE: le mosse rifiutate vanno contate una volta in più. O meglio

quando hai rifiutato una mossa la configurazione di partenza deve essere contata una volta in più nelle medie. Questo perché la probabilità deve essere normalizzata ossia

$$P_{j \rightarrow j} = 1 - \sum_{k \neq j} P_{j \rightarrow k}$$

In pratica l'algoritmo è il seguente:

1) Data la distribuzione asymptotica u_k , decidere la probabilità di transizione a priori $T(j \rightarrow k)$

2) Scegliere una configurazione iniziale del sistema per cui $u_0 \neq 0$

3) Per arrivare allo stato da E_0 a E_1 si procede in 3 fasi:

3a) Calcolare $\tilde{E}_1$ da $T(E_0 \rightarrow \tilde{E}_1)$

3b) Calcolare il rapporto:

$$q = \frac{u_{\tilde{E}_1} T(\tilde{E}_1 \rightarrow E_0)}{u_{E_0} T(E_0 \rightarrow \tilde{E}_1)}$$

3c) Accettare o rifiutare:

- se $q > r$ dove r è un numero random distribuito uniformemente tra $(0,1)$ allora $E_1 = \tilde{E}_1$

- se $q < r$ $E_1 = E_0$

4) Gettare via i primi M stati che servono al sistema per raggiungere la distribuzione asymptotica

5) Calcolare medie di osservabili e fare l'analisi statistica dei dati

Scelta storica di Metropolis: sistema di N p.f.e. classiche.

La probabilità di transizione a priori è la mossa uniforme in un cubo di lato Δ di una p.f.e. $\Rightarrow T_{i \rightarrow j} = T_{j \rightarrow i} \Rightarrow q_{i \rightarrow j} = \frac{u_j}{u_i} = e^{-\beta[V(j) - V(i)]}$

$\Rightarrow$ MOSSE CHE DIMINUISCONO L'ENERGIA POTENZIALE SONO SEMPRE ACCETTATE

MOSSE CHE AUMENTANO L'ENERGIA POT. SONO ACCETTATE SE IL COSTO ENERGETICO (rispetto a kT) È PICCOLO

THE RULE OF THUMB è di scegliere Δ in modo che $A_{acc}/A_{tot} \approx 50\%$

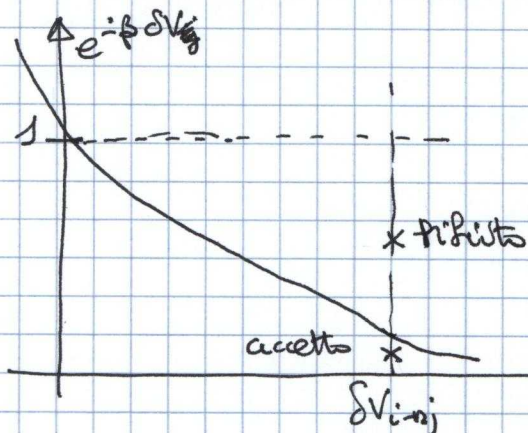
Se $A \ll 1$ sto sempre fermo, se $A \gg 1$ mi muovo sempre un po' di molto poco

$\Rightarrow A \ll 1, A \gg 1$ sono entrambi equivalenti.

Ultima cosa: perché il passo 3??

Nel caso di Metropolis è semplice capire graficamente questo passo.

In fatti abbiamo il seguente grafico: quindi ^{accettare} ~~completare~~ con



probab. $A_{i \to j} = \min[1, q_{i \to j}]$

significa, per $\delta V_{ij} > 0$

trovare con qualche regola δV_{ij}

e poi confrontare A_{ij} con un numero

random ~~da~~ $r \in (0,1)$

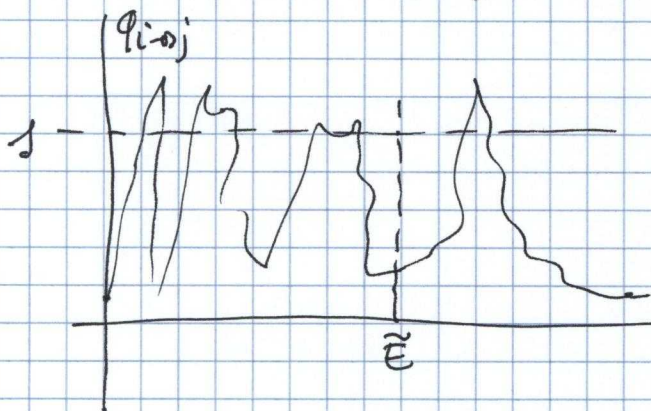
Se $r > A_{ij}$ rifiuto

$r < A_{ij}$ accetto

NOTA: con questa regola il metodo è capace di andare downhill ma anche uphill!!

superclub.phys.uniconed.it user:1 (user1) user_#

Nel caso più generale in cui $T_{ij} \neq T_{ji}$ posso immaginare a E_i fissato di plottare $q_{i \to j}$ per tutti gli stati j



Quindi altamente irregolare

Ora dato uno stato di prova

$\tilde{E}$ l'idea è la stessa;

devo accettare $\tilde{E}$ con probabilità

$A_{i \to \tilde{E}} = \min[1, q_{i \to \tilde{E}}]$

Se immagino di avere un gran numero N di atomi in cui da E_i provo

$\tilde{E}$ questo significa che il numero di volte $\sim N$ che devo accettare

è proporzionale a $q_{i \to \tilde{E}} \cdot N$ ossia se ho un numero a caso ho

$0 < 1$ e $1 > q_{i \to \tilde{E}}$ rifiuto e $1 < q_{i \to \tilde{E}}$ accetto.

CONVERGENZA IN PRATICA

Tale ^{la} convergenza statistica è garantita dal teorema, ~~la~~
la questione di fondamentale importanza è capire se lo spazio
delle configurazioni è esplorato dettagliatamente in un tempo di
calcolo ragionevole. Dobbiamo definire una misura del

TASSO DI CONVERGENZA e dell'EFFICIENZA di un dato processo di Markov.
Questo serve a confrontare diverse possibili scelte delle regole di
transizione, data una catena specifica a studiare quanto deve
durare il run e a ~~calcolare~~ stimare gli errori statistici sulle medie.

Il tempo di convergenza dipende dalle proprietà fisiche in esame:
ci saranno variabili veloci (energia, ...) e variabili lente (variabili
macroscopiche come il parametro d'ordine vicino ad una transizione di fase).

Consideriamo la proprietà O e chiamiamo O_k il suo valore
alla k -esima ^{stato} ~~realizzazione~~ della catena.

Il valore medio e la varianza di O sono definite da:

$$\left\{ \begin{array}{l} \bar{O} = \langle O_k \rangle_{\text{av}} \approx \frac{1}{K} \sum_{k=1}^K O_i \\ \sigma_O^2 = \langle (O_k - \bar{O})^2 \rangle_{\text{av}} \\ \approx \frac{1}{K} \sum_{k=1}^K (O_i - \bar{O})^2 \end{array} \right. \quad \begin{array}{l} \text{dipendono solo dalla distribuzione} \\ \text{di equilibrio e non dalle regole} \\ \text{di transizione.} \end{array}$$

L'errore statistico su $\bar{O}$ in una catena di Markov con P passi
è invece:

$$s_{\bar{O}} = \sqrt{\frac{\sigma_O^2}{P/k_0}}$$

dove ~~tempo di correlazione~~ P/k_0 è il numero di configurazioni indipendenti
lungo la catena.

k_0 è detto "tempo di correlazione" e può essere stimato dalla
conoscenza della funzione di correlazione temporale di O lungo la catena.

Supponiamo di aver fatto k passi della catena.

~~La~~ data una grandezza O la stima del suo valore medio dopo k passi è:

$$\bar{O}_k \approx \frac{1}{k} \sum_{i=0}^k O(i) \xrightarrow{k \rightarrow \infty} \langle O \rangle_u$$

La ~~varianza~~ stima della varianza dopo k passi è invece:

$$\sigma_{\bar{O}_k}^2 \approx \frac{1}{k} \sum_{i=0}^k [(O(i) - \bar{O}_k)^2] \xrightarrow{k \rightarrow \infty} \sigma_O^2$$

~~osservazione~~

Se la lunghezza della "DINAMICA" è statisticamente significativa $\bar{O}_k$ e $\sigma_{\bar{O}_k}^2$ dipendono solo dalla distribuzione limite (stazionaria) del processo e non dalle regole di transizione (come della dinamica stessa).

Siamo interessati a conoscere una stima dell'errore statistico nella nostra stima del valore medio dopo k passi.

La varianza di $\bar{O}_k$ sarà:

$$\sigma_{\bar{O}_k}^2 \approx \langle (\bar{O}_k - \langle O \rangle)^2 \rangle \approx \frac{1}{k^2} \sum_{i=0}^k \underbrace{[O(i) - \langle O \rangle][O(i) - \langle O \rangle]}_{\text{}} \rangle$$

No ci compare la funzione di correlazione di O lungo il processo.

$$C_O(i, j) = \langle [O(i) - \langle O \rangle][O(j) - \langle O \rangle] \rangle$$

Questa è una quantità dinamica e dipende quindi dalle regole di transizione della catena (o della dinamica microscopica...)

Se il processo è stazionario (come noi abbiamo solo stati di equilibrio)

~~si può mostrare~~ che la funzione di Autocorrelazione ha le seguenti proprietà:

- a) $C_O(i, j) = C_O(i - j)$ dipende solo dalla differenza dei tempi
- b) $C_O(i - j) = C_O(j - i)$ per un osservabile temporale

In queste condizioni:

$$J_{0n}^2 \approx \frac{1}{k^2} \sum_i^n \sum_j^n C_0(i-j) = \frac{k}{k^2} \left[\cancel{C_0(0)} \sum_{-k}^k C_0(e) \right] =$$

$$= \frac{1}{k} C_0(0) \left[1 + 2 \sum_1^k \frac{C_0(e)}{C_0(0)} \right] \approx C_0(0) \frac{k_0}{k}$$

la quantità K_0 è detta tempo di correlazione ^{represente} ~~il~~ ~~numero~~ ~~di~~ ~~configurazioni~~ che servono alla "dinamica" per perdere memoria del suo passato

$$k_0 \equiv 1 + 2 \sum_{e=1}^{\infty} e \cdot c_0(e) / c_0(0)$$

quindi se $k \in \mathbb{Z}$ è un qualsiasi $\mathbb{Z}$ intero di

quindi σ_K è maggiore della durata temporale delle
funzioni di correlazione la varianza della media è proprio

$$\sigma_0^2 = \frac{k_0}{k} \sigma_0 = \frac{k_0}{k} \sigma_0^2$$

la stima dell'errore alle medie è quindi

$$S_0 = \sqrt{\sigma_0^2 \frac{k_0}{k}} \Rightarrow \frac{k}{k_0} = \text{numero di configurazioni indipendenti lungo la traiettoria.}$$

The ratio is $\frac{1}{2}$ with the volume

BLOCK-AVERAGES

Quindi per conoscere l'errore alle medie bisogna in linea di principio calcolare le funzioni di correlazione dinamiche. Questo richiede un extra-work sia in CPU che in memoria che in programming. Vogliamo evitare questo e il metodo della media a blocchi ci permette ciò.

Quello che si fa è generare una traiettoria lunga, o sia molto più lunga del tempo di correlazione (almeno di quello che pensa in a il tempo di correlazione) ~~e poi si divide~~.

Dividiamo poi idealmente la traiettoria ~~di~~ di M permi in n_b blocchi di $p_b = M/n_b$ permi ciascuno.

Su ogni blocco calcoliamo la media di $\hat{\theta}_{\text{block}}$.

Arremo così ottenuto n_b stime di $\bar{O}_{P_b}^k = \frac{1}{P_b} \sum_{i=1}^{P_b} O(i, k)$ $k=1, \dots, n_b$

Se la lunghezza del blocco è $>$ di k_0 $P_0 > k_0$ allora le stime ~~dei~~ ottenute sono statisticamente indipendenti e la stima dell'errore z_0 è semplicemente

$$\sigma_{012} = \frac{1}{n_b} \sum_k \left(\frac{1}{n_b} - \langle O \rangle \right)^2$$

$$\langle 0 \rangle \approx \frac{1}{h_b} \sum_i^{h_b} \bar{\phi}_{p_b}^k$$

Il problema è che a priori non conosciamo K_0 e quindi non sappiamo con sicurezza se la lunghezza dei blocchi è ~~sufficiente~~ adeguata.

Per ~~porre~~ trovare P_0 definiamo la quantità:

$$f(p_b) = p_b \cdot \frac{\sigma_0^2}{\sigma_0^2}$$

dalla definizione data in precedenza $\sigma_o^2 = \frac{\sigma_o^2}{n} k_o$

vediamo che per $P_0 \gg$ del tempo di vita di $C_0(t)$

$$f(P_0) = k_0$$

Quindi se all'avventare della lingua di 'blochi'

$f(p_0)$ non cambia, e il valore a cui tende è una stima del tempo di autocorrelazione di 0, e il p_0 minimo a cui $f(p_0)$ è stazionaria è la minima lunghezza di blocco adeguata per ~~stima l'errore~~ ~~con~~ ~~basso~~ una stima non viziate dell'errore.

In pratica possiamo, date una traiettoria rettificata, scrivere un codice che rilegge i dati e li divide in blocchi di lunghezza sempre crescente fino a ~~terminare~~ raggiungere il plateau di $f(P_0)$.

Se il plateau non appare significa che M è troppo corto e bisognerà fare un run più lungo.