

3 L'equazione di Vlasov - derivazione euristica

3.1 Dall'equazione di Liouville alla gerarchia

Le equazioni cinetiche descrivono sistemi di molte particelle, come un gas, in cui l'incognita è la densità di probabilità di una particella nello spazio delle fasi, che indicherò con $f(t, \mathbf{x}, \mathbf{v})$ (o anche $f_t(\mathbf{x}, \mathbf{v})$). Da questa funzione si possono ottenere le misure di qualunque osservabile. In particolare

$$\rho(t, \mathbf{x}) = \int f(t, x, v) dv$$

è la densità di probabilità spaziale della particella. Se ci sono N particelle,

$$N \int_A \rho(t, \mathbf{x}) d\mathbf{x}$$

sarà (circa) il numero di particelle nella regione A .

Quello che però sappiamo di un sistema di N particelle è che la sua dinamica è governata dalle equazioni di Newton, che sono equivalenti all'equazione di Liouville, la cui incognita è la densità di probabilità $f^N(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{v}_1, \dots, \mathbf{v}_N)$, funzione di posizione e velocità di ogni particella. Nota f^N , otteniamo la distribuzione della probabilità della prima particella nello spazio integrando nelle variabili delle altre. Indico con $\mathbf{x}_k^N = (\mathbf{x}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_N)$ e uso l'analoga notazione \mathbf{v}_k^N .

$$f_1^N(\mathbf{x}_1, \mathbf{v}_1) := \int d\mathbf{x}_2^N d\mathbf{v}_2^N f^N(\mathbf{x}_1^N, \mathbf{v}_1^N)$$

Se invece integrassimo nelle variabili relative alla particella 1 e alle particelle 3, 4, \dots , N , otterremmo la distribuzione della seconda particella, e così via.

Richiami - Marginali

Sia \mathcal{X} un qualche insieme/spazio ragionevole. Sia $\mathbf{X} = (X_1, \dots, X_n)$ una variabile aleatoria in \mathcal{X}^n , e sia $\mu(d\mathbf{x}) = \mu(dx_1, \dots, dx_n)$ la sua legge, cioè per ogni $\phi \in C_b(\mathcal{X}^n)$

$$\mathbb{E}(\phi(\mathbf{X})) = \int \phi(\mathbf{x})\mu(d\mathbf{x}).$$

Posso pensare alle singole X_i come variabile aleatorie, in tal caso μ è una **distribuzione congiunta** di X_1, \dots, X_n . Posso chiedermi qual è la legge μ_i della singola X_i . Determinarla equivale a conoscere, per ogni $\alpha \in C_b(\mathcal{X})$, il valore atteso $\mathbb{E}(\alpha(X_i))$. Per definizione

$$\int \mu_i(dx)\alpha(x) = \mathbb{E}(\alpha(X_i)) = \int \alpha(x_i)\mu(d\mathbf{x})$$

In pratica μ_i si ottiene da μ integrando in tutte le variabili tranne x_i .

La misura di probabilità μ_i prende in nome di **i-marginale** di μ .

Si osservi che se $\Pi_i : \mathcal{X}^n \rightarrow \mathcal{X}$ è la i -esima proiezione, cioè la funzione che a $(x_1 \dots x_n)$ associa x_i , allora $\mu_i = \Pi_i \# \mu$.

Introduco una notazione frequente in teorie cinetiche e utile anche in altri contesti. Sia $g_1 : \Omega_1 \rightarrow \mathbb{R}$, e $g_2 : \Omega_2 \rightarrow \mathbb{R}$, la funzione "prodotto tensore" di g_1 e g_2 è definita su $\Omega_1 \times \Omega_2$ come

$$(g_1 \otimes g_2)(\mathbf{z}_1, \mathbf{z}_2) = g_1(\mathbf{z}_1)g_2(\mathbf{z}_2).$$

Nel caso $\Omega_1 = \Omega_2 = \Omega$, questa notazione serve a distinguere il prodotto di funzioni su Ω dato da $g_1(x)g_2(x)$, dal “prodotto” definito su $\Omega \times \Omega$ dato da $g_1(x_1)g_2(x_2)$.

Userò la scrittura compatta

$$g^{\otimes k} = g \otimes g \cdots \otimes g$$

per indicare $g(x_1)g(x_2) \dots g(x_n)$.

Analogamente, date due misure μ_1 su Ω_1 e μ_2 su Ω_2 , la misura $\mu_1 \otimes \mu_2$ è definita da

$$\int \phi(x_1, x_2)(\mu_1 \otimes \mu_2)(dx_1, dx_2) := \int \phi(x_1, x_2)\mu_1(dx_1)\mu_2(dx_2)$$

Ricordo infine che le variabili X_1, \dots, X_n sono **indipendenti** se

$$\mu(d\mathbf{x}) = \mu_1(dx_1) \cdots \mu_n(dx_n)$$

ovvero $\mu = \mu_1 \otimes \mu_2 \cdots \otimes \mu_n$. Infatti questa condizione è equivalente a

$$\mathbb{E}\left(\prod_{i=1}^n \phi_i(X_i)\right) = \prod_{i=1}^n \mathbb{E}(\phi_i(X_i))$$

per ogni scelta degli osservabili $\phi_1 \dots \phi_n$.

Particelle indistinguibili

Se vogliamo sperare di ottenere una descrizione di un sistema di N particelle usando solo la distribuzione di una particella, questa distribuzione non può dipendere dal nome della particella. La condizione affinché sia così è che f^N sia invariante per permutazione degli indici delle particelle. In fisica questa proprietà di f^N viene descritta come proprietà delle particelle, che in questo caso vengono dette **indistinguibili**. Per esempio se $N = 3$, sotto ipotesi di indistinguibilità

$$f_1^3(\mathbf{x}, \mathbf{v}) = \int d\mathbf{x}_2 d\mathbf{x}_3 d\mathbf{v}_2 d\mathbf{v}_3 f^3(\mathbf{x}, \mathbf{x}_2, \mathbf{x}_3, \mathbf{v}, \mathbf{v}_2, \mathbf{v}_3) = \int d\mathbf{x}_1 d\mathbf{x}_3 d\mathbf{v}_1 d\mathbf{v}_3 f^3(\mathbf{x}_1, \mathbf{x}, \mathbf{x}_3, \mathbf{v}_1, \mathbf{v}, \mathbf{v}_3)$$

etc., cioè f_1^3 è sia la distribuzione della prima particella, sia la distribuzione della seconda, sia la distribuzione della terza.

È dunque evidente che se conosco f^N conosco f_1^N . Se voglio che f_1^N rappresenti tutto il sistema devo invece poter ottenere f^N conoscendo f_1^N .

Caos molecolare

In generale fissare le marginali non permette di determinare la distribuzione, ma c'è un caso facile in cui questo accade, ed è quello in cui assumiamo che le particelle siano distribuite indipendentemente con una certa f . In tal caso, poiché per ogni i la particella i -esima ha densità di probabilità $f(\mathbf{x}_i, \mathbf{v}_i)$, per indipendenza sia ha che f^N è la distribuzione prodotto:

$$f^N(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{v}_1, \dots, \mathbf{v}_N) = \prod_{i=1}^N f(\mathbf{x}_i, \mathbf{v}_i).$$

Usando la notazione di prodotto tensore

$$f^N = f^{\otimes N}, \quad \text{con } f = f_1^N$$

(la condizione che f sia la marginale a una particella si ottiene integrando). Questa proprietà prende il nome di **caos molecolare**, ed equivale proprio a pensare che la coppia posizione - velocità delle particelle siano variabili aleatorie i.i.d..

Riassumendo, se il sistema è fatto di particelle identiche ed è in uno stato di caos molecolare, allora conoscere la distribuzione di una particella permette di descrivere l'intero sistema. La domanda che dobbiamo farci è se le equazioni che governano la dinamica consentono a queste due assunzioni di essere verificate. Le equazioni della meccanica permettono di descrivere l'evoluzione della funzione f^N , attraverso l'equazione di Liouville. Per comprendere se è possibile ottenere una descrizione "ridotta" del sistema, cioè attraverso la sola distribuzione di una particella, è necessario studiarla.

Interazioni di coppia

Considero un sistema di particelle di massa unitaria, interagenti a coppie con una forza di potenziale F di potenziale $U(\mathbf{x})$, cioè $\mathbf{F}(\mathbf{x}) = -\partial_x U(\mathbf{x})$.

$$\begin{cases} \dot{\mathbf{x}}_i = \mathbf{v}_i \\ \dot{\mathbf{v}}_i = \mathbf{F}_i = \sum_{j \neq i} \mathbf{F}(\mathbf{x}_i - \mathbf{x}_j) = -\sum_{j \neq i} \partial_x U(\mathbf{x}_i - \mathbf{x}_j) \end{cases} \quad (10)$$

Nei casi fisici più interessanti, U è il potenziale gravitazionale (attrattivo) o il potenziale coulombiano repulsivo, o un potenziale "empirico" di interazione molecolare come quello di Lennard-Jones. Qui per semplicità assumerò che U sia un potenziale simmetrico, repulsivo, a supporto compatto.

Poiché si tratta di un sistema meccanico con la forza posizionale, il sistema conserva la misura e l'equazione di Liouville coincide con l'equazione del trasporto:

$$\partial_t f^N + \sum_{i=1}^N \mathbf{v}_i \cdot \partial_{x_i} f^N + \sum_{i=1}^N \mathbf{F}_i \partial_{v_i} f^N = 0.$$

Lemma 3.1. *La dinamica preserva l'indistinguibilità. Si assuma che f^N al tempo $t = 0$ sia invariante per permutazioni degli indici delle particelle. Allora f^N è invariante a tutti i tempi.*

Dimostrazione per esercizio.

Per esercizio, si mostri che, invece, la dinamica non preserva il caos molecolare (si consideri per semplicità un sistema di sole due particelle). Il motivo è che se anche inizialmente le particelle sono indipendenti, appena interagiscono lo stato di una è correlato allo stato dell'altra, dunque l'indipendenza viene distrutta.

La gerarchia BBGKY

Per capire cosa accade, possiamo provare a scrivere le equazioni soddisfatte dalle distribuzioni marginali a j particelle, sotto l'ipotesi di invarianza per permutazioni. Per definizione, la marginale a j particelle è

$$f_j^N(\mathbf{x}_1^j, \mathbf{v}_1^j) = \int f^N(\mathbf{x}_1^N, \mathbf{v}_1^N) d\mathbf{x}_{j+1}^N d\mathbf{v}_{j+1}^N.$$

Si noti che per l'ipotesi di simmetria, considerare la marginale nelle prime j variabili, o nelle ultime j , o in qualsiasi sottoinsieme delle j particelle dà la stessa funzione. Supponendo che

U sia regolare e che il dato iniziale sia regolare a supporto compatto, si ottiene facilmente che f^N al tempo t è regolare e a supporto compatto. Questo fatto ci permetterà di scambiare integrali e derivate senza problemi.

Integro l'equazione di Liouville in $\mathbf{x}_{j+1}^N, \mathbf{v}_{j+1}^N$. Il termine $\partial_t f^N$ diventa $\partial_t f_j^N$. In $\sum_{i=1}^n \mathbf{v}_i \cdot \partial_{x_i} f^N$ i termini con $i \leq j$ e con $i > j$ danno contributi differenti: se $i > j$, scrivo $\mathbf{v}_i \cdot \partial_{x_i} f^N = \text{div}_{x_i}(\mathbf{v}_i f^N)$ il cui integrale nella variabile $d\mathbf{x}_i$ è nullo perché f^N è a supporto compatto. Se $i \leq j$, posso scambiare l'ordine di integrazione e derivazione. Complessivamente, il termine di trasporto dà

$$\sum_{i=1}^j \mathbf{v}_i \cdot \partial_{x_i} f_j^N.$$

Resta da capire cosa accade al termine con le forze, che è

$$\sum_{i=1}^N \sum_{h \neq i} \int \mathbf{F}(\mathbf{x}_i - \mathbf{x}_h) \cdot \partial_{v_i} f^N d\mathbf{x}_1^j d\mathbf{v}_1^j.$$

Gli integrali dei termini con $\partial_{v_i}, i > j$ sono nulli, perché f^N è a supporto compatto. Rimane dunque solo la somma per i da 1 a j . Separo la somma su h in $h \leq j$ e $h > j$. La prima somma dà l'analogo a j particelle del termine presente nell'equazione di Liouville:

$$\sum_{i=1}^j \sum_{h \neq i} \mathbf{F}(\mathbf{x}_i - \mathbf{x}_h) \partial_{v_i} f_j^N.$$

L'altro termine contiene, per $i \leq j$, la somma per $h > j$ di

$$\int \mathbf{F}(\mathbf{x}_i - \mathbf{x}_k) \cdot \partial_{v_i} f^N dx_{j+1}^N dv_{j+1}^N$$

Questa espressione è indipendente dall'indice k per l'invarianza per permutazioni. Fisso dunque $k = j + 1$, e la somma su h dà $N - j$ termini uguali a

$$\int \mathbf{F}(\mathbf{x}_i - \mathbf{x}_{j+1}) \cdot \partial_{v_i} f_{j+1}^N dx_{j+1} dv_{j+1}.$$

In conclusione, l'equazione per la marginale a j particelle è

$$\begin{aligned} \partial_t f_j^N + \sum_{i=1}^j \mathbf{v}_i \cdot \partial_{x_i} f_j^N + \sum_{i=1}^j \sum_{h \neq i} \mathbf{F}(\mathbf{x}_i - \mathbf{x}_h) \cdot \partial_{v_i} f_j^N + \\ + \sum_{i=1}^j (N - j) \int \mathbf{F}(\mathbf{x}_i - \mathbf{x}_{j+1}) \partial_{v_i} f_{j+1}^N(\mathbf{x}_1^{j+1}, \mathbf{v}_1^{j+1}) d\mathbf{x}_{j+1} d\mathbf{v}_{j+1} = 0. \end{aligned}$$

In particolare l'equazione per la distribuzione a una particella è

$$\partial_t f_1^N + \mathbf{v}_1 \cdot \partial_{x_1} f_1^N + (N - 1) \int \mathbf{F}(\mathbf{x}_1 - \mathbf{x}_2) \partial_{v_1} f_2^N(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2) d\mathbf{x}_2 d\mathbf{v}_2 = 0.$$

Come si vede, l'equazione per j particelle coinvolge l'equazione per $j + 1$ particelle, quindi non abbiamo ottenuto equazioni chiuse, ma solo una **gerarchia** di equazioni, che prende il nome di gerarchia BBGKY (da Bogoliubov e altri quattro che l'hanno ricavata indipendentemente). L'ultima equazione, quella per $j = N$, è l'equazione di Liouville, che invece è un'equazione chiusa.

La speranza di ottenere una descrizione ridotta del sistema, cioè un'equazione chiusa in f_1 , è legata alla possibilità di soddisfare i due seguenti aspetti:

- l'ultimo termine deve essere di ordine $O(N^0)$
- $f_2(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2)$ deve potersi scrivere in termini di f_1 .

Quest'ultima condizione si verifica nel caso di particelle indipendentemente e identicamente distribuite, perché in tal caso f_2 si fattorizza in $f_1(\mathbf{x}_1, \mathbf{v}_1)f_1(\mathbf{x}_2, \mathbf{v}_2)$. Però assumere che il dato iniziale è fattorizzato, cioè che $f^N(\mathbf{x}_1^N, \mathbf{v}_1^N, 0) = \prod_{i=1}^N f(\mathbf{x}_i, \mathbf{v}_i, 0)$ non è sufficiente: come già osservato, l'interazione crea correlazione tra le particelle.

Sunto

- La descrizione "cinetica" di un sistema si ottiene se si riesce a descrivere la dinamica di un sistema di N particelle con un'opportuna dinamica di una singola particella.
- Condizione preliminare per poter avere una descrizione cinetica è l'**indistinguibilità** delle particelle, che, se assunta al tempo 0, è vera per tutti i tempi.
- Poter descrivere un sistema a livello cinetico equivale all'assunzione di **caos molecolare**, cioè che la legge delle N particelle fattorizza.
- La gerarchia BBGKY è il sistema di equazioni integro-differenziali soddisfatto dalle distribuzioni marginali a j -particelle, e mostra come il caos molecolare non può essere assunto.

3.2 Equazioni di campo medio

Nel cosiddetto limite di campo medio, si assume che l'intensità dell'interazione sia di ordine $1/N$. Il sistema di particelle viene modificato in

$$\begin{cases} \dot{\mathbf{x}}_i = \mathbf{v}_i \\ \dot{\mathbf{v}}_i = \mathbf{F}_i = \frac{1}{N} \sum_{j \neq i} \mathbf{F}(\mathbf{x}_i - \mathbf{x}_j) \end{cases} \quad (11)$$

Nel limite $N \rightarrow +\infty$, la correlazione che si crea tra una particella e l'altra va a zero come $1/N$, dunque ci si può aspettare che la gerarchia limite supporti soluzioni fattorizzate (cioè nel limite le particelle sono indipendenti). D'altra parte il contributo totale delle interazioni è di ordine 1, dunque la dinamica sarà non banale.

L'equazione per f_1^N è

$$\partial_t f_1^N(t, \mathbf{x}_1, \mathbf{v}_1) + \mathbf{v}_1 \cdot \partial_{\mathbf{x}_1} f_1^N(t, \mathbf{x}_1, \mathbf{v}_1) + (N-1) \frac{1}{N} \int \mathbf{F}(\mathbf{x}_1 - \mathbf{x}_2) \cdot \partial_{\mathbf{v}_1} f_2^N(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2) d\mathbf{x}_2 d\mathbf{v}_2 = 0,$$

Si vede che se N tende a infinito si ottiene un'espressione che non dipende da N , che rimane non banale.

La gerarchia di Boltzmann e l'equazione di Vlasov

Supponiamo, formalmente, che le f_j^N convergono a delle funzioni f_j . Passando al limite, si ottiene la **gerarchia di Boltzmann**, una gerarchia di equazioni che non dipendono più da N :

$$\partial_t f_j(t, \mathbf{x}_1^j, \mathbf{v}_1^j) + \sum_{i=1}^j \mathbf{v}_i \cdot \partial_{\mathbf{x}_i} f_j(t, \mathbf{x}_1^j, \mathbf{v}_1^j) + \sum_{i=1}^j \int \mathbf{F}(\mathbf{x}_i - \mathbf{x}_{j+1}) \partial_{\mathbf{v}_i} f_{j+1}(\mathbf{x}_1^{j+1}, \mathbf{v}_1^{j+1}) d\mathbf{x}_{j+1} d\mathbf{v}_{j+1} = 0.$$

Si osservi che il contributo di interazione tra le coppie di particelle è svanito, perché di ordine $1/N$, e ne rimane traccia solo nell'ultimo termine.

Il fatto nuovo è che queste equazioni ammettono una soluzione fattorizzata, come mostreremo.

Lemma 3.2. *La gerarchia di Boltzmann ammette soluzioni fattorizzate $f_j = f^{\otimes j}$, dove $f = f(t, \mathbf{x}, \mathbf{v})$ risolve l'equazione*

$$\partial_t f(\mathbf{x}, \mathbf{v}) + \mathbf{v} \cdot \partial_{\mathbf{x}} f(\mathbf{x}, \mathbf{v}) + \left(\int \mathbf{F}(\mathbf{x} - \mathbf{y}) f(\mathbf{y}, \mathbf{w}) d\mathbf{y} d\mathbf{w} \right) \cdot \partial_{\mathbf{v}} f(\mathbf{x}, \mathbf{v}) = 0. \quad (12)$$

Chiamerò l'equazione 12 **equazione di Vlasov** (anche se questo nome si riferisce più precisamente a specifici modelli individuati dall'espressione della forza F). Si osservi che non è un'equazione lineare.

Dimostrazione. Se la gerarchia ammette soluzioni fattorizzate, l'equazione per la prima marginale deve essere un'equazione chiusa. Infatti, chiamando $f_1 = f$ e imponendo $f_2 = f \otimes f$, l'equazione per la prima marginale diventa proprio l'equazione di Vlasov.

Inseriamo l'espressione $f_j = f^{\otimes j}$ nell'equazione per la marginale j -esima. Indico con $\check{f}_i = \prod_{h, h \neq i} f(t, \mathbf{x}_h, \mathbf{v}_h)$.

Il termine $\partial_t f^{\otimes j}$ si scrive come $\sum_{i=1}^j \check{f}_i \partial_t f(t, \mathbf{x}_i, \mathbf{v}_i)$.

Il termine $\sum_{i=1}^j \mathbf{v}_i \cdot \partial_{\mathbf{x}_i} f^{\otimes j}$ si scrive come $\sum_{i=1}^j \check{f}_i \mathbf{v}_i \cdot \partial_{\mathbf{x}_i} f(t, \mathbf{x}_i, \mathbf{v}_i)$.

Infine, $\sum_{i=1}^j \int \mathbf{F}(\mathbf{x}_i - \mathbf{x}_{j+1}) \partial_{\mathbf{v}_i} f^{\otimes(j+1)}(\mathbf{x}_1^{j+1}, \mathbf{v}_1^{j+1}) d\mathbf{x}_{j+1} d\mathbf{v}_{j+1}$ diventa

$$\sum_{i=1}^j \check{f}_i \partial_{\mathbf{v}_i} f(t, \mathbf{x}, \mathbf{v}) \int \mathbf{F}(\mathbf{x}_i - \mathbf{y}) f(t, \mathbf{y}, \mathbf{w}) d\mathbf{y} d\mathbf{w}$$

Dunque l'equazione per la j -esima marginale è soddisfatta perché ha la forma

$$\sum_{i=1}^j \check{f}_i \times (\text{equazione di Vlasov per } f(t, \mathbf{x}_i, \mathbf{v}_i)) = 0$$

□

La procedura di limite che ci dà l'equazione di Vlasov prende il nome di **limite di campo medio**. Per spiegare perché devo fare qualche richiamo.

Richiami - Legge dei grandi numeri

La legge dei grandi numeri afferma che date N variabili *i.i.d.* (cioè indipendenti e identicamente distribuite) \mathbf{x}_i di valore atteso \mathbf{m} , la **media empirica**

$$\frac{1}{N} \sum_i \mathbf{x}_i$$

tende, in un qualche senso probabilistico, al valore atteso \mathbf{m} .

Osservo che la tesi si ottiene mostrando che la media empirica delle variabili $\mathbf{z}_i = \mathbf{x}_i - \mathbf{m}$, *i.i.d.* di media nulla, tende a 0, infatti

$$\frac{1}{N} \sum_i \mathbf{x}_i - \mathbf{m} = \frac{1}{N} \sum_i (\mathbf{x}_i - \mathbf{m}) = \frac{1}{N} \sum_i \mathbf{z}_i$$

Teorema 3.3. *Legge debole dei grandi numeri.*

Siano \mathbf{z}_i , $i = 1 \dots N$, variabili aleatorie i.i.d. di media nulla e varianza σ . Allora il valore atteso di $\frac{1}{N} |\sum_i \mathbf{z}_i|$ tende a 0. Come conseguenza, per ogni $\varepsilon > 0$, $\mathbb{P}(\frac{1}{N} |\sum_i \mathbf{z}_i| > \varepsilon) \rightarrow 0$.

Dimostrazione. Indico con $\mu(d\mathbf{z})$ la legge comune delle variabili \mathbf{z}_i ,

La media empirica $\bar{\mathbf{z}} := \frac{1}{N} \sum_i \mathbf{z}_i$ è una variabile aleatoria di media nulla e, per indipendenza delle \mathbf{z}_i , di varianza pari alla somma delle varianze, cioè σ^2/N . Dunque

$$\mathbb{E}(|\bar{\mathbf{z}}|) = \int |\bar{\mathbf{z}}| \mu^{\otimes N}(d\mathbf{z}_1^N) = \int_{|\bar{\mathbf{z}}| < \varepsilon} |\bar{\mathbf{z}}| \mu^{\otimes N}(d\mathbf{z}_1^N) + \int_{|\bar{\mathbf{z}}| \geq \varepsilon} |\bar{\mathbf{z}}| \mu^{\otimes N}(d\mathbf{z}_1^N)$$

il primo termine è minore di ε , il secondo lo stimo con Chebychev: poiché $|\bar{\mathbf{z}}| \geq \varepsilon$

$$\int_{|\bar{\mathbf{z}}| \geq \varepsilon} |\bar{\mathbf{z}}| \mu^{\otimes N}(d\mathbf{z}_1^N) \leq \frac{1}{\varepsilon} \int |\bar{\mathbf{z}}|^2 \mu^{\otimes N}(d\mathbf{z}_1^N) = \frac{1}{\varepsilon} \frac{1}{N} \sigma^2$$

Scegliendo $\varepsilon = \sigma N^{-1/2}$ si ottiene che $\mathbb{E}(|\bar{\mathbf{z}}|) \leq 2\sigma/N^{1/2}$, da cui la tesi.

Infine, dato $\varepsilon > 0$,

$$\mathbb{P}(|\bar{\mathbf{z}}| > \varepsilon) = \int \mathcal{X}\{|\bar{\mathbf{z}}| > \varepsilon\} \mu^{\otimes N}(d\mathbf{z}_1^N) \leq \frac{1}{\varepsilon} \int |\bar{\mathbf{z}}| \mu^{\otimes N}(d\mathbf{z}_1^N) = \frac{1}{\varepsilon} \mathbb{E}(|\bar{\mathbf{z}}|) \rightarrow 0.$$

□

Si può anche dimostrare, ma non lo farò, che vale la legge forte dei grandi numeri, cioè che $\bar{\mathbf{z}}$ converge a 0 quasi certamente (cioè a meno di eventi di probabilità nulla).

Consideriamo ora N particelle i.i.d. con densità $f(\mathbf{x}, \mathbf{v})$. La variabile aleatoria $\mathbf{F}(\mathbf{x} - \mathbf{x}_i)$ ha come media $\int \mathbf{F}(\mathbf{x} - \mathbf{y}) f(d\mathbf{y}, d\mathbf{w})$. La legge forte dei grandi numeri assicura che

$$\frac{1}{N} \sum_{i=1}^N \mathbf{F}(\mathbf{x} - \mathbf{x}_i) \rightarrow \int \mathbf{F}(\mathbf{x} - \mathbf{y}) f(\mathbf{y}, \mathbf{w}) d\mathbf{y} d\mathbf{w}.$$

Dunque, formalmente, nel limite di campo medio il sistema differenziale per ogni singola particella diventa

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{v} \\ \dot{\mathbf{v}} = \int \mathbf{F}(\mathbf{x} - \mathbf{y}) f(t, \mathbf{y}, \mathbf{w}) d\mathbf{y} d\mathbf{w} \end{cases}$$

Considerando f nota, si può scrivere l'equazione di Liouville per questo sistema e si ottiene esattamente l'equazione di Vlasov per f .

Sunto

- **Le interazioni di campo medio** sono interazioni di coppia, con intensità scalata di $1/N$, dove N è il numero di particelle, dunque la correlazione tra particelle, anche se si crea, svanisce nel limite $N \rightarrow +\infty$.
- Nel limite $N \rightarrow +\infty$ la gerarchia che si ottiene ammette soluzioni fattorizzate, in cui il singolo fattore soddisfa l'equazione di Vlasov.
- La forza che agisce su ogni particella tende alla forza “media”, e l'equazione di Liouville per il sistema limite è l'equazione di Vlasov.

3.3 Altri modelli

Prima di discutere le proprietà matematiche dell'equazione di Vlasov e come si ottenga dal sistema di particelle, introduco qualche altro esempio, in ambito non fisico. Da vari decenni vengono studiati modelli per descrivere il comportamento di folle, il moto di sciami o greggi, o anche come cambiano le opinioni in un collettivo umano (“dinamica delle opinioni”), o come si comportano gli agenti sui mercati finanziari. In questi modelli le variabili riguardano gli “agenti” (posizione e velocità di uccello, opinione o disponibilità economica di un individuo), e queste variabili cambiano per interazione tra gli agenti. Spesso gli agenti sono tutti uguali, e si differenziano solo per il valore delle variabili. Gli aspetti più affascinanti di questi modelli sono le cosiddette **proprietà emergenti**, cioè i comportamenti coordinati a livello collettivo, che non vengono imposte nel modello ma sono conseguenza delle dinamiche microscopiche (in questo senso “emergono” dalla dinamica degli agenti). Fenomeni di questo tipo sono l'allineamento degli uccelli in volo: non esiste una “regia” esterna che dice ai singoli uccelli cosa fare, ma il fatto che si allineano deve essere conseguenza delle regole dell'interazione tra essi. Un fenomeno dello stesso tipo è l'emergere del consenso su una particolare opinione (si pensi all'esplosione che sembra improvvisa di una moda nel vestire, o nel parlare, o nel gusto).

Illustro alcuni di questi modelli differenziali.

3.3.1 Il modello per il flocking di Cucker e Smale

In questo caso gli agenti sono uccelli (o più in generale animali), le variabili sono posizione e velocità ($\mathbf{x}_i, \mathbf{v}_i$) e l'interazione è fatta in modo che ogni uccello modifica la sua velocità per uniformarsi alla velocità degli uccelli vicini.

$$\begin{aligned}\dot{\mathbf{x}}_i &= \mathbf{v}_i \\ \dot{\mathbf{v}}_i &= \sum_j p_{ij}(\mathbf{v}_j - \mathbf{v}_i)\end{aligned}\tag{13}$$

Il contributo dell'agente j all'accelerazione dell'agente i ha intensità $p_{ij}|\mathbf{v}_j - \mathbf{v}_i|$ e fa decrescere la differenza di velocità (si pensi alla soluzione dell'equazione $\dot{\mathbf{v}} = p(\mathbf{w} - \mathbf{v})$, con \mathbf{w} fissato: si ottiene una convergenza esponenziale di $\mathbf{v}(t)$ al valore \mathbf{w}). I coefficienti (positivi o nulli) p_{ij} formano la **matrice di comunicazione**.

Nel modello di Cucker e Smale questi coefficienti sono del tipo

$$p_{ij} = \frac{1}{1 + |\mathbf{x}_i - \mathbf{x}_j|^\alpha}$$

per qualche $\alpha > 0$. In questo modo, la velocità dell'agente i tende ad allinearsi a quella degli uccelli vicini, mentre è poco influenzata dagli uccelli lontani.

Si noti che se p_{ij} è simmetrico, allora la velocità media si conserva (lo si provi per esercizio). Si può dimostrare che se α è abbastanza piccolo, asintoticamente le velocità diventano uguali alla velocità media, e gli uccelli si muovono su traiettorie parallele, mantenendo distanze fisse tra loro.

In modelli di opinioni, si possono considerare invece sistemi di questo tipo nelle sole variabili \mathbf{v} , con opportune matrici di comunicazione. In questo caso l'allineamento delle \mathbf{v}_i corrisponde all'emergere del consenso.

3.3.2 Il modello di Kuramoto per la sincronizzazione

Il fenomeno della sincronizzazione spontanea è molto affascinante e ne esistono esempi naturali e artificiali: potete cercare in rete informazioni e video su lucciole e metronomi che si sincronizzano.

Ogni agente è descritto da una variabile periodica (il tempo per l'alternarsi acceso / spento di una lucciola, o la posizione sull'orbita chiusa dell'asta del metronomo). Assumendo che tutti gli agenti siano uguali, possiamo usare delle **fasi** cioè delle variabili angolari ϑ_i di periodicità 2π . Nel caso più semplice, il modello di Kuramoto è

$$\dot{\vartheta}_i = \omega - \frac{1}{N} \sum_j \sin(\vartheta_i - \vartheta_j)$$

Quando la differenza tra $\vartheta_i - \vartheta_j$ il contributo dell'agente j alla velocità della fase dell'agente i è nella direzione di far diminuire la differenza.

Si può provare che a meno di dati iniziali molto particolari, asintoticamente $\dot{\vartheta}_i \rightarrow \omega$, e $|\vartheta_i - \vartheta_j| \rightarrow 0$, realizzando così la sincronizzazione.

Si noti che le variabili $\gamma_i = \vartheta_i - t\omega$ verificano

$$\dot{\gamma}_i = -\frac{1}{N} \sum_j \sin(\gamma_i - \gamma_j)$$

Questo è un **sistema gradiente**, infatti

$$\begin{aligned} \dot{\gamma}_i &= -\partial_{\gamma_i} U(\boldsymbol{\gamma}) \\ U(\boldsymbol{\gamma}) &= -\frac{1}{2} \sum_{i,j} \cos(\gamma_i - \gamma_j) \end{aligned}$$

(lo si provi per esercizio). Dunque U decresce lungo il moto (ed è una funzione di Liapunov per il sistema). Si noti però che la funzione U ha infiniti minimi globali, che si realizzano per valori coincidenti delle γ_i , e anche infiniti punti stazionari che non sono minimi (quelli in cui γ_i coincidono o differiscono per π , cioè sono opposti).

3.3.3 Un algoritmo per la determinazione dei minimi

In moltissime situazioni applicative è necessario determinare il minimo globale di una funzione $U(\mathbf{x})$. In casi importanti (come l'addestramento delle reti neurali) \mathbf{x} può vivere in uno spazio di dimensione enorme, e U può avere un panorama complesso di punti di minimo locale.

Esistono molti algoritmi per risolvere questo problema, per esempio si può considerare la dinamica

$$\dot{\mathbf{x}} = -\partial_x U$$

che garantisce che U decresce. Ma a seconda del dato iniziale, il sistema finisce in un minimo locale. In genere viene aggiunta una stocasticità che permette alla traiettoria di sfuggire dai minimi troppo alti (“simulated annealing”). Un altro possibile modo per curare questo problema è avere molti agenti che esplorano lo spazio dei minimi, e che tendono a richiamarsi tra loro:

$$\dot{\mathbf{x}}_i = -\partial_x U(\mathbf{x}_i) - \sum_j p_{ij}(\mathbf{x}_i - \mathbf{x}_j).$$

In questo modo è possibile che l’agente che sta nel minimo più profondo, da cui dunque esce con maggiore difficoltà, riesca a richiamare gli altri agenti, permettendo una rapida individuazione del minimo.

3.4 Una descrizione unificata

Tutti i modelli che ho considerato, in cui l’interazione è di ordine $1/N$ possono essere scritti nella forma

$$\dot{\mathbf{z}}_i = \mathbf{H}(\mathbf{z}_i) + \frac{1}{N} \sum_j \mathbf{Q}(\mathbf{z}_i, \mathbf{z}_j)$$

dove \mathbf{H} è la componente di singola particella, mentre \mathbf{Q} tiene conto dell’interazione. Per esercizio si scrivano \mathbf{z} , \mathbf{H} e \mathbf{Q} nel caso del sistema di particelle, per Cucker-Smale, per Kuramoto, per il sistema per la ricerca dei minimi. È facile osservare che questa notazione può essere ulteriormente compattata, nella forma

$$\dot{\mathbf{z}}_i = \frac{1}{N} \sum_j \mathbf{K}(\mathbf{z}_i, \mathbf{z}_j)$$

dove \mathbf{K} tiene conto sia dei termini di singola particella, sia di quelli di interazione. Basta infatti scegliere $\mathbf{K}(\mathbf{z}, \mathbf{w}) = \mathbf{Q}(\mathbf{z}, \mathbf{w}) + \mathbf{H}(\mathbf{z})$.

La corrispondente equazione di Liouville è

$$\partial_t f^N(t, \mathbf{z}_1^N) + \sum_{i=1}^N \operatorname{div}_{z_i} \left(\frac{1}{N} \sum_{h=1}^N \mathbf{G}(\mathbf{z}_i, \mathbf{z}_h) f^N(t, \mathbf{z}_1^N) \right) = 0$$

La gerarchia è

$$\partial_t f_j^N(t, \mathbf{z}_1^j) + \sum_{i=1}^j \operatorname{div}_{z_i} \left(\frac{N-j}{N} \int d\mathbf{z}_{j+1} \mathbf{G}(\mathbf{z}_i, \mathbf{z}_{j+1}) f_{j+1}^N(t, \mathbf{z}_1^{j+1}) \right) = 0$$

Nel limite $N \rightarrow +\infty$, se le marginali convergono, si ottiene la gerarchia infinita:

$$\partial_t f_j(t, \mathbf{z}_1^j) + \sum_{i=1}^j \operatorname{div}_{z_i} \left(\int d\mathbf{z}_{j+1} \mathbf{G}(\mathbf{z}_i, \mathbf{z}_{j+1}) f_{j+1}(t, \mathbf{z}_1^{j+1}) \right) = 0$$

Ipotizzando che

$$f_j^N \rightarrow f^{\otimes j}$$

tutte le equazioni della gerarchia sono soddisfatte se è soddisfatta la prima, che diventa

$$\begin{aligned}\partial_t f_t + \operatorname{div}_z (f_t \mathcal{K}[f_t]) &= 0 \\ \mathcal{K}[f](\mathbf{z}) &:= \int \mathbf{G}(\mathbf{z}, \tilde{\mathbf{z}}) f(\tilde{\mathbf{z}}) = 0\end{aligned}\tag{14}$$

che è appunto l'equazione di campo medio associata al sistema di partenza.

Per discutere dell'esistenza e unicità delle soluzioni di questa equazione e se sia effettivamente derivabile dal modello microscopico ci serve esplorare la nozione di convergenza debole delle misure.