


DIPARTIMENTO
DI MATEMATICA



SAPIENZA
UNIVERSITÀ DI ROMA

IFM 2025-2026

Introduzione alle equazioni cinetiche

© 2026 di Dario Benedetto con licenza attribuzione - non commerciale
- condividi allo stesso modo 4.0 internazionale CC BY-NC-SA 4.0 

Dario Benedetto - <http://brazil.mat.uniroma1.it/dario>

Sapienza Università di Roma
Dipartimento di Matematica
Piazzale Aldo Moro n. 5, 00185 Roma
www.mat.uniroma1.it

IFM 2025/2026: introduzione alle equazioni cinetiche

3 aprile 2026

Queste note sono “appunti” e non “dispense”: vuol dire che trovate gli argomenti e le dimostrazioni, ma le motivazioni sono piuttosto stringate.

Possibile materiale di approfondimento

- Dispense di E. Caglioti <https://sites.google.com/site/ecaglioti/didattica/MR> su
 - Gerarchia e equazione di Vlasov in “Dispense III”
 - Limite di campo medio in “Il limite di campo medio”
 - Equazione di Boltzmann e teorema H in “Dispense III”
 - Limite idrodinamico per l’equazione di Boltzmann in “Hydrodynamic limit for the Boltzmann Equation”

- C. Cercignani, R. Illner, M. Pulvirenti **The Mathematical Theory of Dilute Gases** Applied Mathematical Sciences 106 - Springer

- Sul problema del passaggio da una descrizione irreversibile a una reversibile potete guardare sul testo precedente, e anche informarvi sul “modello di Kac”, un modellino che contiene molte delle difficoltà concettuali del passaggio al limite da sistemi reversibili a sistemi irreversibili.

G.A. Gottwald, M. Oliver: *Boltzmann’s Dilemma: An Introduction to Statistical Mechanics via the Kac Ring* SIAM Review **51** 3, pp 613–635 2009

<https://epubs.siam.org/doi/pdf/10.1137/070705799>

- C. Villani: *Topics in Optimal Transportation* Graduate Studies in Mathematics vol. 58, Springer 2003

This is the first comprehensive introduction to the theory of mass transportation with its many—and sometimes unexpected—applications. In a novel approach to the subject, the book both surveys the topic and includes a chapter of problems, making it a particularly useful graduate textbook.

- F. Santambrogio: *Optimal Transport for Applied Mathematicians* PNLDE vol. 87, Birkhäuser 2015

Testo che contiene anche sviluppi più recenti, in particolare le interpretazioni di varie importanti equazioni come flussi gradiente di particolari funzionali rispetto a opportune metriche.

Indice

1	EDO, flussi e l'equazione di Liouville	4
1.1	EDO e flussi	4
1.2	L'equazione di Liouville	7
1.3	Il pushforward	11
1.4	Flussi a divergenza nulla	12
1.5	Il caso dei sistemi meccanici conservativi	13
1.6	EDO autonome e gruppi a un parametro di diffeomorfismi	14
2	Il teorema del ritorno di Poincaré	15
3	L'equazione di Vlasov - derivazione euristica	17
3.1	Dall'equazione di Liouville alla gerarchia	17
3.2	Equazioni di campo medio	21
3.3	Altri modelli	24
3.3.1	Il modello per il flocking di Cucker e Smale	24
3.3.2	Il modello di Kuramoto per la sincronizzazione	25
3.3.3	Un algoritmo per la determinazione dei minimi	25
3.4	Una descrizione unificata	26
4	Una metrica per le misure di probabilità	28
4.1	La metrica 1-lipschitziana	28
4.2	Il problema di Monge-Kantorovich	31
4.3	La distanza di Wassertein	33
5	La validità dell'equazione di Vlasov	37
5.1	Costruzione delle soluzioni	37
5.2	Il limite di campo medio	41
6	L'equazione di Boltzmann	43
6.1	Modelli discreti	43
6.1.1	Il caso lineare e la master equation	43
6.1.2	Interazioni binare	44
6.2	L'equazione di Boltzmann per le sfere dure	45
6.2.1	La sezione d'urto delle sfere dure	48
7	La gerarchia BBGKY per il modello di sfere dure	51
7.1	L'equazione di Liouville per una dinamica con urto	51
7.2	L'equazione di Liouville per due sfere dure	54

1 EDO, flussi e l'equazione di Liouville

La descrizione fisico-matematica dei sistemi macroscopici, come gas, fluidi, solidi (in generale “mezzi continui”, in cui le scale spazio-temporali alle quali si vuol descrivere il mezzo non permettono l'analisi delle particelle che lo compongono), si fa attraverso opportune equazioni, che vengono derivate in vari modi. Uno è quello di utilizzare i principi primi dei sistemi di particelle (conservazione della massa, bilancio della quantità di moto e dell'energia). In questo modo però non si ottengono equazioni chiuse e vanno definite delle relazioni tra alcune delle grandezze in gioco attraverso le “relazioni costitutive”. Un altro approccio è quello di tentare di determinare il comportamento di un sistema macroscopico a partire dal comportamento dei suoi costituenti atomici (suggerito anche da Hilbert nel suo sesto problema). Seguendo questa idea, si possono “passare al limite” le equazioni per un sistema di N particelle, mandando N a infinito, ottenendo delle equazioni macroscopiche, in un qualche senso più “fondate” delle equazioni che si ottengono imponendo relazioni costitutive. Questo modo di procedere viene chiamato spesso “derivazione formale” delle equazioni.

La vera sfida consiste nel mostrare la **validità** delle equazioni macroscopiche, cioè di derivarle rigorosamente dai modelli microscopici: si tratta di mostrare che le soluzioni del sistema microscopico convergono in qualche senso alle soluzioni delle equazioni macroscopiche. Si noti che questo fatto non è assicurato dalla convergenza formale delle equazioni.

Per alcune equazioni cinetiche (e per alcuni sistemi idrodinamici) questo programma ha avuto parziale successo. In questa parte del corso discuseremo alcuni di questi casi. Inoltre le teorie cinetiche, che si studiano da 150 anni, hanno avuto un rinnovato interesse della comunità scientifica perché molta modellistica recente di agenti in interazione (animali, mercati, opinioni) può essere fatta in termini di equazioni cinetiche. Inoltre per questa parte del corso introduco e utilizzo il linguaggio del trasporto di massa (o meglio di misure di probabilità), che è al centro di molta ricerca recente in analisi e viene usato per spiegare alcuni aspetti del funzionamento dei LLM.

1.1 EDO e flussi

Le equazioni differenziali ordinarie (EDO) sono il principale strumento modellistico che la matematica ha messo a disposizione della fisica e di altre discipline.

Nei corsi di base dell'analisi avete appreso che una EDO viene risolta assegnando un dato iniziale. In molti problemi fisico-matematici, e specialmente nei sistemi dinamici e nella meccanica statistica, non ci si limita a studiare la singola soluzione di una EDO ma si osserva l'evoluzione di un intero insieme di stati possibili. La nozione di **flusso** associato a una EDO è la formalizzazione matematica di questa esigenza.

Definizione - flusso associato a una EDO: Indichiamo con $\Phi^{t,s}(\mathbf{x})$ il *flusso* associato al campo vettoriale $\mathbf{u}(t, \mathbf{x})$ per l'EDO

$$\dot{\mathbf{x}} = \mathbf{u}(t, \mathbf{x}).$$

Formalmente, $\Phi^{t,s}(\mathbf{x})$ è l'operatore di evoluzione che mappa il dato iniziale \mathbf{x} assegnato al tempo s , nella soluzione al tempo t . Esso risolve quindi il seguente problema di Cauchy:

$$\begin{aligned} \frac{d}{dt} \Phi^{t,s}(\mathbf{x}) &= \mathbf{u}(t, \Phi^{t,s}(\mathbf{x})) \\ \Phi^{s,s}(\mathbf{x}) &= \mathbf{x}. \end{aligned} \tag{1}$$

Per semplicità pupporrò che \mathbf{u} sia un campo definito da \mathbb{R}^d in \mathbb{R}^d , e sufficientemente regolare e limitato in modo che le soluzioni dell'EDO esistano per tutti i tempi e siano regolari nel tempo e nel dato iniziale.

Proprietà del Flusso.

- Fissati $s, t \in \mathbb{R}$, $\Phi^{t,s}$ è un diffeomorfismo da \mathbb{R}^d in sé.
- Per ogni $s \in \mathbb{R}$: $\Phi^{s,s} = \mathbb{I}$
- Vale la regola di composizione dei flussi per ogni $r, s, t \in \mathbb{R}$:

$$\Phi^{t,s} \circ \Phi^{s,r} = \Phi^{t,r} \quad (2)$$

- In particolare la mappa inversa coincide con l'evoluzione temporale invertita:

$$(\Phi^{t,s})^{-1} = \Phi^{s,t}.$$

Queste asserzioni si dimostrano dalle ipotesi su \mathbf{u} che garantiscono esistenza e unicità globale (lascio la dimostrazione per esercizio). Qui commento solo l'equazione (2): pensando a $\Phi^{t,r}(\mathbf{x})$ come all'operazione che va avanzare una particella in \mathbf{x} dal tempo r al tempo t , è evidente che questa operazione può essere ottenuta come composizione dell'avanzamento dal tempo r al tempo s di \mathbf{x} , con l'avanzamento del risultato dal tempo s al tempo t .

Queste affermazioni si estendono facilmente al caso di un campo vettoriale \mathbf{u} definito e regolare su un dominio Ω , con la condizione che $\mathbf{u}(t, \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0$, per ogni $\mathbf{x} \in \partial\Omega$, dove $\mathbf{n}(\mathbf{x})$ è la normale esterna a $\partial\Omega$ in \mathbf{x} .

Il flusso, dal punto di vista analitico, non contiene informazioni aggiuntive rispetto al sistema differenziale di partenza; tuttavia, la descrizione globale “a livello di flusso” apre la strada, come vedremo, a nuove indagini che rispondono a domande che riguardano il comportamento del sistema, e non la singola traiettoria.

Analizziamo ora alcune le proprietà geometriche e differenziali del flusso.

Richiami - Matrice jacobiana

Ricordo che se Φ è un diffeomorfismo, il suo jacobiano $\partial_{\mathbf{x}}\Phi(\mathbf{x})$ è un operatore lineare che ci dice come Φ trasforma le strutture lineari intorno a \mathbf{x} . Infatti, se considero \mathbf{x} e $\mathbf{x} + \mathbf{h}$, con \mathbf{h} “infinitesimo”, lo spostamento \mathbf{h} viene mappato da Φ in

$$\Phi(\mathbf{x} + \mathbf{h}) - \Phi(\mathbf{x}) \sim \partial_{\mathbf{x}}\Phi(\mathbf{x})\mathbf{h}.$$

Più formalmente,

$$d\Phi(\mathbf{x}) = \partial_{\mathbf{x}}\Phi(\mathbf{x}) d\mathbf{x}.$$

Il “determinate jacobiano” $J = \det \partial_{\mathbf{x}}\Phi(\mathbf{x})$ ci dice come si trasforma “l'elemento di volume”: se $\mathbf{y} = \Phi(\mathbf{x})$ allora $dy_1 \dots dy_d = J dx_1 \dots dx_d$.

Con un abuso di notazione, userò l'espressione $d\mathbf{x}$ non solo per indicare il vettore dei differenziali dx_i , ma anche per l'elemento di misura di Lebesgue $dx_1 \dots dx_d$.

Proposizione 1.1. *(Evoluzione dello jacobiano del flusso) Sia $\Phi^{t,s}$ il flusso associato al campo $\mathbf{u}(t, \mathbf{x})$ e sia $J(t, s, \mathbf{x}) = \det \partial_{\mathbf{x}}\Phi^{t,s}(\mathbf{x})$ il determinante della matrice jacobiana del flusso.*

Considerando noto il flusso, la matrice jacobiana verifica la seguente equazione differenziale lineare matriciale:

$$\frac{d}{dt} \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}) = \partial_{\mathbf{x}} \mathbf{u}|_{t, \Phi^{t,s}(\mathbf{x})} \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}) \quad (3)$$

Inoltre, il determinante jacobiano J soddisfa l'equazione differenziale scalare:

$$\frac{d}{dt} J(t, s, \mathbf{x}) = \operatorname{div} \mathbf{u}|_{t, \Phi^{t,s}(\mathbf{x})} J(t, s, \mathbf{x}) \quad (4)$$

Ricordando che $J(s, s, \mathbf{x}) = 1$ e considerando noto il flusso, questa equazione è risolta da

$$J(t, s, \mathbf{x}) = e^{\int_s^t d\tau \operatorname{div} \mathbf{u}(\tau, \Phi^{\tau,s}(\mathbf{x}))}.$$

In particolare J rimane strettamente positivo per tutti i tempi.

Dimostrazione:

L'asserzione sulla soluzione dell'equazione per J discende dalla teoria delle equazioni lineari omogenee a coefficienti non costanti. La prima equazione si ricava banalmente scambiando l'ordine di derivazione (rispetto a t e rispetto a \mathbf{x}) e applicando la regola della catena all'equazione che definisce il flusso. La seconda richiede di studiare la derivata del determinante. Il cuore logico è racchiuso nel seguente lemma.

Lemma 1.2.

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det \partial_{\mathbf{x}} \Phi^{t+\varepsilon,t}(\mathbf{x}) = \operatorname{div} \mathbf{u}(t, \mathbf{x})$$

La notazione un po' barocca serve per esplicitare la derivata di $\partial_{\mathbf{x}} \Phi^{t,t}$ rispetto alla prima variabile temporale t .

Accettando per un momento il Lemma 1.2, calcoliamo $\frac{d}{dt} J(t, s, \mathbf{x})$ sfruttando la proprietà di composizione del flusso $\Phi^{t+\varepsilon,s} = \Phi^{t+\varepsilon,t} \circ \Phi^{t,s}$:

$$\begin{aligned} \frac{d}{dt} \det \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}) &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det \partial_{\mathbf{x}} (\Phi^{t+\varepsilon,t} \circ \Phi^{t,s}(\mathbf{x})) \\ &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det \left(\partial_{\mathbf{x}} \Phi^{t+\varepsilon,t} \Big|_{\Phi^{t,s}(\mathbf{x})} \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}) \right) \\ &= \operatorname{div} \mathbf{u}|_{t, \Phi^{t,s}(\mathbf{x})} \det \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}). \end{aligned} \quad (5)$$

Per completare la dimostrazione, provo Lemma 1.2 usando un risultato di algebra lineare sulle perturbazioni dell'identità.

Lemma 1.3. *Sia A una matrice quadrata di ordine n , allora per $\varepsilon \rightarrow 0$ vale:*

$$\det(\mathbf{I} + \varepsilon A) = 1 + \varepsilon \operatorname{Tr} A + O(\varepsilon^2)$$

Dimostrazione. Poniamo $M = \mathbf{I} + \varepsilon A$. Dalla definizione classica di determinante:

$$\det M = \sum_{\sigma} (-1)^{\operatorname{sgn}(\sigma)} \prod_{i=1}^n M_{i\sigma_i}.$$

Isoliamo il termine corrispondente alla permutazione identica ($\sigma_i = i$):

$$\prod_{i=1}^n M_{ii} = \prod_{i=1}^n (1 + \varepsilon A_{ii}) = 1 + \varepsilon \sum_{i=1}^n A_{ii} + O(\varepsilon^2) = 1 + \varepsilon \operatorname{Tr} A + O(\varepsilon^2).$$

Qualsiasi altra permutazione coinvolge almeno due elementi fuori dalla diagonale principale, i quali sono di ordine $O(\varepsilon)$. Moltiplicandoli, produrranno contributi di ordine almeno $O(\varepsilon^2)$. Il lemma è provato.

Dimostrazione del Lemma 1.2:

Sviluppando in serie di Taylor il flusso per tempi piccoli otteniamo:

$$\Phi^{t+\varepsilon,t}(\mathbf{x}) = \mathbf{x} + \varepsilon \mathbf{u}(t, \mathbf{x}) + O(\varepsilon^2).$$

Derivando rispetto a \mathbf{x} e calcolando il determinante tramite il Lemma 1.3 (sostituendo $A = \partial_{\mathbf{x}}\mathbf{u}$):

$$\det \partial_{\mathbf{x}}\Phi^{t+\varepsilon,t}(\mathbf{x}) = \det (\mathbf{I} + \varepsilon \partial_{\mathbf{x}}\mathbf{u}(t, \mathbf{x}) + O(\varepsilon^2)) = 1 + \varepsilon \operatorname{div} \mathbf{u}(t, \mathbf{x}) + O(\varepsilon^2).$$

Qui abbiamo sfruttato che la traccia della matrice jacobiana di un campo vettoriale coincide con la sua divergenza $\operatorname{Tr}(\partial_{\mathbf{x}}\mathbf{u}) = \operatorname{div} \mathbf{u}$. Dividendo per ε e passando al limite per $\varepsilon \rightarrow 0$, si ottiene la tesi.

Sunto

- **Definizione di flusso:** è la famiglia a due parametri di diffeomorfismi $\Phi^{t,s}$, dove s è il tempo iniziale e t quello finale.
- **Proprietà del flusso:** il flusso da s a t è la composizione del flusso dal tempo s al tempo r e dal tempo r al tempo t .
- **Lo jacobiano del flusso** verifica l'EDO lineare omogenea definita attraverso la matrice jacobiana del campo $\partial_{\mathbf{x}}\mathbf{u}(t, \Phi^{t,s}(\mathbf{x}))$; il suo determinante J verifica l'EDO lineare omogenea definita dalla divergenza del campo calcolata lungo il flusso.

1.2 L'equazione di Liouville

La descrizione di una EDO mediante il flusso ci permette di studiare il sistema in presenza di *incertezza sul dato iniziale*. Nella realtà fisica, infatti, è impossibile misurare o conoscere lo stato iniziale di un sistema con precisione infinita.

Descrizione statistica e incertezza.

Immaginiamo di assegnare una densità di probabilità ai dati iniziali al tempo $t = 0$, che indicheremo con $f_0(\mathbf{x})$ (consideriamo $s = 0$ come tempo iniziale senza perdita di generalità). La dinamica è deterministica, dunque l'incertezza "viene trasportata" sul punto di partenza si propaga dinamicamente, traducendosi in una distribuzione di probabilità per lo stato del sistema a qualsiasi istante successivo t . L'equazione di Liouville è l'equazione alle derivate parziali (PDE) che governa proprio l'evoluzione temporale di questa densità.

Cominciamo con un esempio, supponendo che il dato iniziale sia distribuito uniformemente in un dominio misurabile B , ovvero:

$$f_0(\mathbf{x}) = \frac{1}{|B|} \mathcal{X}_B(\mathbf{x})$$

dove $|B|$ è la misura di Lebesgue di B e \mathcal{X}_B è la funzione indicatrice. L'evoluzione microscopica è strettamente **deterministica**, dunque sappiamo con certezza (con probabilità 1) che al tempo t il sistema si troverà nell'**evoluzione dell'insieme B** :

$$B_t := \Phi^{t,0}(B) = \{\Phi^{t,0}(\mathbf{x}) \mid \mathbf{x} \in B\}.$$

Tuttavia, non abbiamo alcuna garanzia che la distribuzione all'interno di B_t rimanga uniforme: il campo vettoriale può comprimere o dilatare i volumi dello spazio delle fasi in modo non omogeneo. Approfondiamo.

Richiami - Misure di probabilità

Ricordo, con un linguaggio non del tutto preciso, che a $f \in L^1(\mathbb{R}^d)$ funzione positiva con integrale 1 si può associare una **distribuzione di probabilità** dichiarando, per ogni misurabile A ,

$$\mathbb{P}(A) = \int_A f(\mathbf{x}) \, d\mathbf{x}$$

In questo modo viene definita una **misura di probabilità**, cioè una misura positiva su \mathbb{R}^d con $\mathbb{P}(\mathbb{R}^d) = 1$, di cui f è la **densità di probabilità**.

Una misura di probabilità ammette densità se e solo se è assolutamente continua rispetto alla misura di Lebesgue, cioè se $|A| = 0$ implica $\mathbb{P}(A) = 0$ (è una conseguenza del teorema di Radon-Nikodym: f è la “derivata di Radon-Nikodym” della misura di probabilità rispetto a quella di Lebesgue).

A ogni misura di probabilità $\mu(d\mathbf{x})$ si può associare una variabile aleatoria \mathbf{X} , dichiarando

$$\mathbb{P}(\mathbf{X} \in A) = \int_A \mu(d\mathbf{x})$$

La misura μ si chiama **legge della variabile aleatoria**. Un **osservabile** della variabile aleatoria \mathbf{X} è una qualunque funzione misurabile $F : \mathbb{R}^d \rightarrow \mathbb{R}$ rispetto alla misura di probabilità. Il suo valore atteso è

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

Se μ ammette densità f rispetto alla misura di Lebesgue, questa espressione diventa $\int_A F(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x}$.

Considerando una misura di probabilità iniziale di densità $f_0(\mathbf{x})$. Vogliamo scoprire come deve essere fatta la densità di probabilità $f(t, \mathbf{x})$ a un tempo $t > 0$. Possiamo pensare che il dato iniziale \mathbf{X}_0 sia una variabile aleatoria distribuita con f_0 . Il sistema al tempo t sarà dato dalla variabile aleatoria $\mathbf{X}_t = \Phi^{t,0}(\mathbf{X}_0)$: scelgo \mathbf{X}_0 in base alla distribuzione di probabilità iniziale, e poi ne considero l'evoluzione data dal sistema determinisco, il risultato è ancora una variabile aleatoria perché il dato iniziale lo è, anche se la dinamica è deterministica. In pratica la probabilità interviene solo nel dato iniziale e viene solo trasportata dal flusso. Vediamo come.

Principio di conservazione della probabilità:

Per definizione di \mathbf{X}_t , per ogni A misurabile

$$\mathbb{P}(\mathbf{X}_0 \in A) = \mathbb{P}(\mathbf{X}_t \in A_t)$$

Questa espressione ci dice che la probabilità si **conserva**. Indicando con f_t lo densità (incognita) al tempo t , la precedente identità si scrive

$$\int_A d\mathbf{x} f_0(\mathbf{x}) = \int_{A_t} d\mathbf{x} f(t, \mathbf{x}) \tag{6}$$

dove $A_t := \Phi^{t,0}(A)$. Questa condizione globale ci permette di identificare puntualmente f al tempo t nel seguente modo. Cambiamo variabile $\mathbf{x} = \Phi^{t,0}(\mathbf{y})$ nell'integrale a sinistra e rinominiamo \mathbf{y} in \mathbf{x} per comodità di lettura

$$\int_A d\mathbf{x} f_0(\mathbf{x}) = \int_{A_t} d\mathbf{x} f(t, \mathbf{x}) = \int_A d\mathbf{x} f(t, \Phi^{t,0}(\mathbf{x})) J(t, 0, \mathbf{x})$$

Portando tutto sotto lo stesso segno di integrale, ricaviamo che per ogni insieme A :

$$\int_A [f(t, \Phi^{t,0}(\mathbf{x}))J(t, 0, \mathbf{x}) - f_0(\mathbf{x})] d\mathbf{x} = 0$$

Per la teoria dell'integrazione di Lebesgue, l'integrando deve annullarsi quasi ovunque:

$$f(t, \Phi^{t,0}(\mathbf{x}))J(t, 0, \mathbf{x}) = f_0(\mathbf{x}). \quad (7)$$

Per rigore, stiamo usando il seguente lemma:

Lemma 1.4 (Lemma fondamentale del calcolo delle variazioni.). *Sia $g(\mathbf{x})$ una funzione localmente integrabile. Se per ogni misurabile A vale*

$$\int_A g(\mathbf{x}) d\mathbf{x} = 0$$

allora g è nulla quasi ovunque.

Dimostrazione. Sia A^+ l'insieme su cui $g > 0$; per definizione è misurabile, e, usando l'ipotesi, si ottiene che ha misura nulla (ha misura nulla $A^+ \cap B$ per ogni palla B , e dunque A^+ ha misura nulla). Analogamente si procede per $g < 0$, e si ottiene che la misura degli \mathbf{x} per cui g è non nulla è zero, cioè g è nulla quasi ovunque.

L'uguaglianza (7) ci definisce esplicitamente il valore di f al tempo t e nella posizione \mathbf{x} :

$$f(t, \mathbf{x}) = \frac{f_0(\Phi^{0,t}(\mathbf{x}))}{J(t, 0, \Phi^{0,t}(\mathbf{x}))}. \quad (8)$$

Interpretazione geometrica: L'equazione (7) è la versione microscopica (locale) della conservazione della probabilità imposta nella (6). Immaginiamo la quantità $f_0(\mathbf{x}) d\mathbf{x}$ come la probabilità contenuta in un cilindro di base infinitesima $d\mathbf{x}$ e altezza $f_0(\mathbf{x})$. Sotto l'azione del flusso, la base si deforma linearmente tramite $\partial_{\mathbf{x}}\Phi^{t,0}$ e la sua misura diventa $J(t, 0, \mathbf{x}) d\mathbf{x}$. Affinché il volume (la probabilità totale) rimanga costante, la "nuova" altezza f deve compensare esattamente il cambiamento di volume espresso da J :

$$f_0(\mathbf{x}) d\mathbf{x} = f(t, \Phi^{t,0}(\mathbf{x})) d\Phi^{t,0}(\mathbf{x}) = f(t, \Phi^{t,0}(\mathbf{x}))J(t, 0, \mathbf{x}) d\mathbf{x}$$

Derivazione dell'equazione di Liouville

Cerchiamo ora l'equazione alle derivate parziali (PDE) soddisfatta da f . Deriviamo rispetto al tempo t la relazione (7). Il membro di destra è costante nel tempo (derivata nulla). Applicando la regola di Leibniz a sinistra:

$$\partial_t \left(f(t, \Phi^{t,0}(\mathbf{x})) \right) J(t, 0, \mathbf{x}) + f(t, \Phi^{t,0}(\mathbf{x})) \partial_t J(t, 0, \mathbf{x}) = 0$$

Il primo termine rappresenta la **derivata materiale** (o lungo il flusso) di f :

$$\partial_t \left(f(t, \Phi^{t,0}(\mathbf{x})) \right) = (\partial_t f + \mathbf{u} \cdot \nabla f)|_{t, \Phi^{t,0}(\mathbf{x})}$$

Nel secondo termine, inseriamo l'espressione di $\partial_t J$ ricavata in precedenza. Sostituendo otteniamo:

$$(\partial_t f + \mathbf{u} \cdot \nabla f + f(\operatorname{div} \mathbf{u})|_{t, \Phi^{t,0}(\mathbf{x})}) J(t, 0, \mathbf{x}) = 0$$

Poiché J non è mai nullo, componendo con il flusso inverso si ottiene

$$\partial_t f(t, \mathbf{x}) + \mathbf{u}(t, \mathbf{x}) \cdot \nabla f(t, \mathbf{x}) + f(t, \mathbf{x}) \operatorname{div} \mathbf{u}(t, \mathbf{x}) = 0.$$

Il primo addendo è la variazione esplicita, il secondo è il termine di trasporto, il terzo tiene conto delle variazioni di misura. Usando la regola di Leibniz vettoriale $\operatorname{div}(f\mathbf{u}) = \mathbf{u} \cdot \nabla f + f \operatorname{div} \mathbf{u}$, si ricava infine l'**equazione di Liouville**:

$$\partial_t f + \operatorname{div}(f\mathbf{u}) = 0. \quad (9)$$

Possiamo riassumere quanto abbiamo fatto dicendo che questa PDE con dato iniziale f_0 si risolve determinando il flusso e utilizzando l'espressione (8). Questo è un esempio particolare del caso in cui una PDE si risolve risolvendo una EDO, le cui soluzioni prendono il nome di **caratteristiche**, e l'equazione si dice risolta con il **metodo delle caratteristiche**.

Interpretazione fisica.

L'equazione di Liouville è una **legge di conservazione** in forma di divergenza (*vedi* anche Buttà pagg. 10–11), cioè è una equazione del tipo

$$\partial_t g(t, \mathbf{x}) + \operatorname{div}(\mathbf{j}(t, \mathbf{x})) = 0$$

dove il campo vettoriale \mathbf{j} prende il nome di **corrente** dell'incognita g . Nel caso dell'equazione di Liouville, la corrente per f è $f\mathbf{u}$.

Il motivo di questa definizione è il seguente: se integriamo ambo i membri su un dominio A e applichiamo il Teorema della divergenza (o Teorema di Gauss-Green), otteniamo:

$$\frac{d}{dt} \int_A g(t, \mathbf{x}) \, d\mathbf{x} = - \int_A \operatorname{div}(\mathbf{j}(t, \mathbf{x})) \, d\mathbf{x} = - \int_{\partial A} \mathbf{j}(t, \mathbf{x}) \cdot \mathbf{n} \, \sigma(d\mathbf{x})$$

dove con $\sigma(d\mathbf{x})$ intendo la misura sul bordo ∂A . Quest'ultima uguaglianza asserisce che la variazione temporale di g all'interno del volume fisso A è pari esclusivamente al flusso della corrente attraverso la superficie di bordo ∂A .

Per chi le conosce, anche le equazioni di diffusione, come quella del calore $\partial_t f = \Delta f$, sono leggi di conservazione; in tal caso la corrente è $-\nabla f$.

Esercizi proposti:

- Si provi, usando l'espressione integrale di J , che:

$$f(t, \Phi^{t,0}(\mathbf{x}), t) = f_0(\mathbf{x}) \exp\left(- \int_0^t \operatorname{div} \mathbf{u}(\tau, \Phi^{\tau,0}(\mathbf{x})) \, d\tau\right)$$

- Si mostri che l'equazione di Liouville è equivalente a

$$\frac{d}{dt} f(t, \Phi^{t,0}(\mathbf{x})) = - \operatorname{div} \mathbf{u}(t, \Phi^{t,0}(\mathbf{x})) f(t, \Phi^{t,0}(\mathbf{x}))$$

Si metta in relazione questa uguaglianza con l'espressione di f al punto precedente.

- Si provi che $J(t, s, \Phi^{s,t}(\mathbf{x})) J(s, t, \mathbf{x}) = 1$.

1.3 Il pushforward

Rileggiamo la soluzione trovata in termini di **pushforward** di misure di probabilità, concetto che ci sarà utile nel seguito.

Supponiamo di considerare una variabile aleatoria \mathbf{X} che prende valori su V (per i nostri scopi un sottoinsieme di \mathbb{R}^d , ma può essere più generale). La sua legge è una misura di probabilità μ su V che permette di esprimere

$$\mathbb{P}(\mathbf{X} \in A) = \int_A \mu(d\mathbf{x}).$$

Richiami - Misure e funzioni continue

Ricordo che una generica misura (anche con segno) $\nu(d\mathbf{x})$ si può identificare con un elemento del duale delle funzioni continue e limitate su V , spazio che indico con $C_b(V)$. In pratica una misura ν è un funzionale lineare continuo (o, equivalentemente, limitato) su $C_b(V)$, cioè esiste c tale che per ogni $\varphi \in C_b$

$$\left| \int \varphi(\mathbf{x})\nu(d\mathbf{x}) \right| \leq c\|\varphi\|_\infty$$

dove l'integrale a sinistra è, per definizione, l'azione di ν su φ .

Dunque per assegnare una misura è sufficiente sapere come agisce sulle funzioni continue e limitate. In termini probabilistici, la legge μ di una variabile aleatoria \mathbf{X} è perfettamente definita dai valori attesi di osservabili in C_b :

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

Sia ora Φ una funzione continua da V a W (per fissare le idee, un sottoinsieme di \mathbb{R}^m). In tal modo si definisce $\mathbf{Y} := \Phi(\mathbf{X})$ variabile aleatoria su W . Cerchiamo la legge di \mathbf{Y} che indicheremo con ν . Sia $\beta \in C_b(W)$. Per definizione

$$\int \beta(\mathbf{y})\nu(d\mathbf{y}) = \mathbb{E}(\beta(\mathbf{Y})) = \mathbb{E}(\beta(\Phi(\mathbf{X}))) = \int \beta(\Phi(\mathbf{x}))\mu(d\mathbf{x}).$$

Poiché le misure sono individuate dai valori degli integrali degli osservabili, la relazione precedente definisce esattamente la legge ν . Formalizzo.

Si chiama **pushforward** di μ mediante $\Phi : V \rightarrow W$ la misura su W , indicata con $\Phi\#\mu$, definita dalla seguente relazione: qualunque sia $\beta \in C_b(\mathbb{R}^m)$

$$\int_W \beta(\mathbf{y})\Phi\#\mu(d\mathbf{y}) := \int_V \beta(\Phi(\mathbf{x}))\mu(d\mathbf{x}).$$

Osservo che questa definizione si estende a Φ misurabile tali che il secondo membro ha senso, e questo accade se Φ è misurabile.

Teorema 1.5. *In termini della misura μ_t la misura di densità f_t , la soluzione dell'equazione di Liouville è il pushforward di μ_0 mediante il flusso:*

$$\mu_t = \Phi^{t,0}\#\mu_0.$$

Si noti che dato il flusso, posso fare pushforward di ogni misura iniziale, non solo quelle assolutamente continue rispetto a Lebesgue, in tal caso, però, otterremo **soluzioni deboli**

perché non possiamo dare senso puntuale a $\operatorname{div}(\mathbf{u}f)$. Per esercizio, si provi che se μ è una misura, allora

$$\mu_t := \Phi^{t,0} \# \mu$$

risolve nel senso debole delle distribuzioni l'equazione di Liouville, cioè per ogni $\varphi(t, \mathbf{x})$ regolare

$$\int \varphi(t, \mathbf{x}) \mu_t(d\mathbf{x}) - \int \varphi(0, \mathbf{x}) \mu_0(d\mathbf{x}) = \int_0^t ds \int \partial_t \varphi(s, \mathbf{x}) \mu_s(d\mathbf{x}) + \int_0^t ds \int \nabla \varphi(s, \mathbf{x}) \cdot \mathbf{u}(s, \mathbf{x}) \mu_s(d\mathbf{x})$$

Questa definizione di soluzione debole non è particolarmente utile. Nel caso dell'equazione di Liouville è preferibile considerare l'identità $\Phi_{\#}^{t,0} \mu_0 = \mu_t$ come definizione di soluzione debole, infatti equivale ad affermare che

$$\int \varphi(\mathbf{x}) \mu_t(d\mathbf{x}) = \int \varphi(\Phi^{t,0}(\mathbf{x})) \mu_0(d\mathbf{x}).$$

e individua esattamente la soluzione.

Sunto

- **Conservazione della probabilità:** l'incertezza sui dati iniziali viene trasportata al tempo t dal flusso deterministico. Imponendo questa regola si ottiene l'espressione per la densità di probabilità al tempo t , in funzione di quella al tempo 0.
- **L'equazione di Liouville:** è la PDE che governa l'evoluzione della densità di probabilità: nella sua espressione c'è il termine di trasporto (la derivata lungo il flusso) e il termine $f \operatorname{div} \mathbf{u}$ che tiene conto delle variazioni di volume dovute al flusso.
- L'equazione di Liouville è risolta dal push-forward della misura iniziale lungo il flusso, anche se il dato iniziale non è regolare.

1.4 Flussi a divergenza nulla

Proposizione 1.6.

$$\frac{d}{dt} |A_t| = \int_{A_t} \operatorname{div} \mathbf{u}(t, \mathbf{x}) d\mathbf{x}$$

In particolare, $\operatorname{div} \mathbf{u} \equiv 0$ se e solo se il flusso conserva la misura.

La dimostrazione si ottiene facilmente calcolando la derivata dopo aver cambiato variabile nell'integrale

$$|A_t| = \int_{A_t} d\mathbf{x} = \int_A J(t, 0, \mathbf{x}) d\mathbf{x}$$

e usando il lemma 1.4

Definizione - mappa che conserva la misura: una mappa Φ da un dominio Ω in sé conserva la misura se e solo se per ogni funzione $\alpha \in C_b(\Omega)$

$$\int \alpha(\Phi(\mathbf{x})) d\mathbf{x} = \int \alpha(\mathbf{x}) d\mathbf{x}$$

Con gli strumenti della teoria della misura di Lebesgue, si prova che questa definizione è equivalente a quella che considera α come funzione caratteristica di misurabili $A \subset \Omega$, e dunque è equivalente a

$$|\Phi^{-1}(A)| = |A|,$$

per ogni A , dove $\Phi^{-1}(A) = \{\mathbf{y} | \Phi(\mathbf{y}) \in A\}$.

Se Φ è biettiva, questa definizione coincide con la proprietà $|\Phi(A)| = |A|$. Se Φ è anche regolare, questa proprietà è ovviamente equivalente a $J \equiv 1$.

La precedente proposizione ha come corollario immediato il seguente fatto.

Proposizione 1.7. *Il flusso $\Phi^{t,s}$ conserva la misura per ogni t se e solo se $\operatorname{div} \mathbf{u} \equiv 0$.*

Si noti che se il flusso di campo \mathbf{u} conserva la misura, data μ_0 assolutamente continua rispetto alla misura di Lebesgue, con densità f_0 , si ha

$$\Phi^{t,0} \# \mu_0(d\mathbf{x}) = f(t, \mathbf{x}) d\mathbf{x} = f_0(\Phi^{0,t}(\mathbf{x})) d\mathbf{x}$$

e quindi $f_t(\Phi^{t,0}(\mathbf{x})) = f_0(\mathbf{x})$, cioè f è **trasportata** dal flusso generato da \mathbf{u} . Derivando in t , si ottiene che f soddisfa

$$\partial_t f + \mathbf{u} \cdot \nabla f = 0,$$

che prende il nome di **equazione del trasporto**, infatti a sinistra c'è la derivata lungo il flusso. Equivalentemente, l'equazione del trasporto (che si può scrivere qualunque sia la divergenza di \mathbf{u}), coincide con l'equazione di Liouville se e solo se $\operatorname{div} \mathbf{u} = 0$, cioè se il flusso conserva la misura. Anche l'equazione del trasporto si risolve con il metodo delle caratteristiche, infatti $f(t, \mathbf{x}) = f_0(\Phi^{0,t}(\mathbf{x}))$; i valori di f vengono solo **trasportati** dal flusso, mentre per l'equazione di Liouville vengono anche corretti dal termine dovuto alla variazione di volume.

1.5 Il caso dei sistemi meccanici conservativi

I sistemi meccanici conservativi sono governati da EDO del secondo ordine, in cui l'accelerazione è pari alla forza \mathbf{F} , che non dipende dalla velocità, ed è l'opposto del gradiente dell'energia potenziale. In questo caso, la densità di probabilità nello spazio delle fasi (velocità e posizione) $f(\mathbf{x}, \mathbf{v}, t)$ soddisfa l'equazione di Liouville

$$\partial_t f(\mathbf{x}, \mathbf{v}, t) + \operatorname{div}_{\mathbf{x}}(\mathbf{v}f(\mathbf{x}, \mathbf{v}, t)) + \operatorname{div}_{\mathbf{v}}(\mathbf{F}(\mathbf{x}, t)f(\mathbf{x}, \mathbf{v}, t)) = 0$$

che è equivalente all'equazione del trasporto

$$\partial_t f(\mathbf{x}, \mathbf{v}, t) + \mathbf{v} \cdot \partial_{\mathbf{x}} f(\mathbf{x}, \mathbf{v}, t) + \mathbf{F}(\mathbf{x}, t) \cdot \partial_{\mathbf{v}} f(\mathbf{x}, \mathbf{v}, t) = 0$$

Infatti, il campo $\begin{pmatrix} \mathbf{v} \\ \mathbf{F} \end{pmatrix}$ ha divergenza nulla nella coppia di variabili (\mathbf{x}, \mathbf{v}) . In altre parole, i sistemi meccanici conservano la misura nello spazio delle fasi, e l'equazione di Liouville coincide con la corrispondente equazione del trasporto (si noti che queste affermazioni non sono invarianti per la scelta di coordinate lagrangiane diverse da quelle rettangolari, ma sono sempre vere se il sistema viene descritto nel formalismo hamiltoniano, come vedremo in seguito).

Infine, il più semplice sistema meccanico è quello costituito da una particella libera. In questo caso la forza è nulla e l'equazione di Liouville si riduce a

$$\partial_t f(\mathbf{x}, \mathbf{v}, t) + \mathbf{v} \cdot \partial_{\mathbf{x}} f(\mathbf{x}, \mathbf{v}, t) = 0$$

che è risolta da

$$f(\mathbf{x}, \mathbf{v}, t) = f_0(\mathbf{x} - t\mathbf{v}, \mathbf{v})$$

che descrive il trasporto del **flusso libero** di una distribuzione di probabilità di dati iniziali.

Sunto

- I flussi associati a campi a divergenza nulla conservano la misura.
- I sistemi meccanici conservativi sono governati da campi a divergenza nulla, dunque il flusso associato conserva la misura nello spazio delle fasi.
- Se il campo ha divergenza nulla l'equazione di Liouville coincide con l'equazione del trasporto lineare, ed è risolta dal trasporto della densità.

1.6 EDO autonome e gruppi a un parametro di diffeomorfismi

Se l'EDO è **autonoma**, cioè il campo \mathbf{u} non dipende dal tempo, la nozione di flusso si semplifica. Non è difficile immaginare che deve valere

$$\Phi^{t,s}(\mathbf{x}) = \Phi^{t-s,0}(\mathbf{x})$$

infatti, se il campo vettoriale è costante nel tempo, l'unica cosa importante è quanto tempo è passato, non quali sono gli istanti iniziali e finali. Si provi l'identità per esercizio, derivando in t entrambi i membri e invocando l'unicità delle soluzioni dell'EDO.

Nel caso di EDO autonome, indicherò con Φ^t il flusso $\Phi^{t,0}$.

Proposizione 1.8. *Il flusso Φ^t associato a una EDO autonoma è un gruppo a un parametro di diffeomorfismi, cioè $\Phi^0 = \mathbb{I}$ e $\Phi^t \circ \Phi^s = \Phi^{s+t}$*

La tesi è conseguenza delle proprietà del flusso, specializzate al caso autonomo.

Teorema 1.9 (Generatori dei flussi). *Sia Φ^t un gruppo a un parametro di diffeomorfismi, cioè*

$$\Phi^0(\mathbf{x}) = \mathbf{x}, \quad e \quad \forall s, t \quad \Phi^t \circ \Phi^s = \Phi^{s+t}$$

Allora esiste un campo vettoriale \mathbf{u} tale che Φ^t è il flusso generato da \mathbf{u} , cioè

$$\frac{d}{dt}\Phi^t(\mathbf{x}) = \mathbf{u}(\Phi^t(\mathbf{x}))$$

Dimostrazione. Se esiste, $\mathbf{u}(\mathbf{x})$ deve coincidere con la derivata in t al tempo $t = 0$ di $\Phi^t(\mathbf{x})$. Definisco dunque

$$\mathbf{u}(\mathbf{x}) = \left. \frac{d}{dt}\Phi^t(\mathbf{x}) \right|_{t=0}.$$

Verifico che genera Φ^t .

$$\frac{d}{dt}\Phi^t(\mathbf{x}) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\Phi^{t+\varepsilon}(\mathbf{x}) - \Phi^t(\mathbf{x})) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\Phi^\varepsilon(\Phi^t(\mathbf{x})) - \Phi^t(\mathbf{x})) = \mathbf{u}(\Phi^t(\mathbf{x}))$$

Nella penultima uguaglianza ho usato la proprietà di gruppo, nell'ultima ho usato la definizione di \mathbf{u} . □

2 Il teorema del ritorno di Poincaré

Tra le principali conseguenze della conservazione della misura per il flusso di fase c'è il teorema del ritorno di Poincaré, che afferma che per un flusso a un parametro che conserva la misura, definito su una regione limitata, "quasi ogni" traiettoria ritorna arbitrariamente vicina al suo dato iniziale.

Teorema 2.1 (Teorema del ritorno I). *Sia Ω un insieme di misura limitata, e sia Φ una biezione su Ω , misurabile con inversa misurabile.*

Sia A un sottoinsieme misurabile di Ω , allora, per ogni $k \in \mathbb{N}$ esiste $n \geq k$ tale che

$$|A \cap \Phi^n(A)| > 0$$

e dunque l'intersezione è non vuota. In particolare, esiste $\mathbf{x} \in A$ tale che $\Phi^n(\mathbf{x}) \in A$.

Dimostrazione. Definisco $\Psi = \Phi^k$. Anche Ψ è biettiva e conserva la misura (verificare). Considero la sequenza di sottoinsiemi $\{\Psi^h(A)\}_{h \geq 0}$. Se le loro intersezioni avessero misura nulla si avrebbe

$$|\Omega| \geq \left| \bigcup_h \Psi^h(A) \right| = \sum_h |\Psi^h(A)| = \sum_h |A| = +\infty$$

che è assurdo poiché Ω è limitato. Dunque esiste $h \geq 0$ ed esiste $m \geq 1$ tali che

$$\Psi^h(A) \cap \Psi^{h+m}(A) \neq \emptyset$$

e ha misura non nulla. Poiché Ψ è biettiva, ne segue che

$$A \cap \Psi^m(A) \neq \emptyset$$

In termini di Φ , questa relazione si legge

$$A \cap \Phi^{mk} \neq \emptyset$$

Scegliendo $n = mk \geq k$ si ottiene la tesi. □

Un corollario di questo teorema è il seguente. Sia $\varepsilon > 0$; dirò che \mathbf{x} è un punto che ε -**ritorna** se per ogni $k \in \mathbb{N}$, esiste $n > k$ tale che $|\Phi^n(\mathbf{x}) - \mathbf{x}| < \varepsilon$.

Teorema 2.2 (Teorema del ritorno II). *L'insieme dei punti che non ε -ritornano ha misura nulla.*

Dimostrazione. Sia N_ε l'insieme dei punti che non ε -ritornano, cioè

$$N_\varepsilon = \{\mathbf{x} \mid |\Phi^k(\mathbf{x}) - \mathbf{x}| \geq \varepsilon \text{ definitivamente in } k\} = \bigcup_{k \geq 1} N_\varepsilon^k$$

dove

$$N_\varepsilon^k = \{\mathbf{x} \mid \forall n \geq k \mid \Phi^n(\mathbf{x}) - \mathbf{x}| \geq \varepsilon\} = \bigcap_{n \geq k} \tilde{N}_\varepsilon^n$$

con

$$\tilde{N}_\varepsilon^n = \{\mathbf{x} \mid |\Phi^n(\mathbf{x}) - \mathbf{x}| \geq \varepsilon\}$$

\tilde{N}_ε^n è misurabile perché la funzione $|\Phi^n(\mathbf{x}) - \mathbf{x}|$ è misurabile, e dunque sono misurabili anche N_ε^k e N_ε .

Dimostro che N_ε^k ha misura nulla. Posso ricoprire Ω con una quantità al più numerabile di palle B_1, B_2, \dots di raggio $\varepsilon/2$. Sia $A = B_i \cap N_\varepsilon^k$, che è misurabile. Se per assurdo avesse misura positiva, per il teorema di Poincaré esisterebbe $n > k$ tale che $\Phi^n(A) \cap A$ avrebbe misura positiva. Ma allora esisterebbe $\mathbf{y} \in A \cap \Phi^n(A)$ che torna nello stesso insieme, in particolare $\mathbf{x} = \Psi^n(\mathbf{y}) \in A$. Ne segue

$$|\mathbf{y} - \Psi^n(\mathbf{y})| = |\mathbf{y} - \mathbf{x}| < \varepsilon$$

perché $\mathbf{x}, \mathbf{y} \in A \subset B_i$. Questo fatto contraddice la definizione di A , che deve dunque avere misura nulla. Ma allora $B_i \cap N_\varepsilon^k$ ha misura nulla per ogni i . Usando la σ -subadditività della misura di Lebesgue

$$N_\varepsilon^k = \bigcup_i B_i \cap N_\varepsilon^k \Rightarrow |N_\varepsilon^k| \leq \sum_i |B_i \cap N_\varepsilon^k| = 0$$

Di nuovo usando la σ -subadditività

$$N_\varepsilon = \bigcup_k N_\varepsilon^k \Rightarrow |N_\varepsilon| \leq \sum_k |N_\varepsilon^k| = 0$$

□

Infine, si può provare la seguente affermazione più forte. Dirò che $x \in \Omega$ **ritorna** se

$$\forall \varepsilon > 0, \forall k \in \mathbb{N}, \exists n \geq k \text{ tale che } |\Phi^n(\mathbf{x}) - \mathbf{x}| < \varepsilon$$

Detto in parole, \mathbf{x} ritorna se ε -ritorna per ogni ε , ovvero se la traiettoria che parte da \mathbf{x} torna infinite volte arbitrariamente vicino a \mathbf{x} .

Teorema 2.3 (Teorema del ritorno III). *L'insieme dei punti che non ritornano ha misura nulla.*

Dimostrazione. Se il punto \mathbf{x} ritorna, allora $1/n$ -ritorna per ogni $n \in \mathbb{N}$ positivo. Dunque l'insieme dei punti che ritornano è l'intersezione per $n \geq 1$ degli insiemi dei punti che $1/n$ -ritornano. Il complementare è l'insieme dei punti che non ritornano, ed è dunque l'unione per $n \geq 1$ dell'insieme dei punti che non $1/n$ -ritornano. Poiché questi insiemi hanno misura nulla, la loro unione numerabile ha misura nulla. □

Per commenti sui paradossi del teorema del ritorno vedi le dispense di Buttà–Negrini. È da notare, comunque, che lo studio delle proprietà a tempi lunghi dei sistemi conservativi (hamiltoniani) sono l'inizio dello studio che porta alla descrizione dei sistemi mediante la Meccanica statistica (parola chiave **ipotesi ergodica**), ma non dirò nulla in merito.

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'analogia 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 $h \geq j$. 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 sciame 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

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

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.

4 Una metrica per le misure di probabilità

L'equazione di Vlasov è di fatto un'equazione di Liouville non lineare, e per fare la teoria di esistenza e unicità servono delle metriche che tengono in conto del fatto che l'equazione di Liouville è l'equazione del trasporto delle misure.

Le usuali norme di funzioni non sono molto adatte; consideriamo per esempio A e B domini disgiunti di \mathbb{R}^d , di misura 1, e consideriamo la distanza in L_1 tra \mathcal{X}_A e \mathcal{X}_B . Questa distanza è esattamente 2, indipendentemente da quanto siano lontani in \mathbb{R}^d i due insiemi. Più in generale, $\|f - g\|_1$ misura la differenza dei valori tra f e g in tutti i possibili punti di \mathbb{R}^d , e poi li somma. Le distanze L_p operano nello stesso modo.

Abbiamo bisogno di una nozione di distanza che tenga conto anche della distanza "geometrica": definirò la distanza 1-lipschitziana tra misure di probabilità, accennando anche al fatto che coincide con la distanza 1-Wassertein, anche detta distanza di Kantorovich-Rubinstein.

4.1 La metrica 1-lipschitziana

Comincio con l'introdurre l'analogo della norma L_1 per le misure finite con segno¹. Data una misura con segno μ , essa si decompone in $\mu = \mu^+ - \mu^-$ dove μ^\pm sono due misure positive a supporto disgiunto (per il Teorema di decomposizione di Hahn-Jordan). La norma di μ in variazione totale è

$$\|\mu\|_{VT} = \int \mu^+ + \int \mu^-.$$

Si può provare che questa è esattamente la norma nello spazio delle misure con segno considerato come duale delle funzioni continue e limitate:

$$\|\mu\|_{VT} = \sup_{\phi \in C_b, \|\phi\|_\infty \leq 1} \left| \int \phi d\mu \right|.$$

È facile osservare che si può rimuovere il modulo da questa definizione, infatti se $\psi = \phi \operatorname{sgn}(\int \phi d\mu)$, si ha $\int \psi d\mu = |\int \phi d\mu|$. Ne segue che la **distanza in variazione totale** tra due misure di probabilità μ e ν si può ottenere come

$$\|\mu - \nu\|_{VT} = \sup_{\phi \in C_b, \|\phi\|_\infty \leq 1} \int \phi d\mu - \int \phi d\nu.$$

Cambiando l'insieme delle funzioni ammissibili ϕ , si ottengono altre distanze, in particolare quella che stiamo cercando.

D'ora un poi mi limito a considerare misure di probabilità definite su un compatto K di \mathbb{R}^d , e indicherò l'insieme di queste misure con $\mathcal{P}(K)$.

Definizione di distanza 1-lipschitziana

Indico con $\operatorname{Lip}_1 = \operatorname{Lip}_1(K)$ l'insieme delle funzioni lipschitziane da K in \mathbb{R} con costante di Lipschitz pari a 1:

$$\operatorname{Lip}_1 = \{\phi \in C(K) \mid \forall \mathbf{x}, \mathbf{y}, |\phi(\mathbf{x}) - \phi(\mathbf{y})| \leq |\mathbf{x} - \mathbf{y}|\}$$

¹Per essere rigorosi, una misura di uno spazio di Hausdorff è una misura di Borel se è una misura σ -additiva definita sulla σ -algebra dei boreliani (la più piccola σ -algebra che contiene gli aperti). Lo spazio delle misure di Borel finite e con segno su \mathbb{R}^d coincide con il duale delle funzioni continue e limitate. Lo spazio delle misure boreliane con segno è sufficiente per fare probabilità in \mathbb{R}^d .

Date due misure μ, ν in $\mathcal{P}(K)$, la loro distanza 1-lipschitziana è

$$D_1(\mu, \nu) = \sup_{\phi \in \text{Lip}_1} \int \phi d\mu - \int \phi d\nu.$$

Per capire come funziona, consideriamo il caso semplice $\mu = \delta_{\mathbf{x}_1}, \nu = \delta_{\mathbf{x}_2}$. Per ogni $\phi \in \text{Lip}_1$,

$$\int \phi d\mu - \int \phi d\nu = \phi(\mathbf{x}_1) - \phi(\mathbf{x}_2) \leq |\mathbf{x}_1 - \mathbf{x}_2|$$

e si raggiunge l'uguaglianza scegliendo come ϕ la funzione $\phi(\mathbf{x}) = |\mathbf{x} - \mathbf{x}_2|$. Dunque $D_1(\mu, \nu) = |\mathbf{x}_1 - \mathbf{x}_2|$. Abbiamo dunque definito una metrica che tiene in conto la distanza spaziale tra le due misure.

Come secondo esempio, consideriamo due mappe Φ^1 e Φ^2 , una misura μ e le misure $\nu_i = \Phi_{\#}^i \mu$, $i = 1, 2$. Mostriamo che la distanza tra queste due misure è stimata dalla differenza tra le due mappe. Per ogni $\phi \in \text{Lip}_1$,

$$\int \phi d\nu_1 - \int \phi d\nu_2 = \int (\phi(\Phi^1(\mathbf{x})) - \phi(\Phi^2(\mathbf{x})))\mu(d\mathbf{x}) \leq \sup_{\mathbf{x}} |\Phi^1(\mathbf{x}) - \Phi^2(\mathbf{x})|$$

quindi $D_1(\Phi_{\#}^1 \mu, \Phi_{\#}^2 \mu)$ è controllato da $\|\Phi^1 - \Phi^2\|_{\infty}$.

Proposizione 4.1. D_1 è una metrica.

Dimostrazione.

• D_1 è ben definita.

Fissato $\mathbf{x}_0 \in K$, ogni funzione 1-lipschitziana ϕ si può scrivere come $\phi(\mathbf{x}) = \phi(\mathbf{x}_0) + \psi(\mathbf{x})$, con $\psi(\mathbf{x}) = \phi(\mathbf{x}) - \phi(\mathbf{x}_0)$. La funzione ψ è ancora 1-lipschitziana, inoltre, poiché K è compatto, $|\psi| \leq a$, dove a è il diametro di K . Ne segue che

$$\int \phi d\mu - \int \phi d\nu = \int \psi d\mu - \int \psi d\nu \leq 2a,$$

dunque D_1 esiste perché l'estremo superiore è limitato.

• D_1 è simmetrica.

È sufficiente scegliere $\phi = -\psi$, e passare all'estremo superiore su ψ .

• D_1 è semidefinita positiva.

Scegliendo $\phi \equiv 0$, si ottiene che $D_1 \geq 0$. Inoltre $D_1(\mu, \mu) = 0$.

• D_1 verifica la disuguaglianza triangolare.

Date $\mu, \nu, \rho \in \mathcal{P}(K)$ si ha

$$\int \phi(d\mu - d\nu) = \int \phi(d\mu - d\rho) + \int \phi(d\rho - d\nu) \leq D_1(\mu, \rho) + D_1(\rho, \nu).$$

Passando all'estremo superiore in ϕ si ottiene la tesi.

• D_1 è definita positiva.

Supponiamo che $D_1(\mu, \nu) = 0$. Allora per ogni $\phi \in \text{Lip}_1$ deve valere $\int \phi d\mu = \int \phi d\nu$. Se così non fosse, si avrebbe $D_1(\mu, \nu) > 0$. Inoltre questa uguaglianza si estende a tutte le funzioni lipschitziane (basta dividere per la costante di Lipschitz). Sia ora α una funzione

continua. Provo che $\int \alpha(d\mu - d\nu) = 0$. Per ogni $\varepsilon > 0$ esiste una ψ_ε lipschitziana tale che $\|\alpha - \psi_\varepsilon\|_\infty < \varepsilon$ (per esempio si può invocare il teorema di Weierstrass). Dunque

$$\left| \int \alpha(d\mu - d\nu) \right| = \int |\alpha - \psi_\varepsilon|(d\mu + d\nu) \leq \varepsilon.$$

Poiché le misure sono identificate da come agiscono sulle funzioni continue, ne segue che μ e ν sono uguali. □

Consideriamo ora l'insieme delle misure di probabilità $\mathcal{P}(K)$ dotato della topologia della *convergenza debole nel senso delle misure*, cioè $\mu_n \rightarrow \mu$ se e solo se $\forall \phi \in C(K)$ si ha

$$\int \phi d\mu_n \rightarrow \int \phi d\mu.$$

Teorema 4.2. D_1 è una metrica per la convergenza debole, inoltre $\mathcal{P}(K)$ è completo e compatto.

Dimostrazione. Se $D_1(\mu_n, \mu)$, per ogni funzione test Lipschitziana ϕ , si ha

$$\left| \int \phi(d\mu_n - d\mu) \right| \leq LD_1(\mu_n, \mu) \rightarrow 0.$$

Sia α continua. Per ogni $\varepsilon > 0$ esiste ϕ_ε lipschitziana tale che $\|\alpha - \phi_\varepsilon\|_\infty < \varepsilon$. Ne segue

$$\left| \int \alpha(d\mu_n - d\mu) \right| \leq \int |\alpha - \phi_\varepsilon|(d\mu_n + d\mu) + \left| \int \phi_\varepsilon(d\mu_n - d\mu) \right|.$$

Il primo termine è minore di 2ε per ogni n ; a ε fissato il secondo tende a 0 in n , dunque è minore di ε per n sufficientemente grande. Se ne conclude che $\int \alpha(d\mu_n - d\mu) \rightarrow 0$, cioè $\mu_n \rightarrow \mu$.

Viceversa, suppongo che $\mu_n \rightarrow \mu$. In particolare, per ogni $\phi \in \text{Lip}_1$,

$$\int \phi(d\mu_n - d\mu) \rightarrow 0.$$

Dalla definizione di D_1 , che è un estremo superiore, si ha che deve esistere $\phi_n \in \text{Lip}_1$ tale che

$$D_1(\mu_n, \mu) \leq \int \phi_n(d\mu_n - d\mu) + \frac{1}{n}.$$

Come abbiamo già fatto, possiamo sostituire a ϕ_n la funzione $\psi_n(\mathbf{x}) = \phi_n(\mathbf{x}) - \phi_n(\mathbf{x}_0)$, dove \mathbf{x}_0 è un qualunque fissato punto di K . Le funzioni ψ_n sono 1-lipschitziane e sono limitate dal diametro di K , dunque sono equicontinue ed equilimitate. Il teorema di Ascoli-Arzelà permette di ottenere che a meno di sottosequenze ψ_n converge uniformemente a $\psi \in \text{Lip}_1$. Indico ancora con n la sottosequenza. Si ha

$$D_1(\mu_n, \mu) \leq \int \psi_n(d\mu_n - d\mu) + \frac{1}{n} = \int \psi(d\mu_n - d\mu) + \int (\psi_n - \psi)(d\mu_n - d\mu) + \frac{1}{n}$$

Il primo termine va a 0 perché $\mu_n \rightarrow \mu$, il secondo perché è stimato da $2\|\psi_n - \psi\|_\infty$. Si ottiene che per la sottosequenza estratta $D_1(\mu_n, \mu) \rightarrow 0$. Per concludere la prova bisogna

dimostrare che tutta la successione tende a 0. L'argomento è il seguente. Per assurdo, se $D_1(\mu_n, \mu)$ non tendesse a 0, esisterebbe $\bar{\varepsilon} > 0$ e una sottosuccessione n_k tale che $D_1(\mu_{n_k}, \mu) \geq \bar{\varepsilon}$. Ripetendo l'argomento precedente, si otterrebbe che per una sottosuccessione di n_k il limite sarebbe 0. Questo è assurdo, dunque la successione tende a 0.

La prova della completezza è semplice: se μ_n è di Cauchy, allora per ogni $\phi \in \text{Lip}_1$, $\int \phi d\mu_n$ è di Cauchy. Passando al limite si ottiene un funzionale lineare sulle funzioni lipschitziane limitato nella norma dell'estremo superiore. Per densità si estende alle funzioni continue e dunque si identifica il limite con un funzionale lineare continuo, cioè una misura boreliana con segno μ . Per ogni n si ha che per ogni osservabile ϕ positivo $\int \phi d\mu_n \geq 0$ e $\int d\mu_n = 1$. Passando al limite in queste due relazioni si ottiene che μ è una misura di probabilità.

La compattezza (sequenziale) dello spazio delle misure per la convergenza debole si può ottenere da risultati di analisi funzionale, qui mi limito a dare un'indicazione del perché è vero con un esempio: sia μ_n una misura di probabilità sull'intervallo $[0, 1]$. La probabilità di $[0, 1/2]$ e quella di $(1/2, 1]$ sono due successioni positive limitate, dunque per sottosequenze convergono e la somma delle probabilità rimane 1 (non può disperdersi massa perché siamo su un compatto). Si può iterare questa procedura dividendo in quattro sottointervalli e ottenere una sottosequenza della sottosequenza per cui convergono le probabilità di $[0, 1/4]$, $(1/4, 1/2]$, $(1/2, 3/4]$, $(3/4, 1]$. Sempre passando a sottosequenze si possono anche le convergenze sulla misura degli estremi di questi intervalli. Iterando e utilizzando la sottosequenza diagonale, si trova una unica sottosequenza per cui esistono i limiti di $\mu_n([j/2^h, (j+1)/2^h])$, con $j = 0, \dots, 2^h - 1$, e degli analoghi intervalli chiusi o semichiusi. È facile convincersi che questi limiti definiscono una misura di probabilità. □

Termino questa sezione con una osservazione: nel caso di misure definite su illimitati, la distanza $D_1(\mu, \nu)$ è infinita se $\int |\mathbf{x}| d\mu(d\mathbf{x})$ è finito e $\int |\mathbf{x}| d\nu(d\mathbf{x}) = +\infty$ (lo si mostri per esercizio). La convergenza debole è ancora metrizzabile, ma non con D_1 .

Sunto

- La distanza D_1 è una metrica per $\mathcal{P}(K)$, con K compatto, che riesce a misurare gli effetti di un trasporto di massa.
- D_1 induce la topologia della convergenza debole nel senso delle misure su $\mathcal{P}(K)$, rispetto alla quale $\mathcal{P}(K)$ è sequenzialmente compatto.

4.2 Il problema di Monge-Kantorovich

Il problema di definire metriche che si comportano bene rispetto al trasporto di misura è strettamente legato a questioni di ottimizzazione nella teoria analitica dei problemi di trasporto di massa. L'idea è che la distanza tra due misure di probabilità sia data da quanto costa trasportare l'una dell'altra. Questo problema si chiama problema di Monge, perché Monge se ne occupò cercando il modo più economico per trasportare palle di cannone da diversi depositi in diverse zone di impiego.

Definizione - il problema di Monge

Siano date due misure di probabilità μ e ν . Il problema di Monge consiste nel trovare la mappa Φ , se c'è, che realizza

$$\inf_{\Phi: \nu = \Phi \# \mu} \int |\mathbf{x} - \Phi(\mathbf{x})| \mu(d\mathbf{x})$$

Per esempio considero due misure singolari $\mu(d\mathbf{x}) = \delta_{\mathbf{x}_1}(d\mathbf{x})$, $\nu(d\mathbf{x}) = \delta_{\mathbf{y}_1}(d\mathbf{x})$, con $\mathbf{x}_1 \neq \mathbf{y}_1$. La distanza in variazione totale tra di esse è 2. Non è difficile comprendere che la migliore mappa di trasporto è la traslazione $\Phi_{\mathbf{h}}(\mathbf{x}) = \mathbf{x} + \mathbf{h}$, con $\mathbf{h} = \mathbf{y}_1 - \mathbf{x}_1$. È immediato verificare che $\Phi_{\mathbf{h}} \# \mu = \nu$, e che il costo è $|\mathbf{h}|$.

Poiché la misura di partenza è singolare, non è necessario specificare tutta Φ , ma solo trasportare la massa che è in \mathbf{x}_1 in \mathbf{y}_1 . Generalizzo.

Definizione - matching bipartito euclideo

Siano date due N -uple di punti di \mathbb{R}^d , $\mathbf{x}_i, \mathbf{y}_i, i = 1, \dots, N$. Sia $\mu = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{x}_i}(d\mathbf{x})$ e sia $\nu = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{y}_i}(d\mathbf{x})$. Il problema di Monge consiste nel trovare il miglior matching bipartito, cioè la migliore permutazione σ degli indici che realizzi

$$\frac{1}{N} \inf_{\sigma} \sum |\mathbf{x}_i - \mathbf{y}_{\sigma_i}|$$

Infatti le permutazioni di indici identificano le mappe che assegnano a ognuno dei punti \mathbf{x}_i uno dei punti \mathbf{y}_i , in modo biiettivo, cioè realizzano il trasporto di μ in ν .

Si può però immaginare un altro modo di procedere, rompendo il vincolo di trasportare tutta la massa $1/N$ in \mathbf{x}_i nel punto \mathbf{y}_j , ma ipotizzando di spostare parte della massa in vari punti di arrivo.

Definizione - il problema di Kantorovich discreto

Siano date due N -uple di punti di \mathbb{R}^d , $\mathbf{x}_i, \mathbf{y}_i, i = 1, \dots, N$. Sia $\mu = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{x}_i}(d\mathbf{x})$ e sia $\nu = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{y}_i}(d\mathbf{x})$. Con il nome di **piano di trasporto** si intende una matrice $\{a_{ij}\}_{i,j=1}^N$ di elementi positivi tali che la somma su ogni riga e su ogni colonna sia $1/N$. Il problema di Kantorovich consiste nel trovare il piano di trasporto che realizza

$$\inf_a \sum_{i,j} a_{ij} |\mathbf{x}_i - \mathbf{y}_j|.$$

Il termine $a_{ij} |\mathbf{x}_i - \mathbf{y}_j|$ è il costo per spostare la parte di massa a_{ij} da \mathbf{x}_i in \mathbf{y}_j . Si può riformulare questo problema osservando che

$$\gamma(d\mathbf{x}, d\mathbf{y}) = \sum_{ij} a_{ij} \delta_{\mathbf{x}_i}(d\mathbf{x}) \delta_{\mathbf{y}_j}(d\mathbf{y})$$

è una distribuzione congiunta di μ e ν , e tutte le distribuzioni congiunte di μ e ν sono di questo tipo. Un piano di trasporto è dunque una distribuzione congiunta, e il problema di Kantorovich consiste nel trovare

$$\inf_{\gamma} \int |\mathbf{x} - \mathbf{y}| \gamma(d\mathbf{x}, d\mathbf{y})$$

È un risultato classico dell'ottimizzazione convessa il fatto che il problema di Kantorovich è risolto da

$$\gamma(d\mathbf{x}, d\mathbf{y}) = \sum_i \frac{1}{N} \delta_{\mathbf{x}_i}(d\mathbf{x}) \delta_{\mathbf{y}_{\sigma_i}}(d\mathbf{y})$$

dove σ è la permutazione che risolve il problema di Monge.

Definizione - il problema di Kantorovich in generale

Siano date due misure μ e ν . Il problema di Kantorovich consiste nel trovare il piano di trasporto che realizza

$$\inf_{\gamma} \int |\mathbf{x} - \mathbf{y}| \gamma(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y})$$

dove γ varia nell'insieme $\Gamma(\mu, \nu)$ delle misure congiunte di μ e ν .

In generale, il problema di Monge potrebbe non avere soluzione, per esempio se μ è una misura discreta su N punti e ν su $M > N$ punti, non esiste nessuna mappa Φ tale che $\nu = \Phi\#\mu$, mentre l'insieme dei piani di trasporto è sempre non vuoto, perché la distribuzione prodotto di μ e ν è una distribuzione congiunta. D'altra parte, se Φ trasporta μ in ν , la misura

$$\gamma = \mu \otimes (\Phi\#\mu)$$

è un piano di trasporto.

Sunto

- Il problema di Kantorovich si può sempre formulare.
- Il costo secondo Kantorovich è minore o uguale al costo secondo Monge.
- Si può provare che se il problema di Monge ha soluzione, allora i due costi coincidono. In questo senso, il problema di Kantorovich è un **rilassamento** del problema di Monge

4.3 La distanza di Wassertein

Dato $p \in [1, +\infty)$, su $\mathcal{P}(K)$, si definisce la distanza p -Wasserstein come

$$W_p(\mu, \nu) = \inf_{\gamma} \left(\int |\mathbf{x} - \mathbf{y}|^p \gamma(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y}) \right)^{1/p}.$$

La distanza 1-Wasserstein prende anche il nome di distanza di Kantorovich-Rubinstein.

Teorema 4.3 (Dualità di Kantorovich).

$$W_1(\mu, \nu) = \sup_{\varphi, \psi \in C(K): \varphi(x) + \psi(y) \leq |\mathbf{x} - \mathbf{y}|} \int \varphi \mathrm{d}\mu + \int \psi \mathrm{d}\nu. \quad (15)$$

Questo teorema ha una interessante interpretazione "economica". La distanza W_1 rappresenta il più piccolo costo che per trasportare μ in ν . Data una γ , $\int |\mathbf{x} - \mathbf{y}| \gamma(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y})$ è il costo che si paga per trasportare μ in ν secondo il piano di trasporto γ .

Un altro modo di guardare a questo costo è misurare il costo per "raccogliere" $\mu(\mathrm{d}\mathbf{x})$ in \mathbf{x} , quantificato in $\phi(\mathbf{x})\mu(\mathrm{d}\mathbf{x})$, e il costo di distribuire quanto raccolto secondo ν in \mathbf{y} quantificato in $\psi(\mathbf{y})\nu(\mathrm{d}\mathbf{y})$. Affinché questo modo sia "conveniente", deve valere il vincolo $\phi(\mathbf{x}) + \psi(\mathbf{y}) \leq |\mathbf{x} - \mathbf{y}|$. In tal caso infatti

$$\int \phi \mathrm{d}\mu + \int \psi \mathrm{d}\nu = \int (\phi(\mathbf{x}) + \psi(\mathbf{y})) \gamma(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y}) \leq \int |\mathbf{x} - \mathbf{y}| \gamma(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y}).$$

La dualità di Kantorovich asserisce che il massimo prezzo che si può chiedere per ritirare μ e ridistribuire in ν è proprio il costo minimo di trasporto.

Dalla disuguaglianza precedente si ottiene facilmente che il membro di destra della (15) è minore o uguale di W_1 . Il viceversa non è semplice, una sua possibile prova richiedere la nozione di trasformata di Legendre su spazi di Banach, è disponibile sulle dispense del 2024/2025. Qui mi limito a dare un'idea del perché questo risultato è vero.

Userò le seguenti notazioni “compatte”: date due misure di probabilità μ e ν , e date due funzioni continue φ e ψ :

$$\begin{aligned} d(\mu \otimes \nu) &= \mu \otimes \nu(d\mathbf{x}, d\mathbf{y}) = \mu(d\mathbf{x})\nu(d\mathbf{y}) \\ \varphi \oplus \psi &= (\varphi \oplus \psi)(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{x}) + \psi(\mathbf{y}) \end{aligned}$$

In particolare

$$\int \varphi \oplus \psi d(\mu \otimes \nu) = \int \varphi d\mu + \int \psi d\nu.$$

Si noti che $d(\mu \otimes \nu) \in \Gamma(\mu, \nu)$, infatti è la misura di probabilità congiunta che si ottiene pensando di estrarre \mathbf{x} con μ , e \mathbf{y} con ν , indipendenti. Inoltre indicherò con $c = c(\mathbf{x}, \mathbf{y})$ il costo $|\mathbf{x} - \mathbf{y}|$.

L'idea per provare la dualità è quella di usare due funzioni φ e ψ per esprimere il vincolo sulle distribuzioni congiunte. Sia $\mathcal{M}(K \times K)$ l'insieme delle misure con segno, cioè il duale di $C(K \times K)$, sia \mathcal{M}^+ il sottoinsieme delle misure positive. Evidentemente

$$\Gamma(\mu, \nu) \subset \mathcal{M}^+ \subset \mathcal{M}.$$

Consideriamo la differenza

$$\int \varphi \oplus \psi d(\mu \otimes \nu) - \int \varphi \oplus \psi d\gamma$$

per una generica $\gamma \in \mathcal{M}$. Si noti che questa differenza vale

$$\begin{aligned} &\int \varphi(\mathbf{x})\mu(d\mathbf{x}) + \int \psi(\mathbf{y})\nu(d\mathbf{y}) - \int (\varphi(\mathbf{x}) + \psi(\mathbf{y}))\gamma(d\mathbf{x}, d\mathbf{y}) = \\ &\int \varphi(\mathbf{x})(\mu(d\mathbf{x}) - \gamma^1(d\mathbf{x})) + \int \psi(\mathbf{y})(\nu(d\mathbf{y}) - \gamma^2(d\mathbf{y})) \end{aligned}$$

dove γ^1 e γ^2 sono le due marginali di γ . Se $\gamma \in \Gamma(\mu, \nu)$, la differenza è 0 qualunque siano φ, ψ . Al contrario, se $\gamma \notin \Gamma(\mu, \nu)$ $\mu \neq \gamma^1$ o $\nu \neq \gamma^2$ (o entrambi), e dunque esistono $\bar{\varphi}, \bar{\psi}$ su cui la differenza da un valore non nullo. Pertanto

$$\sup_{\varphi, \psi \in C(K)} \left(\int \varphi \oplus \psi d(\mu \otimes \nu) - \int \varphi \oplus \psi d\gamma \right) = \begin{cases} 0 & \text{se } \gamma \in \Gamma(\mu, \nu) \\ +\infty & \text{altrimenti} \end{cases}$$

infatti nel primo caso la differenza è nulla per ogni coppia di funzioni, nel secondo caso moltiplicando la coppia $(\bar{\varphi}, \bar{\psi})$ per una costante arbitraria, opportuna, si ottiene un qualunque valore reale, pertanto il sup è infinito.

Nella definizione di W_1 posso considerare invece dell'estremo inferiore su $\gamma \in \Gamma$, all'estremo inferiore su $\gamma \in \mathcal{M}^+$, pagando però $+\infty$ se $\gamma \notin \Gamma$:

$$W_1(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu, \nu)} \int c d\gamma = \inf_{\gamma \in \mathcal{M}^+} \sup_{\varphi, \psi} \int c d\gamma + \int \varphi \oplus \psi d(\mu \otimes \nu) - \int \varphi \oplus \psi d\gamma. \quad (16)$$

(ricordo che il costo c è $c = |\mathbf{x} - \mathbf{y}|$). Per una qualunque funzione F di due variabili si prova facilmente che in generale

$$\inf_{a \in A} \sup_{b \in B} F(a, b) \geq \sup_{b \in B} \inf_{a \in A} F(a, b).$$

In qualche caso vale l'uguaglianza, in particolare si dimostra che se F è definita su un compatto convesso, e se inoltre $F(a, b)$ è concava in a per ogni b fissato, ed è convessa in b per ogni a fissato, allora la coppia (a, b) che realizza gli estremi è la stessa indipendentemente dall'ordine, e individua un punto di sella di F . Nel nostro caso, convessità e concavità sono banali, perché la funzione in (16) è lineare nella coppia (φ, ψ) e dunque concava, e lineare in γ , ed è effettivamente definita su un convesso. Non è però soddisfatta la condizione di compattezza. Esistono estensioni di questo risultato che permettono di provare la tesi della scambiabilità degli estremi, che assumerò dunque vera senza altri dettagli.

Dunque

$$\inf_{\gamma \in \Gamma(\mu, \nu)} \int c \, d\gamma = \sup_{\varphi, \psi} \inf_{\gamma \in \mathcal{M}^+} \int c \, d\gamma + \int \varphi \oplus \psi \, d(\mu \otimes \nu) - \int \varphi \oplus \psi \, d\gamma$$

Noto che il calcolo dell'estremo inferiore riguarda solo i due termini in γ :

$$\inf_{\gamma \in \mathcal{M}^+} \int (c - \varphi \oplus \psi) \, d\gamma$$

È evidente che se $\varphi \oplus \psi \leq c$ l'estremo inferiore è raggiunto su $\gamma = 0$, e vale 0. Al contrario, se esistono valori $(\mathbf{x}_0, \mathbf{y}_0)$ tali che $\varphi(\mathbf{x}_0) + \psi(\mathbf{y}_0) > |\mathbf{x}_0 - \mathbf{y}_0|$ basta concentrare γ su questi valori, con valori divergenti di massa per ottenere che l'inf è $-\infty$. Dunque l'estremo inferiore su \mathcal{M}^+ si ottiene dove la coppia di funzioni φ, ψ soddisfa il vincolo $\varphi \oplus \psi \leq c$. Abbiamo dunque provato che

$$W_1(\mu, \nu) = \sup_{\varphi, \psi \in C(K), \varphi \oplus \psi \leq c} \int \varphi \oplus \psi \, d(\mu \otimes \nu).$$

Teorema 4.4 ($W_1 = D_1$).

La distanza di Wasserstein W_1 è uguale alla distanza 1-lipschitziana D_1 .

Dimostrazione. La prima osservazione è che se $\varphi \in \text{Lip}_1$ la coppia di funzioni $\varphi(\mathbf{x}), \psi(\mathbf{y}) = -\varphi(\mathbf{y})$ verifica il vincolo richiesto, dunque

$$\sup_{\varphi, \psi: \varphi(\mathbf{x}) + \psi(\mathbf{y}) \leq |\mathbf{x} - \mathbf{y}|} \int \varphi \, d\mu + \int \psi \, d\nu \geq \sup_{\varphi \in \text{Lip}_1} \int \varphi \, (d\mu - d\nu),$$

e dunque abbiamo facilmente che $W_1 \geq D_1$. Dimostro la disuguaglianza opposta. Date φ e ψ con $\varphi(\mathbf{x}) + \psi(\mathbf{y}) \leq |\mathbf{x} - \mathbf{y}|$, osservo che

$$\varphi(\mathbf{x}) \leq \psi^*(\mathbf{x}) := \inf_{\mathbf{y}} |\mathbf{x} - \mathbf{y}| - \psi(\mathbf{y}),$$

infatti per ogni \mathbf{y} , $\varphi(\mathbf{x}) \leq |\mathbf{x} - \mathbf{y}| - \psi(\mathbf{y})$, e dunque $\varphi(\mathbf{x})$ è minore dell'estremo inferiore in \mathbf{y} . Si ha dunque

$$\int \varphi \, d\mu \leq \int \psi^* \, d\mu$$

D'altra parte, per la definizione di Ψ^* , si ha

$$\psi^*(\mathbf{x}) + \psi(\mathbf{y}) \leq |\mathbf{x} - \mathbf{y}|$$

da cui segue

$$\psi(\mathbf{y}) \leq \psi^{**}(\mathbf{y}) := \inf_{\mathbf{x}} |\mathbf{x} - \mathbf{y}| - \psi^*(\mathbf{x}).$$

Pertanto

$$\sup_{\varphi, \psi: \varphi(\mathbf{x}) + \psi(\mathbf{y}) \leq |\mathbf{x} - \mathbf{y}|} \int \varphi \, d\mu + \int \psi \, d\nu = \sup_{\psi} \left(\int \psi^* \, d\mu + \int \psi^{**} \, d\nu \right)$$

Mostro che per costruzione ψ^* è una funzione 1-lipschitziana, infatti per ogni \mathbf{x}, \mathbf{x}_1 ,

$$\begin{aligned} \psi^*(\mathbf{x}) &= \inf_{\mathbf{y}} (|\mathbf{x} - \mathbf{y}| - \psi(\mathbf{y})) \leq \inf_{\mathbf{y}} (|\mathbf{x}_1 - \mathbf{y}| + |\mathbf{x} - \mathbf{x}_1| - \psi(\mathbf{y})) \\ &= \inf_{\mathbf{y}} (|\mathbf{x}_1 - \mathbf{y}| - \psi(\mathbf{y})) + |\mathbf{x} - \mathbf{x}_1| = \psi^*(\mathbf{x}_1) + |\mathbf{x} - \mathbf{x}_1|. \end{aligned}$$

Scambiando \mathbf{x} con \mathbf{x}_1 e usando le due disuguaglianze, si ottiene che $\psi^* \in \text{Lip}_1$. Dunque

$$\sup_{\varphi, \psi: \varphi(\mathbf{x}) + \psi(\mathbf{y}) \leq |\mathbf{x} - \mathbf{y}|} \int \varphi \, d\mu + \int \psi \, d\nu \leq \sup_{\varphi \in \text{Lip}_1} \int \varphi \, d\mu + \int \varphi^* \, d\nu.$$

La prova si conclude facilmente mostrando che se φ è una funzione 1-lipschitziana, allora

$$\varphi^*(\mathbf{y}) = -\varphi(\mathbf{y}).$$

Dimostro quest'ultimo fatto.

$$\varphi^*(\mathbf{y}) = \inf_{\mathbf{x}} (|\mathbf{x} - \mathbf{y}| - \varphi(\mathbf{x})) = \inf_{\mathbf{x}} (|\mathbf{x} - \mathbf{y}| - (\varphi(\mathbf{x}) - \varphi(\mathbf{y})) - \varphi(\mathbf{y}))$$

La funzione di cui si cerca l'estremo inferiore è non negativa, perché $|\mathbf{x} - \mathbf{y}| - (\varphi(\mathbf{x}) - \varphi(\mathbf{y})) \geq 0$, d'altra parte è nulla in $\mathbf{x} = \mathbf{y}$, dunque l'estremo cercato è 0, e da questo fatto segue la tesi. \square

Sunto

- La dualità di Kantorovich trasforma il problema del calcolo del minimo costo per spostare μ in ν , in un problema di massimo con vincoli.
- Nel caso di costo dato dalla distanza euclidea, la formulazione duale è equivalente al calcolo della distanza 1-lipschitziana tra μ e ν .

5 La validità dell'equazione di Vlasov

5.1 Costruzione delle soluzioni

Ricordo che vogliamo risolvere l'equazione di Vlasov

$$\begin{aligned} \partial_t f_t + \operatorname{div}_z (f_t \mathcal{K}[f_t]) &= 0 \\ \mathcal{K}[f](\mathbf{z}) &:= \int \mathbf{G}(\mathbf{z}, \tilde{\mathbf{z}}) f(\tilde{\mathbf{z}}) = 0. \end{aligned} \tag{17}$$

La riscrivo in forma debole per una “traiettoria” di misure μ_t , $t \geq 0$. Sia

$$\mathcal{K}[\mu](t, \mathbf{z}) := \int \mathbf{G}(\mathbf{z}, \tilde{\mathbf{z}}) \mu_t(d\tilde{\mathbf{z}})$$

e, supposta nota μ_t , sia $\mathbf{Z}_t^\mu(\mathbf{z})$ generato dal campo $\mathcal{K}[\mu]$ (a rigore il flusso è a due parametri, sto sottoindicando che considero solo dati iniziali al tempo 0). L'equazione si riformula per le misure in forma debole come

$$\mu_t = \mathbf{Z}_t^\mu \# \mu_0.$$

Questa espressione non è la soluzione, perché non conosco il flusso se non conosco la soluzione stessa. Osservo che non è necessario supporre regolarità per μ , perché il push forward non lo richiede e per l'esistenza del flusso è sufficiente la regolarità di \mathbf{G} , come vedremo.

Queste considerazioni si concretizzeranno in un teorema di costruzione delle soluzioni per dati iniziali misure di probabilità. Come ipotesi supporrò che μ_0 sia a supporto limitato, e che la funzione \mathbf{G} sia localmente limitata e lipschitziana. L'idea è che, data ν_t con $\nu_0 = \mu_0$, costruisco un flusso \mathbf{Z}_t^ν , e con questo flusso costruisco $\mathbf{Z}_t^\nu \# \mu_0$, ottenendo una nuova traiettoria nello spazio delle misure di probabilità. Mostrando che questa mappa è contrattiva otterrò l'esistenza e unicità delle soluzioni.

Lo faremo in parte attraverso due proposizioni; la prima che assicura che data μ_t possiamo costruire un flusso con buone proprietà, la seconda che assicura che dato un flusso possiamo trasportare μ_0 in μ_t in modo da soddisfare le ipotesi che permettono di costruire il flusso. Infine, dovremo mettere insieme questi due aspetti.

Lemma 5.1 (Regolarità del campo e del flusso). *Sia \mathbf{G} una funzione limitata e lipschitziana di costante di Lipschitz L . Sia μ_t una famiglia di misure di probabilità, con $t \in [0, T]$, con i supporti tutti contenuti in un chiuso limitato B , e sia inoltre debolmente continua in t , cioè $D_1(\mu_t, \mu_s) \rightarrow 0$ se $s \rightarrow t$.*

Allora $\mathcal{K}[\mu](t, \mathbf{z})$ è un campo limitato e lipschitziano di costante di Lipschitz L , inoltre è continuo in $t \in [0, T]$. Dunque il flusso associato \mathbf{Z}_t^μ esiste, ed è lipschitziano di costante di Lipschitz e^{Lt} .

Dimostrazione. Il campo è dato da

$$\mathcal{K}[f](t, \mathbf{z}) = \int \mathbf{G}(\mathbf{z}, \mathbf{w}) \mu_t(d\mathbf{w}).$$

Dunque se $|G| \leq M$ per una qualche costante M ,

$$|\mathcal{K}[f](t, \mathbf{z})| \leq M \int \mu_t(d\mathbf{z}) = M.$$

Inoltre, dati \mathbf{z}_1 e \mathbf{z}_2 ,

$$\begin{aligned} |\mathcal{K}[f](t, \mathbf{z}_1) - \mathcal{K}[f](t, \mathbf{z}_2)| &\leq \int |\mathbf{G}(\mathbf{z}_1, \mathbf{w}) - \mathbf{G}(\mathbf{z}_2, \mathbf{w})| \mu_t(d\mathbf{w}) \\ &\leq L|\mathbf{z}_1 - \mathbf{z}_2| \int \mu_t(d\mathbf{w}) = L|\mathbf{z}_1 - \mathbf{z}_2|. \end{aligned}$$

Provo la continuità nel tempo. Si ha

$$|\mathcal{K}[f](t, \mathbf{z}) - \mathcal{K}[f](s, \mathbf{z})| \leq \left| \int \mathbf{G}(\mathbf{z}, \mathbf{w})(\mu_t(d\mathbf{w}) - \mu_s(d\mathbf{w})) \right| \leq LD_1(\mu_t, \mu_s)$$

dove, nell'ultimo passaggio, ho usato la definizione di distanza D_1 , sfruttando la lipschitzianità di \mathbf{G} nella seconda variabile.

Quanto provato finora garantisce l'esistenza del flusso. La lipschitzianità del flusso è conseguenza della lipschitzianità del campo. Infatti, dati \mathbf{z}_1 e \mathbf{z}_2 ,

$$\begin{aligned} |\mathbf{Z}_t^\mu(\mathbf{z}_1) - \mathbf{Z}_t^\mu(\mathbf{z}_2)| &\leq |\mathbf{z}_1 - \mathbf{z}_2| + \int_0^t ds |\mathcal{K}(s, \mathbf{Z}_s^\mu(\mathbf{z}_1)) - \mathcal{K}(s, \mathbf{Z}_s^\mu(\mathbf{z}_2))| \\ &\leq |\mathbf{z}_1 - \mathbf{z}_2| + L \int_0^t ds |\mathbf{Z}_s^\mu(\mathbf{z}_1) - \mathbf{Z}_s^\mu(\mathbf{z}_2)| \end{aligned}$$

(per la prima disuguaglianza ho usato la versione integrale dell'equazione che definisce il flusso, per la seconda ho usato la lipschitzianità del campo). Applicando il lemma di Gronwall si ottiene che $\mathbf{Z}_t^\mu(\mathbf{z})$ ha costante di Lipschitz stimata da e^{Lt} . □

Il caso fisicamente più interessante è quello dei sistemi conservativi con interazione di coppia, da cui siamo partiti, in tal caso \mathbf{G} è lipschitziana, ma non è limitata (una delle sue componenti è \mathbf{v}). Non è difficile superare questa difficoltà, perché come si vedrà serve solo la limitatezza sul supporto di μ_t . Lascio i dettagli al lettore. Ricordo anche che il caso reale è quello di interazione con potenziale gravitazionale o elettrostatico, caso per cui è tutto più difficile e interessante.

Proposizione 5.2 (Dipendenza del flusso dalla distribuzione).

Siano μ_t e ν_t due famiglie di misure di probabilità, con $t \in [0, T]$, con i supporti tutti contenuti in un chiuso limitato B . Siano inoltre debolmente continue in t , cioè $D_1(\mu_t, \mu_s), D_1(\nu_t, \nu_s) \rightarrow 0$ se $s \rightarrow t$. Vale

$$|\mathbf{Z}_t^\mu(\mathbf{z}_1) - \mathbf{Z}_t^\nu(\mathbf{z}_2)| \leq e^{Lt} \left(|\mathbf{z}_1 - \mathbf{z}_2| + L \int_0^t ds D_1(\mu_s, \nu_s) \right). \quad (18)$$

dove L è la costante di Lipschitz di G .

Dimostrazione. Per la proposizione precedente, $\mathcal{K}[\mu]$ e $\mathcal{K}[\nu]$ sono funzioni continue in t , limitate e lipschitziane uniformemente in $t \in [0, T]$, pertanto i corrispondenti flussi sono ben definiti. Dimostro la (18). Scrivo i due flussi in forma integrale:

$$\mathbf{Z}_t^\mu(\mathbf{z}_1) = \mathbf{z}_1 + \int_0^t ds \mathcal{K}[\mu](s, \mathbf{Z}_s^\mu(\mathbf{z}_1))$$

$$\mathbf{Z}_t^\nu(\mathbf{z}_2) = \mathbf{z}_2 + \int_0^t ds \mathcal{K}[\nu](s, \mathbf{Z}_s^\nu(\mathbf{z}_2))$$

Sottraggo e stimo:

$$|\mathbf{Z}_t^\mu(\mathbf{z}_1) - \mathbf{Z}_t^\nu(\mathbf{z}_2)| \leq |\mathbf{z}_1 - \mathbf{z}_2| + \int_0^t ds |\mathcal{K}[\mu](s, \mathbf{Z}_s^\mu(\mathbf{z}_1)) - \mathcal{K}[\nu](s, \mathbf{Z}_s^\nu(\mathbf{z}_2))|$$

Per stimare l'ultimo termine, sommo e sottraggo il campo generato da μ ma calcolato nel flusso generato da ν , ottenendo i due termini

$$\begin{aligned} (1) &= |\mathcal{K}[\mu](s, \mathbf{Z}_s^\mu(\mathbf{z}_1)) - \mathcal{K}[\mu](s, \mathbf{Z}_s^\nu(\mathbf{z}_2))| \\ (2) &= |\mathcal{K}[\mu](s, \mathbf{Z}_s^\nu(\mathbf{z}_2)) - \mathcal{K}[\nu](s, \mathbf{Z}_s^\nu(\mathbf{z}_2))| \end{aligned}$$

Il primo termine si stima semplicemente usando la lipschitzianità del campo \mathcal{K} , con la costante di Lipschitz L che è quella di \mathbf{G} :

$$|(1)| \leq L |\mathbf{Z}_s^\mu(\mathbf{z}_1) - \mathbf{Z}_s^\nu(\mathbf{z}_2)|$$

Nel secondo caso si tratta di stimare in $\mathbf{z} = \mathbf{Z}_s^\nu(\mathbf{z}_2)$ la differenza

$$|\mathcal{K}[\mu](s, \mathbf{z}) - \mathcal{K}[\nu](s, \mathbf{z})|.$$

Questa stima è meno immediata, e serve la distanza D_1

$$\mathcal{K}[\mu](s, \mathbf{z}) - \mathcal{K}[\nu](s, \mathbf{z}) = \int \mathbf{G}(\mathbf{z} - \mathbf{z}_1) (\mu_s(d\mathbf{z}_1) - \nu_s(d\mathbf{z}_1)) \leq LD_1(\mu_s, \nu_s)$$

Mettendo insieme le due stime si ha

$$|\mathbf{Z}_t^\mu(\mathbf{z}_1) - \mathbf{Z}_t^\nu(\mathbf{z}_2)| \leq |\mathbf{z}_1 - \mathbf{z}_2| + L \int_0^t ds D_1(\mu_s, \nu_s) + L \int_0^t ds |\mathbf{Z}_s^\mu(\mathbf{z}_1) - \mathbf{Z}_s^\nu(\mathbf{z}_2)|$$

Usando il lemma di Gronwall “congelando” la dipendenza temporale del termine in D_1 , si ottiene la tesi (dettagli al lettore). □

Proposizione 5.3 (Dipendenza della distribuzione dal flusso).

Sia μ_0 una distribuzione di probabilità con supporto in un compatto B_0 . Siano μ_t^i , $i = 1, 2$ le soluzioni per $t \in [0, T]$ dell'equazione di Liouville di flussi $\mathbf{Z}_t^{\nu^i}$, generati dai campi determinati dalle misure ν_t^i , di dati iniziali μ_0^i . Le misure μ_t^i hanno supporto in un compatto B che dipende solo da T , e sono debolmente continue in t . Inoltre per T sufficientemente piccolo esistono $C > 0$ e $\gamma \in (0, 1)$ tali che

$$\sup_{t \in [0, T]} D_1(\mu_t^1, \mu_t^2) \leq CD_1(\mu_0^1, \mu_0^2) + \gamma \sup_{t \in [0, T]} D_1(\nu_t^1, \nu_t^2). \quad (19)$$

Dimostrazione. Poiché \mathbf{G} è limitato, chiamando M il suo massimo è immediato concludere che $|\mathbf{Z}_t^{\nu^i}(\mathbf{z}) - \mathbf{z}| \leq Mt$. Questa disuguaglianza implica che μ_t^i per $t \in [0, T]$ hanno supporto in un opportuno compatto B che dipende solo da B_0 e da T .

Sia $\varphi \in \text{Lip}_1$.

$$\begin{aligned} \int \varphi d\mu_t^1 - \varphi d\mu_t^2 &= \int \varphi(Z_t^{\nu_1}) d\mu_0^1 - \int \varphi(Z_t^{\nu_2}) d\mu_0^2 \\ &= \int \varphi(Z_t^{\nu_1}) d\mu_0^1 - \int \varphi(Z_t^{\nu_1}) d\mu_0^2 + \int (\varphi(Z_t^{\nu_1}) - \varphi(Z_t^{\nu_2})) d\mu_0^2 \end{aligned}$$

Il primo addendo è stimato da $D_1(\mu_0^1, \mu_0^2)$ per la costante di Lipschitz di $\varphi(Z^{\nu_1})$, che è stimata da e^{Lt} come provato precedentemente. Il secondo addendo è stimato da $\|Z_t^{\nu_1} - Z_t^{\nu_2}\|_\infty$, che abbiamo stimato nella proposizione precedente. Passando all'estremo superiore su φ si ha

$$D_1(\mu_t^1, \mu_t^2) \leq e^{Lt} D_1(\mu_0^1, \mu_0^2) + Le^{Lt} \int_0^t ds D_1(\nu_s^1, \nu_s^2).$$

Scegliendo T tale che $\gamma = LTe^{LT} < 1$ si ottiene la tesi passando all'estremo superiore in s nell'integrale temporale. \square

A questo punto è semplice costruire la soluzione.

Teorema 5.4. *Sia data μ_0 misura di probabilità al tempo 0, con supporto in un compatto B_0 . L'equazione di campo medio ammette un'unica soluzione debole $\{\mu_t\}_{t \geq 0}$ di dato iniziale μ_0 per $t \in [0, +\infty)$. La soluzione è debolmente continua nel dato iniziale.*

Dimostrazione. Sia T il tempo determinato nella precedente proposizione, e sia $\{\nu_t\}_{t=0}^T$ una famiglia di misure di probabilità a supporto in B e debolmente continua in t , con $\nu_0 = \mu_0$. Considero la mappa \mathcal{M} che associata alla famiglia di misure di probabilità ν_t la famiglia di misure di probabilità $\mathcal{M}[\nu](t) = Z_t^\nu \# \mu_0$. Per quanto provato nelle proposizioni precedenti, La mappa \mathcal{M} è definita da $C([0, T]; \mathcal{P}(B))$ in sé, ed è una contrazione, infatti

$$\sup_{t \in [0, T]} D_1(\mathcal{M}[\nu^1](t), \mathcal{M}[\nu^2](t)) \leq \gamma \sup_{t \in [0, T]} D_1(\nu_t^1, \nu_t^2),$$

con $\gamma < 1$. Quindi in un intorno di μ_0 esiste $\{\mu_t\}_{t \in [0, T]}$ tale che

$$\mathcal{M}[\mu](t) = \mu_t,$$

che dunque risolve l'equazione di Vlasov. L'estesione di questa soluzione a tempi arbitrari segue dal fatto che il tempo T non dipende né dalla misura iniziale, dunque la costruzione della soluzione può essere iterata in ogni intervallo $[kT, (k+1)T]$, $K \in \mathbb{N}$.

Siamo ora μ_t^1 e μ_t^2 due soluzioni, di dato iniziale μ_0^1 e μ_0^2 , rispettivamente. La (19) diventa

$$\sup_{t \in [0, T]} D_1(\mu_t^1, \mu_t^2) \leq CD_1(\mu_0^1, \mu_0^2) + \gamma \sup_{t \in [0, T]} D_1(\mu_t^1, \mu_t^2)$$

da cui si ottiene

$$\sup_{t \in [0, T]} D_1(\mu_t^1, \mu_t^2) \leq \frac{C}{1 - \gamma} D_1(\mu_0^1, \mu_0^2).$$

Questa espressione garantisce l'unicità e la continuità nel dato iniziale. \square

5.2 Il limite di campo medio

Teorema 5.5 (Limite di campo medio).

Sia B_0 compatto, e sia μ_t la soluzione debole dell'equazione di Vlasov di dato iniziale $\mu_0 \in \mathcal{P}(B_0)$. Sia $\mathbf{z}_1^N(0) \in B_0$ dato iniziale di $\mathbf{z}_1^N(t)$ soluzione di

$$\frac{d}{dt}\mathbf{z}_i(t) = \frac{1}{N} \sum_{j=1}^N \mathbf{G}(\mathbf{z}_i(t), \mathbf{z}_j(t)). \quad (20)$$

con dato iniziale $\mathbf{z}_1^N(0)$. Sia

$$\pi(\mathbf{z}_1^N(t), d\mathbf{z}) = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{z}_i(t)}(d\mathbf{z})$$

la **misura empirica** associata. Se $\pi(\mathbf{z}_1^N(0)) \rightharpoonup \mu_0$ allora $\pi(\mathbf{z}_1^N(t)) \rightharpoonup \mu_t$.

Dimostrazione. La dimostrazione si basa sulla continuità nel dato iniziale delle soluzioni deboli dell'equazione di Vlasov, dopo aver notato che la misura empirica $\pi(\mathbf{z}_1^N(t), d\mathbf{z})$ è una soluzione debole dell'equazione di Vlasov. Il punto chiave è il seguente: considero il campo generato dalla misura empirica π

$$\mathcal{K}[\pi](t, \mathbf{z}) = \int \mathbf{G}(\mathbf{z}, \tilde{\mathbf{z}}) \pi(\mathbf{z}_1^N(t), d\tilde{\mathbf{z}}) = \frac{1}{N} \sum_j \mathbf{G}(\mathbf{z}, \mathbf{z}_j(t))$$

Sia \mathbf{Z}_t^π il flusso generato, cioè

$$\frac{d}{dt} \mathbf{Z}_t^\pi(\mathbf{z}) = \mathcal{K}[\pi](t, \mathbf{Z}_t^\pi(\mathbf{z}))$$

Si ha, evidentemente, che per ogni i

$$\mathbf{z}_i(t) = \mathbf{Z}_t^\pi(\mathbf{z}_i(0))$$

infatti queste due traiettorie soddisfano la stessa equazione differenziale con lo stesso dato iniziale. Ne segue che

$$\mathbf{Z}_t^\pi \# \pi(\mathbf{z}_1^N(0), d\mathbf{z}) = \pi(\mathbf{z}_1^N(t), d\mathbf{z}).$$

A questo punto, per continuità delle soluzioni deboli rispetto alla convergenza debole dei dati iniziali, si ottiene la tesi. □

Abbiamo provato un risultato molto forte di validità dell'equazione di campo medio: se prendo dati iniziali che approssimano (debolmente) una distribuzione iniziale di probabilità, questa approssimazione vale anche al tempo t .

Si tratta però di un punto di vista completamente diverso da quello da cui siamo partiti, in cui la domanda era sulla convergenza della marginale a una particella della soluzione dell'equazione di Liouville N particelle, a una soluzione dell'equazione di campo medio. Più in generale ci si può chiedere se, ipotizzando che $f_0^N = f_0^{\otimes N}$,

$$\lim_{N \rightarrow +\infty} f_j^N(t) = f_t^{\otimes N}$$

dove f_t risolve l'equazione di campo medio di dato iniziale f_0 .

La risposta a questa domanda è positiva e si può ottenere come sottoprodotto del teorema precedente, ma non ne parlo quest'anno.

Sunto

- L'equazione di Vlasov per una misura di probabilità consiste nel trovare una famiglia debolmente continua nel tempo che sia il push-forward del dato iniziale per un flusso determinato dal campo medio generato dalla misura stessa.
- Le ipotesi sul nucleo G garantiscono che data la famiglia di misure il campo ha la sufficiente regolarità per definire il flusso e il relativo push-forward.
- La prova di esistenza e unicità si ottiene per contrazione, e garantisce anche la continuità nel dato iniziale.
- La validità dell'equazione di Vlasov si ottiene elegantemente invocando la continuità nel dato iniziale, dopo aver mostrato che la misura empirica associata al sistema di particelle risolve l'equazione di Vlasov,

6 L'equazione di Boltzmann

In tempi recenti sono state introdotte equazioni di tipo Boltzmann come modelli matematici per lo studio di sistemi con interazioni binarie. Questi modelli hanno una natura intrinsecamente probabilistica, al contrario dell'equazione di Boltzmann che proviene da una dinamica deterministica. Ne parlo qui perché permettono al lettore di familiarizzarsi con la struttura dell'equazione e soprattutto con la sua interpretazione probabilistica.

6.1 Modelli discreti

Considero un sistema di N particelle (o agenti) in cui una particella può trovarsi in uno di un numero finito di stati, che indicherò con le cifre da 1 a ℓ . Per fissare le idee, si può pensare a uno stato come a una scatola, o a un valore di energia, o come a una "opinione" di un agente. Indico con n_i il numero di particelle nello stato i .

Descrivo due delle possibili dinamiche probabilistiche per questo sistema.

6.1.1 Il caso lineare e la master equation

Immagino che in un intervallo di tempo Δt piccolo, una particella nello stato i possa saltare nello stato j con probabilità $r(i, j)\Delta t$. La matrice $r(i, j)$ specifica i **rate di salto** ("rate" = "tasso").² Come cambia il numero di particelle nello stato i in questo intervallo di tempo?

$$\Delta n_i = \text{gain} - \text{loss}$$

dove "loss" è il termine di perdita, dovuto al fatto che nel tempo Δt le particelle nello stato i saltano in altri stati, mentre il termine di "gain" è dovuto al fatto che particelle in stati diversi da i saltano nello stato i . È relativamente semplice convincersi che

$$\begin{aligned} \text{loss} &= \sum_{j=1}^{\ell} r(i, j)n_i\Delta t \\ \text{gain} &= \sum_{j=1}^{\ell} r(j, i)n_j\Delta t \end{aligned}$$

Infatti se nel tempo Δt la probabilità di saltare di una particella da i a j è $r(i, j)\Delta t$, in questo tempo ne salteranno $r(i, j)n_i\Delta t$. Indico ora con $p_i(t)$ la frequenza delle particelle nello stato i :

$$p_i(t) = n_i(t)/N, \quad \text{cioè } n_i(t) = Np_i(t)$$

che posso anche interpretare come probabilità che una singola particella sia nello stato i . Posso scrivere,

$$\Delta n_i(t) \approx N \partial_t p_i \Delta t = N \Delta t \left(\sum_{j=1}^{\ell} r(j, i)p_j(t) - \sum_{j=1}^{\ell} r(i, j)p_i(t) \right)$$

²La versione precisa di questo sistema è la seguente: una particella nello stato i salta nello stato j ad tempo t_{ij} che è una variabile esponenziale di valore atteso $1/r(i, j)$. Quindi in un tempo $r(i, j)$ la particella fa circa un salto verso j . Si può pensare a questa variabile come a una sveglia che ricorda alla particella di saltare. Lo stato j in cui finisce la particella è determinato da quale sveglia suona prima.

Dividendo per $N\Delta t$ e passando al limite $\Delta t \rightarrow 0$, $N \rightarrow +\infty$ ottengo

$$\partial_t p_i(t) = \sum_{j=1}^{\ell} (r(j, i)p_j(t) - r(i, j)p_i(t)) \quad (21)$$

Questa equazione prende il nome di **master equation**.

La master equation è l'equazione di evoluzione della distribuzione di probabilità per una catena di Markov, a stati finiti, a tempo continuo con rate di salto r . Si osservi che la master equation vale per la distribuzione di probabilità di ogni singola particella, e per ottenerla non è realmente necessario usare N particelle con N divergente, perché le particelle sono indipendenti. Si osservi infine che si tratta di una equazione lineare.

6.1.2 Interazioni binare

Considero ora la stessa sistema, cioè N particelle in ℓ possibili stati, ma con una dinamica differente, perché considero il caso in cui una particella cambia stato solo se interagisce con un'altra particella. Assumo dunque che in tempo Δt due particelle negli stati i e j interagiscono con probabilità $\lambda(i, j)\Delta t$, per una opportuna matrice a termini non negativi λ . L'interazione porta le due particelle negli stati i', j' con probabilità assegnata $P(i', j'|i, j)$. Complessivamente, in un tempo Δt , una coppia di particelle negli stati (i, j) salta negli stati (i', j') con probabilità $B(i, j, i', j')\Delta t$, dove

$$B(i, j, i', j') = \lambda(i, j)P(i', j'|i, j).$$

Nel tempo Δt , il numero n_i di particelle nello stato i ha una variazione pari a

$$\Delta n_i = \Delta t \left(\sum_{i', j', j} B(i', j', i, j) n_{i'} n_{j'} - \sum_{i', j', j} B(i, j, i', j') n_i n_j \right)$$

Osservo che il termine di sinistra è di ordine N , infatti è circa $N \partial_t p_i(t) \Delta t$, mentre quello di destra è di ordine N^2 perché compare il numero di coppie di particelle. In queste condizioni non può esistere un'equazione limite per p_i . Per ottenerla è necessario che la probabilità di interazione tra due particelle sia di ordine $1/N$. In pratica si divide λ (e dunque B) per N . Si ottiene formalmente, dividendo per $N\Delta t$ e passando al limite,

$$\partial_t p_i(t) = \sum_{i', j', j} B(i', j', i, j) p_{i'} p_{j'} - \sum_{i', j', j} B(i, j, i', j') p_i p_j.$$

Questa è un'equazione di Boltzmann discreta, ed è quadratica in p . L'abbiamo formalmente ottenuta assumendo che ogni particella interagisce con un'altra con probabilità che tende a 0 come a $1/N$. Però poiché le particelle sono N , ogni particella interagisce con qualche altra particella con probabilità non nulla. Abbiamo descritto il sistema nel limite di **interazioni rare**. Si noti la differenza con i modelli di tipo Vlasov: in quel caso l'interazione non è per niente rara, ma l'effetto dell'interazione che va a 0 come $1/N$. Nel caso appena descritto l'effetto dell'interazione è di ordine 1, ma la frequenza dell'interazione va a 0 come $1/N$. Si osservi infine che per derivare questa equazione è indispensabile mandare N a infinito, e che si tratta di una equazione quadratica.

6.2 L'equazione di Boltzmann per le sfere dure

In questa sezione comincio a descrivere come si ricava euristicamente l'equazione di Boltzmann per i gas rarefatti, nel caso più classico, quello con l'interazione di coppia data dalla dinamica di sfere dure, cioè il caso di particelle sferiche che interagiscono solo tramite urti elastici. L'equazione di Boltzmann si può anche ottenere a partire da particelle puntiformi che interagiscono con potenziali centrali repulsivi a corto range (cioè a supporto compatto); l'equazione differirà da quella che introduco solo per l'espressione del cosiddetto nucleo di collisione B .

Considero particelle di diametro $r > 0$, e che a un certo istante di tempo vengano in contatto urtandosi. All'istante dell'urto le loro posizioni \mathbf{x}_1 e \mathbf{x}_2 devono soddisfare la seguente relazione $\mathbf{x}_2 = \mathbf{x}_1 + r\mathbf{n}$ dove \mathbf{n} è il versore di $\mathbf{x}_2 - \mathbf{x}_1$. Indico con \mathbf{v}_1 e \mathbf{v}_2 le velocità "entranti", cioè prima dell'urto. Affinché l'urto possa accadere, le particelle si devono avvicinare. Vediamo come questa condizione si traduca in una espressione simbolica. Supponiamo che l'istante di urto sia $t = 0$. Prima dell'urto, cioè per $t < 0$ e le posizioni delle particelle sono $\mathbf{x}_i + t\mathbf{v}_i$, $i = 1, 2$, e dunque

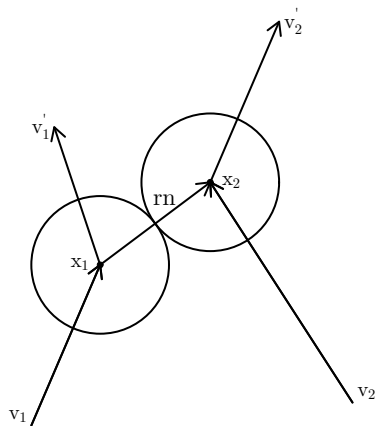
$$|\mathbf{x}_2 - \mathbf{x}_1 + t(\mathbf{v}_2 - \mathbf{v}_1)|^2 = r^2 + 2tr\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) + t^2|\mathbf{v}_1 - \mathbf{v}_2|^2.$$

Questa funzione di t deve essere decrescente fino a $t < 0$; la sua derivata in t è $r\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) + 2t|\mathbf{v}_1 - \mathbf{v}_2|^2$ che è negativa per $t < 0$ se e solo se

$$\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0.$$

(la stessa conclusione si poteva trarre da considerazioni geometriche).

Subito dopo l'urto, la posizione delle particelle è invariata, ma cambiano le velocità. Indico con $\mathbf{v}'_1, \mathbf{v}'_2$ le velocità "uscenti" e ne determino l'espressione in termini delle velocità entranti.



Anche se l'urto è una descrizione semplificata di un'interazione, devono valere i principi della fisica. In particolare, deve conservarsi la quantità di moto (perché non ci sono forze esterne), e deve conservarsi l'energia. Per imporre queste due condizioni è utile scrivere le velocità in termini della velocità del baricentro e della semi-differenza:

$$\bar{\mathbf{v}} = \frac{\mathbf{v}_1 + \mathbf{v}_2}{2}, \quad \mathbf{w} = \frac{\mathbf{v}_1 - \mathbf{v}_2}{2},$$

così che

$$\begin{aligned} \mathbf{v}_1 &= \bar{\mathbf{v}} + \mathbf{w} & \mathbf{v}'_1 &= \bar{\mathbf{v}} + \mathbf{w}' \\ \mathbf{v}_2 &= \bar{\mathbf{v}} - \mathbf{w} & \mathbf{v}'_2 &= \bar{\mathbf{v}} - \mathbf{w}'. \end{aligned} \quad \text{e}$$

Nell'ultima espressione ho usato che la velocità del baricentro si conserva, cioè $\frac{\mathbf{v}'_1 + \mathbf{v}'_2}{2} = \bar{\mathbf{v}}$ e ho indicato con \mathbf{w}' la semi-differenza delle velocità uscenti. In questo modo mi sono assicurato la conservazione della quantità di moto nell'urto, e ho ridotto il problema di trovare le velocità uscenti a quello di determinare \mathbf{w}' . Impongo l'uguaglianza dell'energia cinetica E prima e dopo l'urto:

$$E := \frac{1}{2}(|\mathbf{v}_1|^2 + |\mathbf{v}_2|^2) = |\bar{\mathbf{v}}|^2 + |\mathbf{w}|^2 = E' := \frac{1}{2}(|\mathbf{v}'_1|^2 + |\mathbf{v}'_2|^2) = |\bar{\mathbf{v}}|^2 + |\mathbf{w}'|^2.$$

La conservazione dell'energia si riduce alla condizione

$$|\mathbf{w}'|^2 = |\mathbf{w}|^2.$$

Questa uguaglianza dà diverse possibilità a seconda della dimensione: in \mathbb{R} ci sono solo due possibilità: $w' = w$ o $w' = -w$. Nel primo caso le velocità non cambiano nell'urto, nel secondo le particelle si scambiano le velocità (che è quello che accade colpendo in linea una biglia ferma con un'altra). Nel limite $r \rightarrow 0$ queste due possibilità danno lo stesso sistema: infatti nel secondo caso stiamo semplicemente scambiando il nome delle particelle dopo l'urto, senza nessuna variazione rispetto al moto libero. In pratica la dinamica di urti rigidi in una dimensione è in un certo senso indistinguibile dal flusso libero, e per questo non esiste l'equazione di Boltzmann unidimensionale.

In dimensione maggiore, la condizione di conservazione dell'energia non è evidentemente sufficiente a determinare il valore di \mathbf{w}' . Per andare avanti è indispensabile chiederci di quale sistema quello delle sfere dure sia un modello semplificato. È facile convincersi che stiamo considerando (al limite) un sistema meccanico con interazione di coppia data da un potenziale repulsivo a simmetria radiale che impedisca alle particelle di avvicinarsi più di r , e che non abbia effetto per distanze maggiori. Formalizzo: sia $U_\varepsilon(x)$ un potenziale su \mathbb{R} positivo, regolare, simmetrico rispetto a 0, decrescente tra 0 e $+\infty$, a supporto compatto e tale che

$$\lim_{\varepsilon \rightarrow 0} U_\varepsilon(x) = \begin{cases} +\infty & \text{se } |x| < r \\ 0 & \text{se } |x| > r \end{cases}$$

Mi aspetto che la legge dell'urto elastico si ottenga passando al limite il sistema meccanico di energia potenziale $U_\varepsilon(|\mathbf{x}_1 - \mathbf{x}_2|)$. L'informazione in più che mi da questa supposizione, è che, poiché la forza è parallela alla congiungente $\mathbf{x}_1 - \mathbf{x}_2$, la variazione di impulso avviene lungo la congiungente, e quindi nell'urto può solo cambiare la componente dell'impulso lungo \mathbf{n} . Vediamo le conseguenze di questo fatto.

Indico con $\mathbf{a} \otimes \mathbf{b}$ la matrice di componenti $a_i b_j$ che dunque agisce sui vettori in questo modo

$$\mathbf{a} \otimes \mathbf{b} \mathbf{v} = \mathbf{a} (\mathbf{b} \cdot \mathbf{v})$$

(la chiamerò un po' impropriamente "prodotto tensore"). Con questa definizione, $\mathbf{n} \otimes \mathbf{n}$ è l'operatore di proiezione lungo \mathbf{n} , mentre $\mathbb{I} - \mathbf{n} \otimes \mathbf{n}$ è il proiettore sul sottospazio ortogonale. Riscrivo la condizione di conservazione dell'energia imponendo che le componenti di \mathbf{w} ortogonali a \mathbf{n} non cambino, cioè che

$$(\mathbb{I} - \mathbf{n} \otimes \mathbf{n})\mathbf{w}' = (\mathbb{I} - \mathbf{n} \otimes \mathbf{n})\mathbf{w}.$$

Poiché $\mathbb{I} = (\mathbb{I} - \mathbf{n} \otimes \mathbf{n}) + \mathbf{n} \otimes \mathbf{n}$ ottengo

$$(\mathbf{n} \cdot \mathbf{w}')^2 = (\mathbf{n} \cdot \mathbf{w})^2$$

che mi da due scelte, una che porta a $\mathbf{w}' = \mathbf{w}$ (situazione non fisica in cui le particelle proseguono indisturbate), e l'altra che corrisponde a

$$\mathbf{n} \cdot \mathbf{w}' = -\mathbf{n} \cdot \mathbf{w};$$

in questo caso la componente della velocità relativa lungo \mathbf{n} viene invertita dall'urto. Ora posso scrivere le velocità uscenti in funzione delle entranti:

$$\begin{aligned} \mathbf{v}'_1 &= \bar{\mathbf{v}} + (\mathbb{1} - \mathbf{n} \otimes \mathbf{n}) \mathbf{w}' + \mathbf{n} \otimes \mathbf{n} \mathbf{w}' \\ &= \bar{\mathbf{v}} + (\mathbb{1} - \mathbf{n} \otimes \mathbf{n}) \mathbf{w} - \mathbf{n} \otimes \mathbf{n} \mathbf{w} \\ &= \bar{\mathbf{v}} + (\mathbb{1} - \mathbf{n} \otimes \mathbf{n}) \mathbf{w} + \mathbf{n} \otimes \mathbf{n} \mathbf{w} - 2\mathbf{n} \otimes \mathbf{n} \mathbf{w} \\ &= \mathbf{v}_1 - ((\mathbf{v}_1 - \mathbf{v}_2) \cdot \mathbf{n}) \mathbf{n} \end{aligned}$$

L'espressione della velocità \mathbf{v}'_2 si ottiene facilmente scambiando gli indici, ma viene usualmente scritta in modo leggermente differente:

$$\mathbf{v}'_2 = \mathbf{v}_2 - ((\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}) \mathbf{n} = \mathbf{v}_2 + ((\mathbf{v}_1 - \mathbf{v}_2) \cdot \mathbf{n}) \mathbf{n}$$

Osservazioni.

1. La condizione di conservazione dell'energia e di trasmissione normale dell'impulso sono invarianti per lo scambio tra $\mathbf{v}_1, \mathbf{v}_2$ e $\mathbf{v}'_1, \mathbf{v}'_2$, dunque è facile convincersi che la mappa $(\mathbf{v}_1, \mathbf{v}_2) \rightarrow (\mathbf{v}'_1, \mathbf{v}'_2)$ coincide con la sua inversa, cioè

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{v}'_1 - ((\mathbf{v}'_1 - \mathbf{v}'_2) \cdot \mathbf{n}) \mathbf{n} \\ \mathbf{v}_2 &= \mathbf{v}'_2 + ((\mathbf{v}'_1 - \mathbf{v}'_2) \cdot \mathbf{n}) \mathbf{n} \end{aligned}$$

Nell'espressione della legge di urto le velocità entranti e quelle uscenti sono indistinguibili. La differenza si nota solo ricordando che \mathbf{n} è il versore di $\mathbf{x}_2 - \mathbf{x}_1$: le velocità entranti soddisfano la condizione **pre-collisionale**

$$\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0$$

mentre le velocità uscenti soddisfano la condizione **post-collisionale**

$$\mathbf{n} \cdot (\mathbf{v}'_2 - \mathbf{v}'_1) > 0.$$

2. In quanto limite di un sistema di particelle conservativo, il modello di sfere dure è reversibile, il flusso associato conserva l'energia, che è solo cinetica, e conserva la misura nello spazio delle fasi.

3. La regola d'urto conserva la misura anche nel solo spazio delle velocità, infatti

$$d\mathbf{v}_1 d\mathbf{v}_2 = 2^d d\bar{\mathbf{v}} d\mathbf{w} = 2^d d\bar{\mathbf{v}} d\mathbf{w}' = d\mathbf{v}'_1 d\mathbf{v}'_2$$

Sunto

- La dinamica di sfere dure è un modello di interazioni di coppia ispirato a quelli con potenziali repulsivi a simmetria radiale e a corto range (cioè con supporto compatto).
- La regola d'urto discende dunque dalla conservazione dell'impulso, dalla conservazione dell'energia e dal fatto che la variazione di impulso è lungo al congiungete, come accade per tutti i sistemi con potenziali
- Detto \mathbf{n} il versore di $\mathbf{x}_2 - \mathbf{x}_1$, la condizione di velocità pre-collisionali è

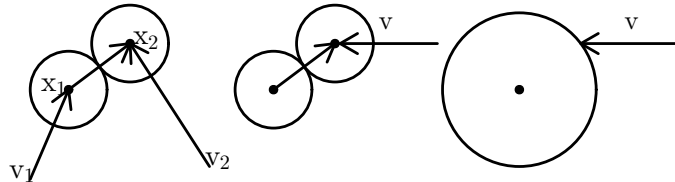
$$\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0$$

La condizione di velocità post-collisionali ha il segno opposto.

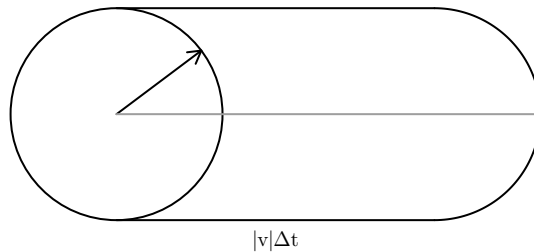
6.2.1 La sezione d'urto delle sfere dure

Per ricavare euristicamente l'equazione di Boltzmann serve ora quantificare la frequenza con cui in un tempo Δt una coppia di particelle di velocità $\mathbf{v}_1, \mathbf{v}_2$ urta e cambia velocità in $\mathbf{v}'_1, \mathbf{v}'_2$. Mettiamoci nel sistema di riferimento inerziale della particella 1, in cui una particella di velocità \mathbf{v}_2 appare avere velocità $\mathbf{v} = \mathbf{v}_2 - \mathbf{v}_1$.

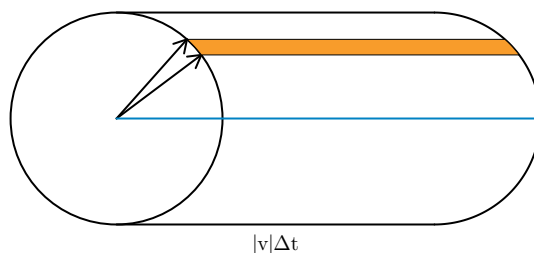
Se le due particelle urtano, il centro della particella 2 si trova sulla circonferenza di centro \mathbf{x}_1 e raggio r .



Le particelle che hanno centro nella regione cilindrica che ha come basi le due semisfere come in figura, sono tutte e sole quelle di velocità \mathbf{v} che nell'intervallo Δt urtano con la particella 1.



Il volume nello spazio delle fasi di questa regione è $\pi r^2 |\mathbf{v}| \Delta t$. Poiché le velocità di uscita sono perfettamente individuate da $\mathbf{v}_1, \mathbf{v}_2$ e dal parametro \mathbf{n} , è necessario determinare il numero di particelle che urtano con direzione \mathbf{n} . Questo conto va fatto considerando una regione di versori di ampiezza $d\mathbf{n}$, in modo da avere un volume di misura positiva. Va determinato il volume della regione arancione, che ha come base una calotta sferica di area $r^2 d\mathbf{n}$, che va proiettata sul piano ortogonale a \mathbf{v} .



Il volume cercato è

$$r^2 |\hat{\mathbf{v}} \cdot \mathbf{n}| d\mathbf{n} |\mathbf{v}| \Delta t = r^2 |\mathbf{v} \cdot \mathbf{n}| d\mathbf{n} \Delta t$$

Assumendo che ci sia **caos molecolare**, supponendo cioè che tutte le particelle abbiano la stessa distribuzione $f(t, \mathbf{x}, \mathbf{v})$, nel tempo Δt il numero di particelle in \mathbf{x} che si trovano ad avere velocità \mathbf{v}'_1 e \mathbf{v}'_2 come conseguenza dell'urto è

$$\Delta t N^2 r^2 |\mathbf{v} \cdot \mathbf{n}| f(t, \mathbf{x}_1, \mathbf{v}_1) f(t, \mathbf{x}_1 + r\mathbf{n}, \mathbf{v}_2)$$

(si noti che per Δt piccolo la seconda particella è in $\mathbf{x}_1 + r\mathbf{n}$ a meno di un ordine Δt). Valuto come cambia il numero di particelle con velocità \mathbf{v}_1 nel tempo a causa delle collisioni. Per ogni osservabile $\varphi(\mathbf{v})$ calcolo

$$\Delta_{\text{coll}} N \int d\mathbf{v}_1 \varphi(\mathbf{v}_1) f(t, \mathbf{x}_1, \mathbf{v}_1).$$

cioè il contributo alla variazione dovuto alle sole collisioni (l'altro termine è quello di flusso libero). La collisione trasforma \mathbf{v}_1 in \mathbf{v}'_1 , dunque

$$\begin{aligned} \Delta_{\text{coll}} N \int d\mathbf{v}_1 \varphi(\mathbf{v}_1) f(t, \mathbf{x}_1, \mathbf{v}_1) \\ = \Delta t N^2 r^2 \int d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} |(\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}| f(t, \mathbf{x}_1, \mathbf{v}_1) f(t, \mathbf{x}_1 + r\mathbf{n}, \mathbf{v}_2) (\varphi(\mathbf{v}'_1) - \varphi(\mathbf{v}_1)) \end{aligned}$$

Si comprende come sia necessario che $Nr^2 = \lambda$ sia di ordine 1. Dividendo per Δt e passando al limite $r \rightarrow 0$, $N \rightarrow +\infty$, $\Delta t \rightarrow 0$ si ottiene che la parte di derivata temporale di $\int f \varphi$ dovuta alle collisioni è

$$\begin{aligned} \lambda \int d\mathbf{v}_1 d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} |(\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}| f(t, \mathbf{x}_1, \mathbf{v}_1) f(t, \mathbf{x}_1, \mathbf{v}_2) (\varphi(\mathbf{v}'_1) - \varphi(\mathbf{v}_1)) \\ = \frac{1}{2} \lambda \int d\mathbf{v}_1 d\mathbf{v}_2 \int d\mathbf{n} |(\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}| f(t, \mathbf{x}_1, \mathbf{v}_1) f(t, \mathbf{x}_1, \mathbf{v}_2) (\varphi(\mathbf{v}'_1) - \varphi(\mathbf{v}_1)) \\ = \frac{1}{2} \lambda \int d\mathbf{v}_1 d\mathbf{v}_2 \int d\mathbf{n} |(\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}| (f(t, \mathbf{x}_1, \mathbf{v}'_1) f(t, \mathbf{x}_1, \mathbf{v}'_2) - f(t, \mathbf{x}_1, \mathbf{v}_1) f(t, \mathbf{x}_1, \mathbf{v}_2)) \varphi(\mathbf{v}_1) \end{aligned}$$

Nel primo passaggio ho sostituito l'integrale sulla semisfera all'integrale sulla sfera, con la metà dell'integrale sulla sfera. Nel secondo passaggio il termine in $\varphi(\mathbf{v}'_1)$ è stato riscritto scambiando il nome delle variabili $\mathbf{v}_1, \mathbf{v}_2$ con il nome delle variabili $\mathbf{v}'_1, \mathbf{v}'_2$. Posso farlo perché le relazioni tra le variabili entranti e uscenti sono le stesse. Poi ho usato che $d\mathbf{v}_1 d\mathbf{v}_2 = d\mathbf{v}'_1 d\mathbf{v}'_2$, e che $|(\mathbf{v}'_1 - \mathbf{v}'_2) \cdot \mathbf{n}| = |(\mathbf{v}_1 - \mathbf{v}_2) \cdot \mathbf{n}|$. In dettaglio:

$$\begin{aligned} \int d\mathbf{v}_1 d\mathbf{v}_2 \int d\mathbf{n} |(\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}| f(t, \mathbf{x}_1, \mathbf{v}_1) f(t, \mathbf{x}_1, \mathbf{v}_2) \varphi(\mathbf{v}'_1) \\ = \int d\mathbf{v}'_1 d\mathbf{v}'_2 \int d\mathbf{n} |(\mathbf{v}'_2 - \mathbf{v}'_1) \cdot \mathbf{n}| f(t, \mathbf{x}_1, \mathbf{v}'_1) f(t, \mathbf{x}_1, \mathbf{v}'_2) \varphi(\mathbf{v}_1) \\ = \int d\mathbf{v}_1 d\mathbf{v}_2 \int d\mathbf{n} |(\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}| f(t, \mathbf{x}_1, \mathbf{v}'_1) f(t, \mathbf{x}_1, \mathbf{v}'_2) \varphi(\mathbf{v}_1) \end{aligned}$$

A questo punto, per l'arbitrarietà di φ , e prendendo anche in considerazione la parte di flusso libero, si ottiene l'equazione di Boltzmann

$$\begin{aligned} \partial_t f + \mathbf{v} \cdot \partial_x f = \lambda Q(f, f) \\ Q(f, f) := \int_S d\mathbf{n} \int d\mathbf{v}_* |(\mathbf{v} - \mathbf{v}_*) \cdot \mathbf{n}| (f' f'_* - f f_*). \end{aligned} \quad (22)$$

dove ho inglobato in λ il fattore $1/2$, e indicato con $(\mathbf{v}, \mathbf{v}_*)$ le velocità entranti, e con $(\mathbf{v}', \mathbf{v}'_*)$ le velocità uscenti, e ho usato la notazione compatta

$$f = f(\mathbf{x}, \mathbf{v}), \quad f_* = f(\mathbf{x}, \mathbf{v}_*), \quad f' = f(\mathbf{x}, \mathbf{v}'), \quad f'_* = f(\mathbf{x}, \mathbf{v}'_*)$$

Il regime in cui $N \rightarrow +\infty$, $r \rightarrow 0$, con Nr^2 costante prende il nome di **limite di bassa densità** o **limite di Boltzmann-Grad**

Sunto

- Euristicamente, l'equazione di Boltzmann descrive l'evoluzione di un sistema di N particelle, in caso di urti rari.
- È necessario assumere che sia vero il caos molecolare.
- La rarità degli urti è garantita dall'assunzione Nr^2 costante, nel limite $N \rightarrow +\infty$, perché la sezione d'urto totale per le particelle di velocità $\mathbf{v}_1, \mathbf{v}_2$ è proporzionale a $r^2|\mathbf{v}_1 - \mathbf{v}_2|$, che è di ordine $1/N$.

7 La gerarchia BBGKY per il modello di sfere dure

Ricavare la gerarchia per la dinamica di N sfere dure presenta alcune difficoltà concettuali, ed è necessario prima comprendere bene come è fatto il flusso associato.

Ricordo che indico con $\mathbf{z}_i = (\mathbf{x}_i, \mathbf{v}_i)$, velocità e posizione dell' i -esima particella. Introduco la notazione semplificata $\mathbf{x} = \mathbf{x}_1^N$, $\mathbf{v} = \mathbf{v}_1^N$, $\mathbf{z} = \mathbf{z}_1^N$. Indicherò il flusso con

$$\mathbf{Z}_t(\mathbf{z}) = (\mathbf{X}_t(\mathbf{x}, \mathbf{v}), \mathbf{V}_t(\mathbf{x}, \mathbf{v}))$$

Fino a che non urtano due particelle, il flusso è libero, cioè

$$\mathbf{V}_t(\mathbf{x}, \mathbf{v}) = \mathbf{v}, \quad \mathbf{X}_t(\mathbf{x}, \mathbf{v}) = \mathbf{x} + t\mathbf{v}.$$

All'istante di contatto le velocità cambiano con le regole dell'urto; dopo il flusso è di nuovo libero, fino all'urto successivo. Si comprende facilmente che $\mathbf{X}_t(\mathbf{x}, \mathbf{v})$ è una funzione continua affine a tratti, mentre $\mathbf{V}_t(\mathbf{x}, \mathbf{v})$ è una funzione costante a tratti, con discontinuità di salto agli istanti di urto. Indicherò con $\mathbf{Z}_{t-}(\mathbf{z}) = (\mathbf{X}_t(\mathbf{z}), \mathbf{V}_{t-}(\mathbf{z}))$ e $\mathbf{Z}_{t+}(\mathbf{z}) = (\mathbf{X}_t(\mathbf{z}), \mathbf{V}_{t+}(\mathbf{z}))$ i differenti valori del flusso prima e dopo l'istante di urto t , in particolare quelli della velocità. Per definire il flusso c'è un'ulteriore difficoltà da considerare: bisogna escludere il caso di urti contemporanei di più particelle, in cui la dinamica non è ben definita e non si può definire per continuità (si pensi a due particelle di velocità opposte che collidono frontalmente, e una terza, che urta con entrambe da una direzione ortogonale: perturbando di poco il tempo di urto delle prime due, la terza collide o con una o con l'altra, andando in direzioni differenti). Inoltre, bisogna assicurarsi che non si verificano infiniti urti in tempo finito dopo il quale non è chiaro come proseguirebbe la dinamica. Entrambe queste patologie si possono curare mostrando che l'insieme dei dati iniziali per cui la dinamica comporta urti multipli o infiniti urti in tempo finito è di misura nulla, grazie al fatto che il flusso conserva la misura.

Scopo di questa sezione è scrivere la gerarchia per la densità di probabilità $f_t^N(\mathbf{z}_1^N)$. Si noti che nella regione $\Omega_r = \{(\mathbf{x}_1^N, \mathbf{v}_1^N) | \forall i \neq j, |\mathbf{x}_i - \mathbf{x}_j| > r\}$ l'equazione di Liouville coincide con l'equazione del trasporto:

$$\partial_t f_t^N + \sum_{i=1}^N \mathbf{v}_i \cdot \partial_{\mathbf{x}_i} f_t^N = 0$$

Però a questa equazione vanno associate le condizioni al contorno su $\partial\Omega_r$, che si ottengono dalle proprietà del flusso. Per marginalizzarla va integrata fuori dai "buchi" $|\mathbf{x}_i - \mathbf{x}_j| < r$, tenendo conto delle condizioni al bordo. Non seguo qui questa lunga e complessa strada (trovate i dettagli negli appunti del 24/25). Preferisco scrivere una versione distribuzionale dell'equazione di Liouville, senza condizioni al contorno, e per mostrare come si fa considero un esempio semplice di dinamica con urto.

7.1 L'equazione di Liouville per una dinamica con urto

Considero una particella di massa 1 sulla semiretta $x > 0$, che si muove libera e meno che non urti con una parete in $x = 0$; in tal caso viene riflessa invertendo la velocità. Il flusso è dunque così definito: se $v > 0$

$$X_t(x, v) = x + vt, \quad V_t(x, v) = v$$

se $v < 0$, detto $\tau = -x/v$ l'istante di tempo in cui la particella urta la parete,

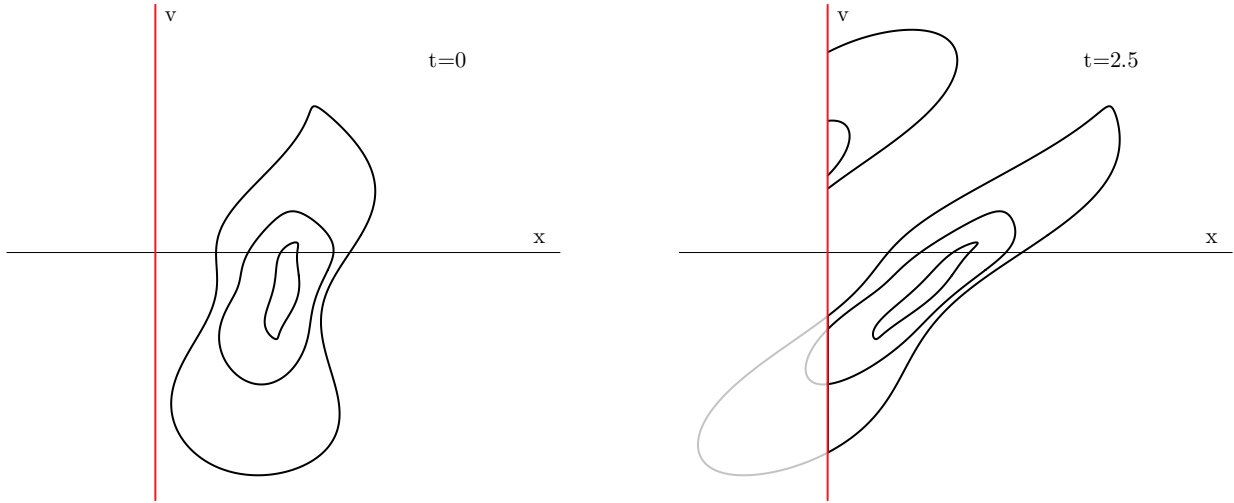
$$\begin{cases} X_t(x, v) = x + vt, & V_t(x, v) = v & \text{per } t < \tau \\ X_t(x, v) = -(t - \tau)v, & V_t(x, v) = -v & \text{per } t > \tau \end{cases}$$

Considero la soluzione dell'equazione di Liouville di dato iniziale $f_0(x, v)$, con supporto in $x > 0$. Non è difficile comprendere che la soluzione si ottiene nel seguente modo. Prima si costruisce la soluzione come se non ci fosse la parete, cioè si definisce $f_t^L(x, v) = f_0(x - vt, v)$. Poi si considera la simmetrica di f_t^L rispetto all'origine: $f_t^R(x, v) = f_t^L(-x, -v)$. La soluzione è data da

$$f_t(x, v) = (f_t^L(x, v) + f_t^R(x, v))\mathcal{X}\{x > 0\}$$

Si noti che f deve avere la seguente proprietà di continuità al bordo:

$$f_t(0, v) = f_t(0, -v).$$



Scrivo l'equazione soddisfatta da f_t nella forma seguente:

$$\partial_t f_t(x, v) + v \partial_x f_t(x, v) = \mathcal{C}$$

dove \mathcal{C} è una distribuzione che tiene conto del possibile urto, con supporto in $x = 0$, perché per $x > 0$ il flusso è libero. Per farlo, integro la soluzione contro una funzione test regolare $\alpha(x, v)$, ottenendo, per definizione di equazione di Liouville,

$$\int_{\mathbb{R}} dx \int_{\mathbb{R}} dv \alpha(x, v) f_t(x, v) = \int_{\mathbb{R}} dx \int_{\mathbb{R}} dv \alpha(X_t(x, v), V_t(x, v)) f_0(x, v)$$

(ricordo che il supporto di f_t è in $x > 0$). Decompongo l'integrale a destra a seconda che l'urto avvenga o meno:

$$\begin{aligned} & \int_{\mathbb{R}} dx \int_{\mathbb{R}} dv \alpha(X_t(x, v), V_t(x, v)) f_0(x, v) = (1) + (2) + (3) \\ (1) &= \int_{\mathbb{R}} dx \int_0^{+\infty} dv \alpha(X_t(x, v), V_t(x, v)) f_0(x, v) \\ (2) &= \int_{-\infty}^0 dv \int_{-vt}^{+\infty} dx \alpha(X_t(x, v), V_t(x, v)) f_0(x, v) \\ (3) &= \int_{-\infty}^0 dv \int_0^{-vt} dx \alpha(X_t(x, v), V_t(x, v)) f_0(x, v) \end{aligned}$$

Solo nel terzo integrale il flusso è calcolato su traiettorie che hanno subito l'urto. La derivata temporale si ottiene considerando due contributi, quello dovuto alla derivata di $\alpha(X_t(x, v), V_t(x, v))$ che vale $V_t(x, v) \partial_x \alpha(X_t(x, v), V_t(x, v))$ in tutte e tre le regioni, e quello dovuto alla derivata degli estremi di integrazione. Il contributo della derivata lungo il flusso di α è, complessivamente,

$$\int_{\mathbb{R}} dx \int_{\mathbb{R}} dv V_t(x, v) \partial_x \alpha(X_t(x, v), V_t(x, v)) f_0(x, v) = \int_{\mathbb{R}} dx \int_{\mathbb{R}} dv v \partial_x \alpha(x, v) f_t(x, v)$$

Distribuzionalmente, questo termine è $-v \partial_x f_t$ integrato contro α .

Determiniamo il contributo che si ottiene derivando in t gli estremi di integrazione di x . Nel caso dell'integrale (2), derivo rispetto all'estremo inferiore $-vt$ ottenendo un coefficiente $+v$:

$$\int_{-\infty}^0 dv v \alpha(X_t(-vt, v), V_{t-}(-vt, v)) f_0(-vt, v)$$

Si noti che se $v < 0$ si ha $X_t(-vt, v) = 0$, dunque sto valutato il contributo della collisione. Poiché l'integrando in (2) = $\int_{-vt}^{+\infty}$ dipende solo dal flusso prima dell'urto, nella derivata temporale compare la velocità prima dell'urto: $V_{t-}(-vt, v) = v$.

Ricordando che sto integrando nella regione $v < 0$, e che $f_0(-vt, v) = f_t(0, v)$, posso riscrivere il termine come

$$- \int_{-\infty}^0 dv |v| \alpha(0, v) f_0(-vt, v) = - \int_{\mathbb{R}} \delta_0(dx) \int_{\mathbb{R}} dv \mathcal{X}\{v < 0\} |v| \alpha(x, v) f_t(x, v).$$

Nel caso dell'integrale (3), derivo rispetto all'estremo superiore $-vt$ ottenendo un coefficiente $-v$:

$$- \int_{-\infty}^0 dv v \alpha(X_{t+}(-vt, v), V_{t+}(-vt, v)) f_0(-vt, v)$$

Questa volta però il flusso è calcolato dopo l'urto: $\alpha(X_{t+}(-vt, v), V_{t+}(-vt, v)) = \alpha(0, -v)$. Dunque si ha

$$\begin{aligned} - \int_{-\infty}^0 dv v \alpha(0, -v) f_t(0, v) &= \int_{\mathbb{R}} \delta_0(dx) \int_{\mathbb{R}} dv \mathcal{X}\{v < 0\} |v| \alpha(x, -v) f_t(x, v) \\ &= \int_{\mathbb{R}} \delta_0(dx) \int_{\mathbb{R}} dv \mathcal{X}\{v > 0\} |v| \alpha(x, v) f_t(x, -v) \end{aligned}$$

Mettendo insieme tutti i contributi si ottiene l'identità distribuzionale

$$\partial_t f_t(x, v) + v \partial_x f_t(x, v) = \delta_0(dx) |v| (\mathcal{X}\{v > 0\} f_t(x, -v) - \mathcal{X}\{v < 0\} f_t(x, v))$$

Si noti che, anche se $f_t(0, v) = f_t(0, -v)$ per continuità, i due termini a destra non si annullano perché hanno supporti disgiunti in v : si tratta per $v < 0$ del termine di loss (le particelle con velocità negativa spariscono in $x = 0$) e per $v > 0$ del termine di gain (le particelle riappaiono in 0 con velocità positiva). Il rate con cui avviene questa trasformazione è dato dalla sezione d'urto $|v|$.

7.2 L'equazione di Liouville per due sfere dure

Considero ora il caso di due sole sfere dure. Assumo che $f_0^2(\mathbf{z}_1, \mathbf{z}_2)$ abbia supporto, compatto, nella regione $|\mathbf{x}_1 - \mathbf{x}_2| > r$, e sia continua. Chiediamoci cosa accade a f_t^2 quando il suo supporto raggiunge la regione di urto. Consideriamo un dato iniziale che urta al tempo t nella configurazione $\mathbf{x}_1, \mathbf{x}_2$. Poiché f^2 deve essere costante lungo il flusso, deve accadere che

$$f_t^2(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2) = f_0(\mathbf{z}) = f_t^2(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}'_1, \mathbf{v}'_2)$$

Sto imponendo la condizione di riflessione al bordo della regione $|\mathbf{x}_1 - \mathbf{x}_2| = r$.

Voglio scrivere la derivata distribuzionale di f_t^2 . Per farlo devo integrare contro una funzione test e distinguere le traiettorie che hanno urtato da quelle libere. A questo scopo introduco la funzione $\Delta_t = \Delta_t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2)$, data da $\Delta_t = \mathbf{x}_2 - \mathbf{x}_1 + (\mathbf{v}_2 - \mathbf{v}_1)t$, il cui modulo misura la distanza tra le due traiettorie libere $\mathbf{x}_1 + \mathbf{v}_1 t$ e $\mathbf{x}_2 + \mathbf{v}_2 t$. Se $\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) > 0$, $|\Delta_t|$ è una funzione crescente, se $\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0$ è una una funzione inizialmente decrescente, poi ha un minimo, e poi torna a crescere (se ne calcoli la derivata temporale). Impongo che Δ_t rimanga costante dopo che $|\Delta_t|$ ha raggiunto il minimo. In questo modo, sul supporto di f_0^2 , esiste al più un solo tempo $t > 0$ tale che $|\Delta_t| = r$.

Sia $\alpha(\mathbf{z})$ una un funzione test.

$$\begin{aligned} \int \alpha(\mathbf{z}) f_t^2(\mathbf{z}) d\mathbf{z} &= \int \alpha(\mathbf{Z}_t(\mathbf{z})) f_0^2(\mathbf{z}) d\mathbf{z} \\ &= \int \alpha(\mathbf{Z}_t(\mathbf{z})) f_0^2(\mathbf{z}) \mathcal{X}\{|\Delta_t| > r\} + \int \alpha(\mathbf{Z}_t(\mathbf{z})) f_0^2(\mathbf{z}) \mathcal{X}\{|\Delta_t| < r\} \end{aligned}$$

La derivata temporale ha due contributi. Il primo si ottiene derivando la funzione test lungo il flusso:

$$\frac{d}{dt} \alpha(\mathbf{Z}_t(\mathbf{z})) = \mathbf{V}_t(\mathbf{x}, \mathbf{v}) \cdot \partial_x \alpha(\mathbf{Z}_t(\mathbf{z}))$$

Il termine che si ottiene è la derivata distribuzionale $-\mathbf{v} \cdot \partial_x f_t^2(\mathbf{z})$ testata contro α . L'altro contributo viene dalla derivata della funzione indicatrice. Tenendo conto del fatto che $\mathcal{X}\{|\Delta_t| < r\} = 1 - \mathcal{X}\{|\Delta_t| > r\}$ si ha

$$\int d\mathbf{z} f_0(\mathbf{z}) \frac{d}{dt} \mathcal{X}\{|\Delta_t| > r\} (\alpha(\mathbf{Z}_{t-}(\mathbf{z})) - \alpha(\mathbf{Z}_{t+}(\mathbf{z})))$$

dove con $\mathbf{Z}_{t\pm}$ indico il valore del flusso al tempo t dopo e prima dell'urto.

Vale

$$\frac{d}{dt} \mathcal{X}\{|\Delta_t| > r\} = \delta(|\Delta_t| - r) \frac{d}{dt} |\Delta_t|$$

Il supporto di questa derivata è nell'insieme in cui $|\Delta_t| = r$, cioè nei dati iniziali per cui si ha l'urto al tempo t . Inoltre

$$\frac{d}{dt} |\Delta_t| = \frac{1}{|\Delta_t|} \Delta_t \cdot (\mathbf{v}_2 - \mathbf{v}_1) = \mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)$$

dove \mathbf{n} è il versore di $\mathbf{x}_2 + \mathbf{v}_2 t - (\mathbf{x}_1 + \mathbf{v}_1 t)$ all'istante di urto. In definitiva

$$\frac{d}{dt} \mathcal{X}\{|\Delta_t| < r\} = \delta(|\Delta_t| - r) \mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1).$$

Si noti che fino al tempo di urto, compreso, $\mathbf{v}_1 = \mathbf{V}_t^1(\mathbf{z})$, $\mathbf{v}_2 = \mathbf{V}_t^2(\mathbf{z})$ e $\Delta_t = \mathbf{X}_t^2(\mathbf{z}) - \mathbf{X}_t^1(\mathbf{z})$. Il primo termine dell'integrale è dunque

$$\int d\mathbf{z} f_0(\mathbf{z}) \alpha(\mathbf{Z}_{t-}) \delta(|\mathbf{X}_t^2(\mathbf{z}) - \mathbf{X}_t^1(\mathbf{z})| - r) [\mathbf{n} \cdot (\mathbf{V}_{t-}^2(\mathbf{z}) - \mathbf{V}_{t-}^1(\mathbf{z}))]^-$$

dove con $[x]^-$ si intende la parte negativa di x , cioè x se x è negativo, 0 altrimenti. Invertendo il push-forward, l'integrale diventa

$$\int d\mathbf{z} \alpha(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2) f_t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2) \delta(|\mathbf{x}_2 - \mathbf{x}_1| - r) [\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)]^-$$

dove $r\mathbf{n} = \mathbf{x}_2 - \mathbf{x}_1$. Esplicito parzialmente la δ , notando che

$$\delta(|\mathbf{x}_1 - \mathbf{x}_2| - r) d\mathbf{x}_2 = r^2 d\mathbf{n} \delta_{\mathbf{x}_1+r\mathbf{n}}(d\mathbf{x}_2)$$

Posso quindi riscrivere il termine più esplicitamente come

$$-r^2 \int d\mathbf{x}_1 d\mathbf{v}_1 d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} \delta_{\mathbf{x}_1+r\mathbf{n}}(d\mathbf{x}_2) |\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)| d\mathbf{n} f_t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2) \alpha(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2)$$

L'altro termine è

$$\begin{aligned} & - \int d\mathbf{z} f_0(\mathbf{z}) \alpha(\mathbf{Z}_t^+) \delta(|\mathbf{x}_2(t) - \mathbf{x}_1(t)| - r) [\mathbf{n} \cdot (\mathbf{V}_t^2(\mathbf{z}) - \mathbf{V}_t^1(\mathbf{z}))]^- \\ & = - \int d\mathbf{z} f_0(\mathbf{z}) \alpha(\mathbf{X}_t^1(\mathbf{z}), \mathbf{X}_t^2(\mathbf{z}), (\mathbf{V}_t^1(\mathbf{z}))', (\mathbf{V}_t^2(\mathbf{z}))') \delta(|\mathbf{X}_t^2(\mathbf{z}) - \mathbf{X}_t^1(\mathbf{z})| - r) \\ & \quad [\mathbf{n} \cdot (\mathbf{V}_t^2(\mathbf{z}) - \mathbf{V}_t^1(\mathbf{z}))]^- . \end{aligned}$$

Invertendo il push-forward, ottengo che l'osservabile α è calcolato nelle velocità uscenti:

$$- \int d\mathbf{z} f_t(\mathbf{z}) \alpha(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}'_1, \mathbf{v}'_2) \delta(|\mathbf{x}_2 - \mathbf{x}_1| - r) [\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)]^- .$$

Procedendo esattamente come sopra ottengo

$$r^2 \int d\mathbf{x}_1 d\mathbf{v}_1 d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} \delta_{\mathbf{x}_1+r\mathbf{n}}(d\mathbf{x}_2) |\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)| f_t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2) \alpha(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}'_1, \mathbf{v}'_2)$$

Ora è necessario fare dei cambi di variabile per riportare α nelle variabili originali.

Lemma 7.1 (Cambi di variabile). *Sia F una funzione delle velocità entranti, di quelle uscenti, e di \mathbf{n} . Si ha*

$$\int d\mathbf{v}_1 d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} F(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}'_1, \mathbf{v}'_2, \mathbf{n}) = \int d\mathbf{v}_1 d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} F(\mathbf{v}'_1, \mathbf{v}'_2, \mathbf{v}_1, \mathbf{v}_2, -\mathbf{n})$$

Dimostrazione. Sia I l'integrale a sinistra. Ricordo che $\mathbf{v}'_1, \mathbf{v}'_2$ sono determinate dalle regole di urto in funzione di $\mathbf{v}_1, \mathbf{v}_2, \mathbf{n}$. Per prima cosa cambio nome alle variabili di integrazione, $\mathbf{v}_1 \rightarrow \mathbf{v}'_1, \mathbf{v}_2 \rightarrow \mathbf{v}'_2$; con questo cambio $\mathbf{v}'_1, \mathbf{v}'_2$ diventano $\mathbf{v}_1, \mathbf{v}_2$ perché la regola d'urto coincide con la sua inversa. Si ottiene così

$$I = \int d\mathbf{v}'_1 d\mathbf{v}'_2 \int_{\mathbf{n} \cdot (\mathbf{v}'_2 - \mathbf{v}'_1) < 0} d\mathbf{n} F(\mathbf{v}'_1, \mathbf{v}'_2, \mathbf{v}_1, \mathbf{v}_2, \mathbf{n})$$

Ora cambio variabili di integrazione $d\mathbf{v}'_1 d\mathbf{v}'_2 = d\mathbf{v}_1 d\mathbf{v}_2$, ricordando che lo jacobiano è 1:

$$I = \int d\mathbf{v}_1 d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}'_2 - \mathbf{v}'_1) < 0} d\mathbf{n} F(\mathbf{v}'_1, \mathbf{v}'_2, \mathbf{v}_1, \mathbf{v}_2, \mathbf{n})$$

Infine, ricordando che $\mathbf{n} \cdot (\mathbf{v}'_2 - \mathbf{v}'_1) = -\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)$ e cambiando la variabile di integrazione $\mathbf{n} \rightarrow -\mathbf{n}$ si ottiene la tesi (le regole di urto sono invarianti nel cambio $\mathbf{n} \rightarrow -\mathbf{n}$, perché quadratiche in \mathbf{n}). \square

Usando questo lemma, il termine di gain si scrive

$$r^2 \int d\mathbf{x}_1 d\mathbf{v}_1 d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} \delta_{\mathbf{x}_1 - r\mathbf{n}}(d\mathbf{x}_2) |\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)| f_t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}'_1, \mathbf{v}'_2) \alpha(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2)$$

L'equazione di Liouville nel caso di due particelle è

$$\begin{aligned} \partial_t f_t(\mathbf{z}) + \sum_{i=1}^2 \mathbf{v}_i \cdot \nabla f_t(\mathbf{z}) &= \mathcal{T}_{12} f_t := \\ r^2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} |\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)| & (\delta_{\mathbf{x}_1 - r\mathbf{n}}(d\mathbf{x}_2) f_t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}'_1, \mathbf{v}'_2) - \delta_{\mathbf{x}_1 + r\mathbf{n}}(d\mathbf{x}_2) f_t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2)) \end{aligned}$$

Anche in questo caso il secondo membro è una distribuzione, che ha supporto solo sul bordo. Abbiamo ottenuto un'espressione relativamente comprensibile: nella regione al bordo data da $|\mathbf{x}_1 - \mathbf{x}_2| = r$, dove avviene l'urto, f varia perché compaiono i valori calcolati nelle velocità uscenti e si azzerano i termini nelle velocità entranti.

È ora semplice riscrivere l'equazione di Liouville nel caso generale:

$$\begin{aligned} \partial_t f_t^N(\mathbf{z}) + \sum_{i=1}^N \mathbf{v}_i \cdot \partial_{x_i} f_t(\mathbf{z}) &= \sum_{i < j} \mathcal{T}_{ij} f^N \\ \mathcal{T}_{ij} f^N(\mathbf{z}) &:= r^2 \int_{\mathbf{n} \cdot (\mathbf{v}_j - \mathbf{v}_i) < 0} d\mathbf{n} |\mathbf{n} \cdot (\mathbf{v}_j - \mathbf{v}_i)| (\delta_{\mathbf{x}_i - r\mathbf{n}}(d\mathbf{x}_j) f^N(\tilde{\mathbf{z}}_{ij}) - \delta_{\mathbf{x}_i + r\mathbf{n}}(d\mathbf{x}_j) f^N(\mathbf{z})) \end{aligned}$$

dove $\tilde{\mathbf{z}}_{ij}$ si ottiene da \mathbf{z} considerando le velocità uscenti per le particelle i e j .

A questo punto possiamo marginalizzare questa equazione per ottenere la gerarchia. Ci servirà il seguente semplice lemma, il cui contenuto non deve sorprendere, perché la probabilità totale si deve conservare, dunque l'integrale in tutte le variabili del termine di collisione deve essere nullo.

Lemma 7.2.

$$\int d\mathbf{x}_i d\mathbf{x}_j d\mathbf{v}_i d\mathbf{v}_j \mathcal{T}_{ij} f^N = 0$$

Dimostrazione. Consideriamo senza mancare di generalità il caso \mathcal{T}_{12} : l'integrale diventa

$$r^2 \int d\mathbf{x}_1 d\mathbf{v}_1 d\mathbf{v}_2 \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} |\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)| (f(\mathbf{x}_1, \mathbf{x}_1 - r\mathbf{n}, \mathbf{v}'_1, \mathbf{v}'_2) - f^2(\mathbf{x}_1, \mathbf{x}_1 + r\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2))$$

Operando sul termine di gain come nel lemma 7.1 si ottiene che esso cancella esattamente il termine di loss. \square

Considero ora l'equazione di Liouville, e integro in $d\mathbf{z}_2^N$. Il termine di trasporto dà $\mathbf{v}_1 \cdot \partial_{x_1} f_1^N$, l'integrale dei termini a destra dà zero nel caso $2 \leq i < j$. Sopravvivono gli $N - 1$ termini dati da $1 = i < k$. Si vede facilmente che sono tutti uguali, dunque si ottiene la seguente equazione per la marginale a una particella

$$\begin{aligned} \partial_t f_1^N + \mathbf{v}_1 \cdot \partial_{x_1} f_1^N &= (N - 1)r^2 \int d\mathbf{v}_2 \\ \int_{\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1) < 0} d\mathbf{n} |\mathbf{n} \cdot (\mathbf{v}_2 - \mathbf{v}_1)| & (f_2^N(\mathbf{x}_1, \mathbf{x}_2 - r\mathbf{n}, \mathbf{v}'_1, \mathbf{v}'_2) - f_2^N(\mathbf{x}_1, \mathbf{x}_2 + r\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2)) \end{aligned} \quad (23)$$

Non è difficile scrivere le successive equazioni della gerarchia.

$$\begin{aligned} \partial_t f_j^N(t, \mathbf{x}_1^j, \mathbf{v}_1^j) + \sum_{i=1}^j \mathbf{v}_i \cdot \partial_{x_i} f_j^N(t, \mathbf{x}_1^j, \mathbf{v}_1^j) &= \sum_{i < h \leq j} \mathcal{T}_{ih} f_j^N + (N - j)r^2 \sum_{i=1}^j \int d\mathbf{v}_{j+1} \\ \int_{\mathbf{n} \cdot (\mathbf{v}_{j+1} - \mathbf{v}_i) < 0} d\mathbf{n} |\mathbf{n} \cdot (\mathbf{v}_{j+1} - \mathbf{v}_i)| & (f_{j+1}^N(t, \mathbf{x}_1^j, \mathbf{x}_i - r\mathbf{n}, \mathbf{v}_1, \dots, \mathbf{v}'_i \dots \mathbf{v}'_{j+1}) - f_{j+1}^N(\mathbf{x}_1^j, \mathbf{x}_i + r\mathbf{n}, \mathbf{v}_1^{j+1})) \end{aligned} \quad (24)$$

Tornando all'equazione per la marginale a una particella, nel limite $N \rightarrow +\infty$, $r \rightarrow 0$, $Nr^2 \rightarrow 2\lambda$, ipotizzando che $f_2^N \rightarrow f_2$ e $f_1^N \rightarrow f_1$ si ottiene

$$\partial_t f_1 + \mathbf{v}_1 \cdot \partial_{x_1} f_1 = 2\lambda \int d\mathbf{v}_2 \int_{S^-} d\mathbf{n} |(\mathbf{v}_1 - \mathbf{v}_2) \cdot \mathbf{n}| (f_2(\mathbf{x}_1, \mathbf{x}_1, \mathbf{v}'_1, \mathbf{v}'_2) - f_2(\mathbf{x}_1, \mathbf{x}_1, \mathbf{v}_1, \mathbf{v}_2))$$

e equazioni analoghe per le f_j con $j \geq 2$, limiti dell'equazione per la marginale a j particelle. La corrispondente gerarchia di equazioni ammette soluzioni fattorizzate se $f_1 = f$ risolve l'equazione di Boltzmann, che si ottiene supponendo che

$$f_2(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2) = f(\mathbf{x}_1, \mathbf{v}_2) f(\mathbf{x}_2, \mathbf{v}_2),$$

ed è

$$\partial_t f + \mathbf{v} \cdot \partial_x f = 2\lambda \int d\mathbf{v}_* \int_{S^-} d\mathbf{n} |(\mathbf{v} - \mathbf{v}_*) \cdot \mathbf{n}| (f' f'_* - f f_*) = \lambda \int_S d\mathbf{n} \int d\mathbf{v}_* |(\mathbf{v} - \mathbf{v}_*) \cdot \mathbf{n}| (f' f'_* - f f_*)$$

dove

$$f = f(\mathbf{x}, \mathbf{v}), \quad f_* = f(\mathbf{x}, \mathbf{v}), \quad f' = f(\mathbf{x}, \mathbf{v}'), \quad f'_* = f(\mathbf{x}, \mathbf{v}_*)$$

Un'ultima osservazione: nel termine di collisione della prima equazione della gerarchia, sia nel loss che nel gain f_2 è calcolata nelle velocità pre-collisionali. Dunque per ottenere Boltzmann assumiamo che f_2 sia fattorizzata nelle variabili precollisionali. Se f_2 nel termine di gain si fattorizzasse nelle variabili uscenti i due termini si sarebbero compensati dando 0.

La prova rigorosa della validità dell'equazione di Boltzmann è concettualmente sottile, ed è stata ottenuta da Lanford nel 1973, per tempi dell'ordine del tempo di collisione. La difficoltà non dovrebbe sorprendere, perché l'equazione di Boltzmann è un'equazione irreversibile, come vedremo nel prossimo capitolo.

Scaling e limite di bassa densità

C'è un modo più formale per giustificare il nome di limite di bassa densità. Fissiamo r , scegliamo un parametro ε , e definiamo le funzioni $f^{N,\varepsilon}$ nel seguente modo:

$$f_j^N(t, \mathbf{x}_1^j, \mathbf{v}_1^j) = \varepsilon^{3j} f_j^{N,\varepsilon}(\varepsilon t, \varepsilon \mathbf{x}_1^j, \mathbf{v}_1^j).$$

Si noti che le velocità non scalano perché spazio e tempo sono scalate nello stesso modo. Le nuove funzioni sono normalizzate a 1, cioè sono ancora densità di probabilità. Ipotizziamo che $f_j^{N,\varepsilon}$ vari su scala 1 in t e \mathbf{x} ; ciò equivale ad assumere che f_j^N vari spazialmente e temporalmente su scala $1/\varepsilon$, molto grande rispetto alla dimensione delle sfere. Nell'equazione per la marginale a una particella di $F^{N,\varepsilon}$ si ottiene che il termine di collisione ha un coefficiente dato da

$$\frac{1}{\varepsilon}\varepsilon^{-3}\varepsilon^6(N-1)r^2 = (N-1)(\varepsilon r)^2,$$

dove ε^{-3} viene dalla normalizzazione di $f_1^{N,\varepsilon}$, ε^6 da quella di $f_2^{N,\varepsilon}$, mentre $1/\varepsilon$ viene dal rapporto tra le derivate in t e \mathbf{x} microscopiche e le corrispondenti derivate nelle variabili macroscopiche.

Il limite di Boltzmann-Grad si ottiene mandando N a $+\infty$ e $\varepsilon \rightarrow 0$, in modo che $N\varepsilon^2r^2 = 2\lambda$ fissato non nullo. Fissando questo valore λ , si ha che il numero di particelle va come $N \approx \lambda/(r^2\varepsilon^2)$, dunque in cubo micoscopico di lato $1/\varepsilon$ e volume $1/\varepsilon^3$ ci saranno ordine $V^{2/3}$ particelle. Dunque si tratta di un limite di **bassa densità**.