UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea in Fisica

Tesi di Laurea

Introduzione alla geometria del trasporto ottimale con qualche applicazione alla fluidodinamica

Relatore

Prof. Davide Vittone

Laureando

Pietro Locatelli

Anno Accademico 2018/2019

Introduzione

Il problema del trasporto ottimale è volto a studiare come trasportare una massa nel "modo migliore possibile" da un luogo ad un altro. Si può formulare in questo modo: lo spazio di partenza e di arrivo sono genericamente due spazi topologici X, Y. Per quantificare la massa spostata viene naturale utilizzare una misura di probabilità μ sullo spazio X ed una ν sullo spazio Y. E' chiaro che se si sposta una massa da $A \subset X$ a $B \subset Y$ allora affinchè la quantità spostata rimanga la stessa si impone $\mu(A) = \nu(B)$. Si possono poi normalizzare ad 1 le misure, così diventano misure di probabilità. Tuttavia, la domanda da cui nasce il problema è: quanto "costa" lo spostamento del materiale da X a Y? Si definisce dunque una funzione di costo $c(x, y)$ che deve essere almeno misurabile e affinchè abbia una corrispondenza con il modello descritto sia non negativa. A priori non si esclude nemmeno che possa assumere il valore $+\infty$. E' chiaro che se si vogliono sommare le variazioni punto per punto di questa funzione, a livello infinitesimo, è necessario utilizzare integrali.
Il problema è stato creato da Gaspard Monge nel 1781, con il famoso paper [6]. Egli era una figura eclettica, che lavorò sia nei campi della fisica, dell'ingegneria e della matematica. Egli formulò il problema come la ricerca del minimo di un integrale il cui argomento è la funzione $c(x, y)$. Riscrivendo quello appena detto in simboli

$$inf\{I[\pi]; \pi \in \Pi(\mu, \nu)\}$$

dove

$$I[\pi] = \int_{X \times Y} c(x, y) d\pi(x, y)$$

e

$$\Pi(\mu, \nu) = \{\pi \in P(X \times Y) \mid \pi[A \times Y] = \mu[A], \pi[X \times B] = \nu[B] \ \forall A, B \text{ misurabili}\}$$

La formulazione di Monge era però limitata al caso di mappe di trasporto, ovvero $\nu(y) = \delta[y = T(x)]$, dove $T : X \mapsto \mathbb{R}$. Dopodichè ci fu un successivo sviluppo ad opera del matematico Leonid Kantorovich (premio nobel per l'economia nel 1975) che trattò il problema come un vero e proprio problema di Ottimizzazione nel 1942, ovvero formulandolo come il modo migliore di allocare delle risorse. Egli, a differenza di Monge, lasciò la misura ν come una misura di probabilità generica sullo spazio Y. Si può facilmente intuire che in questo caso μ indica la densità di probabilità della produzione, mentre ν quella di quanto viene acquistato. Kantorovich fu importante perchè semplificò il problema con strumenti di analisi funzionale, quali la dualità, e poichè sviluppò un concetto appropriato di distanza fra misure di probabilità. In seguito, il matematico francese Yann Brenier riutilizzò questa teoria per risolvere dei problemi in fluidodinamica. Il teorema più importante dal punto di vista dell'analisi matematica è il teorema 3. Esso dimostra che la soluzione al problema di Kantorovich si riduce naturalmente a

$$\nu(y) = \delta[y = \nabla\phi(x)]$$

dove ϕ è una funzione $\phi : X \mapsto \mathbb{R}$ le cui proprietà si analizzeranno nella tesi. La seconda parte di quest'ultima è incentrata sulle motivazioni che hanno portato Brenier ad utilizzare il trasporto ottimale nelle sue ricerche, che sono sfociate nella dimostrazione del teorema di fattorizzazione polare. Con il lavoro [3] è iniziato uno sviluppo del trasporto ottimale nell'ambito dell'equazioni alle derivate parziali, della probabilità, della analisi funzionale, della geometria, e della fluidodinamica. Viceversa, il trasporto ottimale si è spesso rivelato uno strumento per semplificare dei problemi nei campi sopracitati. Ci sono state anche due medaglie Fields che hanno contribuito in tempi recenti allo sviluppo del problema. La prima è Cedric Villani, che è anche lo scrittore di due libri di riferimento sull'argomento: il primo, da cui ho preso quasi tutto il materiale della mia tesi, è [13], che è dal punto di vista dell'analisi matematica; il secondo, [12], si focalizza più sulla geometria, sulla probabilità, e i sistemi dinamici. Più di recente, Alessio Figalli, è stato il secondo matematico premiato con la medaglia Fields che ha dato notevoli sviluppi all'argomento. In particolare, ha utilizzato il trasporto ottimale per risolvere equazioni riguardanti la meteorologia e la struttura dei cristalli. Un'ulteriore referenza è il libro scritto da Filippo Santambrogio: [9].

1 Il problema di Kantorovich per il costo quadratico

.
Il problema formulato da Kantorovich è definito come la minimizzazione di un integrale il cui argomento è una funzione di costo $c(x, y)$. Ovvero,

$$inf\{I[\pi]; \pi \in \Pi(\mu, \nu)\}$$

dove I è l'integrale che definisce il problema e

$$\pi \in \Pi(\mu, \nu) = \{\pi \in P(X \times Y) \mid \pi[A \times Y] = \mu[A], \pi[X \times B] = \nu[B] \ \forall A, B \text{ misurabili}\}$$

con (X, μ), (Y, ν) spazi di probabilità. Da qui in poi assumo che $X = Y = \mathbb{R}^n$.
La funzione di costo che si analizza è la norma euclidea al quadrato. Tuttavia è più comodo definire $c(x, y) = \frac{|x-y|^2}{2}$ invece di $|x - y|^2$. Si può dunque scrivere

$$I[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|x - y|^2}{2} d\pi(x, y).$$

Siano μ e ν due misure di probabilità(di Borel) su $\mathbb{R}^n$ con momenti finiti del secondo ordine, cioè

$$M_2 := \int_{\mathbb{R}^n} \frac{|x|^2}{2} d\mu(x) + \int_{\mathbb{R}^n} \frac{|y|^2}{2} d\nu(y) < +\infty \tag{1}$$

Questa condizione assicura che il funzionale I è sempre finito su $\Pi(\mu, \nu)$. Certamente, qualsiasi sia $\pi \in \Pi(\mu, \nu)$,

$$I[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|x - y|^2}{2} d\pi(x, y) \leq \int (|x|^2 + |y|^2) d\pi(x, y).$$

Inoltre, enuncio il principio di dualità di Kantorovich, il quale è il punto di partenza per dimostare alcune proposizioni. Esso è definito come

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = \sup_{(\phi,\psi) \in \Phi_c} J(\phi, \psi). \tag{2}$$

dove la famiglia Φ_c è lo spazio delle coppie $(\phi, \psi) \in L^1(\mu) \times L^1(\nu)$ tali per cui $\phi(x) + \psi(y) \leq c(x, y)$, mentre $J(\phi, \psi) := \int_{\mathbb{R}^n \times \mathbb{R}^n} (\phi(x) + \psi(y)) d\pi(x, y)$. Questa formula viene derivata esplicitamente nel capitolo 1 di *Topics in optimal trasportation* di Cedric Villani [13].

1.1 Il problema principale

Si inizia a dare risposta sommaria al problema di minimizzazione di Kantorovich

Teorema 1 (Esistenza di una misura ottimale)**.** *Il problema di minimizzazione* $inf\{I[\pi]; \pi \in \Pi(\mu, \nu)\}$ *ammette un minimo.*

Dimostrazione. Prima di tutto si nota che $\Pi(\mu, \nu)$ è non vuoto dato che $\mu \times \nu \in \Pi(\mu, \nu)$. Il punto cruciale è capire che $\Pi(\mu, \nu)$ è compatto rispetto alla topologia debole delle misure di probabilità. L'ultima affermazione si può dimostrare partendo dalla seguente disuguaglianza

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} (|x|^2 + |y|^2) d\pi(x, y) \leq 2M_2 < +\infty$$

ma questa è una proprietà generale che non ha bisogno del fatto che M_2 sia finito(Infatti se M_2 è finito si ottengono due integrali finiti $\int_{\mathbb{R}^n} |x|^2 d\mu(x) < +\infty$, $\int_{\mathbb{R}^n} |y|^2 d\nu(y) < +\infty$. Allora è un fatto noto di teoria della misura per cui $\forall \delta > 0$ esiste un insieme $E = E_x \times E_y$ con $\mu(E_x) < +\infty$ e $\nu(E_y) < +\infty$ tale per cui $\int_{\mathbb{R}^n \setminus E_x} |x|^2 d\mu < \delta$, $\int_{\mathbb{R}^n \setminus E_y} |y|^2 d\nu < \delta$. Non è restrittivo prendere E_x, E_y tali per cui contengano entrambi una palla in $\mathbb{R}^n$ di raggio 1. Questo implica che $\mu[X \setminus E_x] < \int_{\mathbb{R}^n \setminus E_x} |x|^2 d\mu < \delta$ e $\nu[Y \setminus E_y] < \int_{\mathbb{R}_n \setminus E_y} |y|^2 d\nu < \delta$. Allora, sia $\delta > 0$ fissato e siano $E_x \subset X$ e $E_y \subset Y$ tali che

$$\mu[X \setminus E_x] < \delta \text{ e } \nu[Y \setminus E_y] < \delta.$$

Qualsiasi sia $\pi \in \Pi(\mu, \nu)$,

$$\pi[(X \times Y) \setminus (E_x \times E_y]) \leq \pi[(Y \setminus E_y)] + \pi[(X \setminus E_x) \times L] \leq 2\delta$$

Questo prova che la famiglia di misure $\Pi(\mu, \nu)$ sia tight, e quindi il fatto che sia compatta rispetto alla topologia debole. Allora questo implica che esiste un minimizzatore di I. Posso dunque dire che se $(\pi_k)_{k \in N}$ è una successione minimizzante, essa ammette un punto di accumulazione $\pi_* \in \Pi(\mu, \nu)$. Ora pongo la funzione di costo $c(x, y) = \frac{|x-y|^2}{2}$ come l'estremo superiore di una successione crescente $(c_l)_{l \in \mathbb{N}}$ di funzioni continue e limitate. Utilizzando il teorema della convergenza monotona in (1) e il fatto che π_* è un punto di accumulazione di (π_k), la disuguaglianza $c_l \leq c$ e la proprietà minimizzante di π_*, si ottiene

$$\begin{aligned} \int c(x, y) d\pi_*(x, y) &\overset{(1)}{=} \lim_{l \to \infty} \int c_l(x, y) d\pi_*(x, y) \\ &\leq \lim_{l \to +\infty} \limsup_{k \to \infty} \int c_l(x, y) d\pi_k(x, y) \\ &\leq \limsup_{k \to \infty} \int c(x, y) d\pi_k(x, y) = inf I \end{aligned}$$

□

1.2 Il problema duale

La trattazione duale aiuta ad ottenere un risultato più preciso. La condizione affinchè (ψ, ϕ) appartengano a Φ_c è $\phi(x) + \psi(y) \leq \frac{|x-y|^2}{2}$ che vale per $\mu-$quasi ogni x e per $\nu-$quasi ogni $y \in \mathbb{R}^n$. Sviluppando il quadrato nel termine a destra, spostando i termini quadratici da destra a sinistra e invertendo il senso della disuguaglianza moltiplicando per (-1) si ottiene $x \cdot y \leq [\frac{|x|^2}{2} - \psi(x)] + [\frac{|y|^2}{2} - \phi(y)]$. Conviene quindi che per adeguarsi alla teoria delle funzioni convesse coniugate si considerino le nuove funzioni $\widetilde{\phi}(x) = \frac{|x|^2}{2} - \psi(x)$ e $\widetilde{\psi}(y) = \frac{|y|^2}{2} - \phi(y)$. Per non appesantire la notazione d'ora in poi la tilde sarà sottointesa. Si possono quindi scrivere le due disuguaglianze

$$\inf_{\pi \in \Pi(\mu, \nu)} I[\pi] = M_2 - \sup\{\int (x \cdot y) d\pi(x, y); \pi \in \Pi(\mu, \nu)\}$$

e

$$\sup_{\Phi_c} J = M_2 - inf\{J(\psi, \phi); (\phi, \psi) \in \widetilde{\Phi}\}$$

dove $\widetilde{\Phi}$ è il l'insieme di tutte le coppie (ψ, ϕ) in $L^1(d\mu) \times L^1(d\nu)$ con valori in $\mathbb{R} \cup \{+\infty\}$ tali per cui

$$x \cdot y \leq \psi(x) + \phi(y). \tag{3}$$

E' facile vedere che la condizione (1) diventa

$$\sup\left\{\int (x \cdot .y) d\pi(x, y); \pi \in \Pi(\mu, \nu)\right\} = \inf\{J(\psi, \phi); (\phi, \psi) \in \widetilde{\Phi}\}$$

Ora introduciamo il double convexification trick per ottenere le migliori coppie possibili nel problema duale. Siano $(\psi, \phi) \in \widetilde{\Phi}$, per (3) si può scrivere per $d\nu$ quasi ogni $y \in Y$

$$\psi(y) \geq \sup_x [x \cdot y - \phi(x)] =: \phi^*(y) \tag{4}$$

Osservazione 1. *sup_x può intendersi come l'estremo superiore essenziale rispetto a μ. Ma se si prende questa convenzione si entra in problemi di teoria della misura. Per questa ragione si può adottare una strategia differente. Dal momento che (3) vale per ogni coppia $(x, y) \in N_x^c \times N_y^c$ tale per cui $\mu(N_x) = 0$ e $\nu(N_y) = 0$ posso ridefinire ψ che valga $+\infty$ su N_x e ϕ che valga $+\infty$ su N_y. Il valore di $J(\psi, \phi)$ è ancora lo stesso dal momento che le modifiche apportate riguardano solo insiemi di misura nulla (dato che $\phi \in L^1(\mu)$ e $\psi \in L^1(d\nu)$). Si può quindi scrivere sup_x inteso come un vero e proprio estremo superiore. Altrimenti si può richiedere semplicemente che (ϕ, ψ) siano continue oltre al fatto che $x \cdot y \leq \psi(x) + \phi(y)$ valga per ogni x, y. In quest'ultima opzione il problema appena citato non sussisterebbe a causa del fatto che l'estremo superiore e l'estremo superiore essenziale coinciderebbero.* $\square$

Come conseguenza di (4) $J(\psi, \phi) \geq J(\phi, \phi^*)$.
Inoltre $\phi(x) \geq [x \cdot y - \phi^*(y)] =: \phi^{**}(x)$ e perciò

$$J(\phi, \phi^*) \geq J(\phi^{**}, \phi^*). \tag{5}$$

Si può concludere affermando la seguente proposizione

$$\inf_{(\psi,\phi)\in\widetilde{\Phi}} J(\psi, \phi) \geq \inf_{\phi\in L^1(d\mu)} J(\phi^{**}, \phi^*) \tag{6}$$

.
Se si ammette che $(\phi^{**}, \phi^*) \in L^1(d\mu) \times L^1(d\nu)$ allora sicuramente $(\phi^{**}, \phi^*) \in \widetilde{\Phi}$ e deriva da (6) che l'estremo inferiore di J rimanga lo stesso nel caso uno restringa i valori di J al sottoinsieme di $\widetilde{\Phi}$ costituito dalle coppie (ϕ^{**}, ϕ^*).
(ϕ^*, ϕ^{**}) sono delle funzioni molto particolari. Infatti quest'ultime sono funzioni convesse semicontinue inferiormente perchè definite come l'estremo superiore di una famiglia di funzioni lineari.

1.3 Richiami di analisi convessa

In questa sezione richiamo dei risultati di Analisi convessa utili per alcune dimostrazioni. Questi risultati si possono trovare in qualsiasi testo classico sull'argomento come Rockafellar [7].

Definizione 1 (funzione convessa)**.** *Una funzione ψ è convessa su $\mathbb{R}^n$ se è una funzione $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, non identicamente $+\infty$ tale per cui*

$$\forall x, y \in [0, 1],\ \psi(tx + (1-t)y) \leq \psi(x) + (1-t)\psi(y). \tag{7}$$

E' strettamente convessa se l'uguaglianza (7) implica che $x = y$ oppure $t = 0$, o $t = 1$. Si definisce $Dom(\phi)$ l'insieme (convesso) dei punti dove ψ è finita. Poichè è un insieme convesso il suo bordo è un insieme piccolo. Per convenzione gli insiemi tali per cui la dimensione di Hausdorff è al più $n - 1$ in $\mathbb{R}^n$ vengono chiamati insiemi piccoli. Questi hanno ovviamente misura di Lebesgue pari a 0.

1) differenziabilità
Si prova che ogni funzione convessa ϕ è automaticamente continua e localmente Lipschtiziana su $Int(Dom(\phi))$, il più grande insieme aperto incluso in $Dom(\phi)$. Come conseguenza, dal teorema di Rademacher, $\nabla\phi$ è ben definita quasi ovunque e localmente limitata in $Int(Dom(\phi))$. Inoltre l'insieme di punti dove $\nabla\phi$ non esiste è un insieme piccolo. I valori di ϕ sulla frontiera $\partial(Dom(\phi))$ possono essere modificati in modi differenti, senza cambiare le proprietà di convessità di ϕ. Ma se ϕ si assume che sia convessa semicontinua inferiormente, allora questo condiziona i valori di ϕ sulla frontiera. Infatti due funzioni convesse semicontinue inferiormente ψ e ϕ tali per cui $int(Dom(\psi)) = int(Dom(\phi))$ e tali per cui i loro valori coincidono sull'insieme appena scritto, sono uguali.
2) Il grafico giace sopra la sua tangente
Per tutti i punti x dove ϕ è differenziabile si ha la seguente relazione

$$\forall z \in \mathbb{R}^n,\ \phi(z) \geq \phi(x) + \nabla\phi(x) \cdot (z - x),$$

che esprime il fatto che il grafico di ϕ giace sopra il suo iperpiano tangente nel punto x. Una conseguenza è il fatto che $\nabla\phi$ è monotona, ovvero per qualsiasi ϕ differenziabile nei punti x e z,

$$\big\langle \nabla\phi(x) - \nabla\phi(z), x - z \big\rangle \geq 0.$$

3) sottodifferenziabilità.
E' utile introdurre una generalizzazione del concetto di differenziale chiamato sottodifferenziale. Il sottodifferenziale $\partial\phi$ di una funzione convessa è l'insieme di punti definito nel seguente modo

$$y \in \partial\phi(x) \Longleftrightarrow [\forall z \in \mathbb{R}^n, \phi(z) \geq \phi(x) + \big\langle y, z - x \big\rangle] \tag{8}$$

Ho deciso di porre come convenzione che si indichi la mappa $\partial\phi$ riferendosi al suo grafo, ovvvero a $Graph(\partial\phi)$. Utilizzando il teorema di Hahn Banach, si può dimostrare che per tutti gli $x \in Int(Dom(\phi))$, il sottodifferenziale $\partial\phi$ è non vuoto. In più ϕ è differenziabile in un punto x se solo se $\partial\phi(x)$ contiene un solo elemento che è $\nabla\phi(x)$.
Se ϕ è semicontinua inferiormente, allora la mappa del sottodifferenziale è sempre continua su tutto $\mathbb{R}^n$, nel senso che

$$\left.\begin{cases} x_k \longrightarrow x \\ \partial\phi(x_k) \ni y_k \longrightarrow y \end{cases}\right\} \Longrightarrow y \in \partial\phi(x) \tag{9}$$

- La mappa del sottodifferenziale genera il cono normale ai sottoinsiemi di livello di ϕ. Questo significa che se $C_x = \{z \in \mathbb{R}^n; \phi(z) \leq \phi(x)\}$ allora le direzioni di tutti i vettori in $\partial\phi(x)$ generano il cono normale $N_x = \{n_x \in \mathbb{R}^n; \forall z \in C_x, \langle n_x, z - x\rangle \leq 0\}$.

  Questo è la generalizzazione della proprietà che il gradiente nel caso di funzioni liscie è ortogonale alle curve di livello.

- Su $Int(dom(\phi))$, $\partial\phi(x) = \overline{convex}(\lim_{x_k \longrightarrow x} \nabla f(x_k))$

Continuità del gradiente: In seguito verrà inoltre utilizzata la seguente proprietà: Sia ϕ una funzione convessa semicontinua inferiormente, sia x un punto di differenziabilità di ϕ, e sia $y = \nabla\phi(x)$. Allora vale che $\forall\epsilon > 0$ $\exists\delta > 0$ tale per cui $\nabla\phi(B_\delta(x)) \subset \partial\phi(B_\delta(x)) \subset B_\epsilon(y)$ dove $B_r(x_0)$ indica la palla di raggio r centrata in x_0.
4) Monotonia:
E' facile ottenere un'altra proprietà del sottodifferenziale che è in comune con il gradiente:

$$\forall y_1 \in \partial\phi(x_1), \forall y_2 \in \partial\phi(x_2),\ \langle y_2 - y_1, x_2 - x_1\rangle \geq 0$$

5) funzioni coniugate.
Per ogni funzione $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, uno può definire la sua funzione convessa coniugata, o trasformata di Legendre, come

$$\phi^*(y) = \sup_{x \in \mathbb{R}^n} (x \cdot y - \phi(x)). \tag{10}$$

Allora ϕ^* è una funzione convessa semicontinua inferiormente. Dalla definizione

$$\forall x, y \in \mathbb{R}^n\ x \cdot y \leq \phi(x) + \phi^*(y) \tag{11}$$

Proposizione 1 (Caratterizzazione del sottodifferenziale)**.** *Sia ϕ una funzione convessa semicontinua inferiormente su $\mathbb{R}^n$. Allora*

$$\forall\ x, y\ \in \mathbb{R}^n\ x \cdot y = \phi(x) + \phi^*(y) \Longleftrightarrow y \in \partial\phi(x) \Longleftrightarrow x \in \partial\phi^*(y)$$

Dimostrazione

$$\begin{aligned} x \cdot y &= \phi(x) + \phi^*(y) \\ &\overset{(10)}{\Longleftrightarrow} x \cdot y \geq \phi(x) + \phi^*(y) \\ &\overset{(11)}{\Longleftrightarrow} \forall z \in \mathbb{R}^n, x \cdot y \geq \phi(x) + y \cdot z - \phi(z) \\ &\Longleftrightarrow \forall z \in \mathbb{R}^n, \phi(z) \geq \phi(x) + \langle y, z - x \rangle \\ &\overset{(8)}{\Longleftrightarrow} y \in \partial\phi(x) \end{aligned}$$

Ripetendo gli stessi passaggi e utilizzando la proposizione 2 vale anche $x \in \partial\phi^*(y)$.
6) Regolarizzazione. In analisi convessa, il ruolo che usualmente svolge la convoluzione affinchè la funzione sia più regolare (rende più smooth la funzione)lo svolge la inf convoluzione: qualsiasi siano ϕ e ψ due funzioni convesse, definisco $(\phi \square \psi)(z) = \inf_{x+x'=z} [\phi(x) + \psi(x')]$ La inf convoluzione ha un comportamento interessante sotto la trasformata di Legendre:

$$(\phi \square \psi)^* = \phi^* + \psi^*$$

7) Dualità e semicontinuità inferiore

Proposizione 2 (Dualità di Legendre per funzioni semicontinue inferiormente)**.** *Sia* $\phi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ *una funzione. Allora le seguenti proprietà sono equivalenti:*
(i) ψ *è convessa semicontinua inferiormente*
(ii) $\phi = \psi^*$ *per una certa funzione* ψ
(iii)$\phi^{**} = \phi$

Dimostrazione. E' ovvio che (iii) implica (ii) che, poichè le trasformate di Legendre sono semicontinue inferiormente, a sua volta implica (i). Rimane solamente da provare (i) $\Rightarrow$ (iii) Spezzo il ragionamento in 3 passaggi.
1)
Da (10), $\phi(x) \geq \sup_y [x \cdot y - \phi^*(y)] := \phi^{**}(x)$
2)
Sia $x \in Int(Dom(\phi))$. Dal momento che $\partial\phi(x) \neq \emptyset$ per il teorema di Hahn Banach, posso scegliere $y \in \partial\phi(x)$. Dalla proposizione 1, $\phi(x) + \phi^*(y) = x \cdot y$ da cui deriva che $\phi(x) \leq \sup_y [x \cdot y - \phi^*(y)] = \phi^{**}(x)$. Quindi ϕ e ϕ^{**} coincidono su $Int(Dom(\phi))$. Se $Dom(\phi) = \mathbb{R}^n$, allora $\phi^{**} = \phi$.
3)
Per trattare il caso generale possiamo rendere regolare ϕ con la inf convoluzione. Sia $\psi_\epsilon = \frac{|x|^2}{2\epsilon}$, e sia $\phi_\epsilon = \phi \square \psi_\epsilon$ Utilizzando il fatto che ϕ è semicontinua inferiormente e limitata dal basso da una funzione affine si dimostra che

$$\forall x \in \mathbb{R}^n \quad \phi(x) = \lim_{\epsilon \to 0} \phi_\epsilon(x).$$

Dal momento che $Dom(\phi_\epsilon) = \mathbb{R}^n$ si sa da 2) che $\phi_\epsilon^{**} = \phi_\epsilon$. D'altra parte $\phi_\epsilon \leq \phi$, cioè $\phi_\epsilon^* \geq \phi^*$. Quindi, utilizzando ancora una volta la trasformata di Legendre $\phi_\epsilon^{**} \leq \phi^{**}$. Si conclude che

$$\forall x \in \mathbb{R}^n \quad \phi^{**}(x) \geq \liminf_{\epsilon \to 0} \phi_\epsilon^{**}(x) = \liminf_{\epsilon \to 0} \phi_\epsilon(x) = \phi(x).$$

□

Osservazione 2. *(i) E' facile costruire funzioni convesse che non sono semicontinue inferiormente tali per cui* $\phi \neq \phi^{**}$*. Questo tuttavia può avvenire solamente su* $\partial Dom(\phi)$*.*
La dualità di Legendre ha delle conseguenze importanti. Se ϕ *è strettamente convessa nell'intorno di un certo* $x \in \mathbb{R}^n$*, allora* ϕ^* *è differenziabile su* $\partial\phi(x)$ *e* $\nabla\phi^*(y) = x$ *per tutti gli* $y \in \partial\phi(x)$*.*

Dalla relazione $x \cdot y = \phi(x) + \phi^*(y)$ si ottiene per differenziazione

$$(\nabla\phi)^{-1} = \nabla\phi^*. \tag{12}$$

Se ϕ è superlineare, cioè

$$\lim_{+\infty} \frac{\phi(x)}{|x|} = +\infty$$

allora $\nabla\phi(\mathbb{R}^n) = \mathbb{R}^n$, caso in cui $\nabla\phi$ è una biiezione $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ e dunque $\nabla\phi^*$ è la sua inversa.

Osservazione 3. *Nel caso ϕ e ϕ^* siano due volte differenziabili c'è una corrispondenza duale fra la convessità stretta di ϕ e la derivabilità di ϕ^*, che può essere esplicitato nel seguente modo*

$$D^2\phi^*(\nabla\phi(x)) = [D^2\phi(x)]^{-1}$$

Questa formula si ottiene differenziando (12)*.*

Teorema 2 (Esistenza di un paio ottimo di funzioni convesse coniugate)**.** *Siano μ, ν due misure di probabilità su $\mathbb{R}^n$, con momenti finiti del secondo ordine. Sia $\widetilde{\Phi}$ definito da* (3)*.*
Allora esistono una coppia (ϕ, ϕ^*) di funzioni convesse semicontinue inferiormente su $\mathbb{R}^n$, tali per cui

$$\inf_{\widetilde{\Phi}} J = J(\phi, \phi^*)$$

Strategia della dimostrazione:
Considero una successione minimizzante $(\phi_k, \psi_k)_{k\in N}$. Dal double convexification trick, considero da subito che (ϕ_k, ψ_k) siano funzioni convesse coniugate(dimentichiamoci per un attimo che si deve dimostrare che le funzioni $(\phi_k, \psi_k) \in L^1(d\mu) \times L^1(d\nu)$). Ora vorrei dimostrare che estraendo una sottosuccesione di funzioni $(\phi_k, \psi_k) \longrightarrow (\phi, \psi) \in (L^1(d\mu), L^1(d\nu))$ e

$$(\phi, \psi) \in \widetilde{\Phi}, J(\phi, \psi) \leq \liminf_{k\to+\infty} J(\phi_k, \psi_k)$$

Quindi (ϕ, ψ) sarebbe la coppia ottimale, ed è sufficiente utilizzare il double convexification trick per ottenere la coppia adatta.

□

In ogni caso questo risultato non è vero in totale generalità poichè il problema è invariante per addizione e sottrazione successiva di una stessa costante:

$$\forall a \in \mathbb{R}, J(\phi - a, \psi + a) = J(\phi, \psi) \tag{13}$$

Quindi nel caso (ϕ_k, ψ_k) sia una successione minimizzante convergente a (ψ, ϕ) anche $(\phi_k - k, \psi_k + k)$ è una successione minimizzante, ma non ci si può aspettare che converga ad una coppia di funzioni appartenenti ad L^1.

Lemma 1 (Double convexification lemma)**.** *Siano μ, ν due probabilità di misura con il supporto rispettivamente in $X, Y \in \mathbb{R}^n$, che soddisfano $M_2 = \int_X \frac{|x|^2}{2} d\mu(x) + \int_Y \frac{|y|^2}{2} d\nu(y) < +\infty$.*
Sia $\widetilde{\Phi}$ definito da (3) *e sia (ϕ_k, ψ_k) una successione minimizzante per J su $\widetilde{\Phi}$. Allora*
(i) Uno può modificare (ϕ_k, ψ_k) su insiemi di misura trascurabile rispetto a ν e a μ in modo che la disuguaglianza (3) *sia rispettata per tutti gli $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ senza cambiare i valori di $J(\phi_k, \psi_k)$.*
(ii)Esiste una seccessione di numeri reali $(a_k)_k \in N$ tale per cui $(\overline{\phi_k}, \overline{\psi_k}) = (\phi_k^{**} - a_k, \phi_k^* + a_k)$ continua a essere una successione minimizzante di J su $\widetilde{\Phi}$ e soddisfa le stime dal basso $\forall x \in X$, $\forall y \in Y$

$$\overline{\phi_k}(x) \geq -\frac{|x|^2}{2}, \quad \overline{\phi_k}(y) \geq -\frac{|y|^2}{2}, \tag{14}$$

insieme alle stime dall'alto

$$\liminf_{k\to+\infty}\inf_{x\in X}(\overline{\phi_k}(x)+\frac{|x|^2}{2})\le\inf_{\widetilde{\Phi}}J+M_2\quad\liminf_{k\to+\infty}\inf_{y\in Y}(\overline{\psi_k}(y)+\frac{|y|^2}{2})\le\inf_{\widetilde{\Phi}}J+M_2. \tag{15}$$

(iii) In particolare con la scelta X=Y=$\mathbb{R}^n$ si ha

$$\inf_{\widetilde{\Phi}}J(\phi,\psi)=\inf_{\phi\in L^1(d\nu)}J(\phi^{**},\phi^*).$$

Allora l'estremo inferiore di J su $\widetilde{\Phi}$ non cambia a patto di restringere J all'insieme di coppie di funzioni convesse coniugate in $\widetilde{\Phi}$.

Osservazione 4. *(i) Non c'è niente di speciale circa la scelta di $\mathbb{R}^n$ in questo lemma. Funziona tutto allo stesso modo se X è uno spazio arbitrario di Banach E e Y è lo spazio topologico duale E^*, a condizione che il prodotto scalare di $\mathbb{R}^n$ sia sostituito dal prodotto di dualità tra X e Y.*
(ii)La strategia della dimostrazione sotto può essere usata per un problema di Monge-Kantorovich più generale, non necessariamente in uno spazio di Banach, solamente con la condizione che la funzione di costo c soddisfa una stima del tipo $c(x,y)\le c_X(x)+c_Y(y)$ con $c_X\in L^1(d\mu)$, $c_Y\in L^1(d\nu)$. Per ottenere un esempio è sufficiente prendere $c_X(x)=|x|^2$ e $c_Y(y)=|y|^2$. Esse, infatti, se vengono integrate sui rispettivi domini e sommate danno come risultato M_2, che per ipotesi è finito e maggiora $I[\pi]$; l'integrale che per definizione è associato alla funzione di costo $c(x,y)=\frac{|x-y|^2}{2}$.

Dimostrazione. Abbiamo già dato prova di (i) nella osservazione 1, e (iii) è una conseguenza diretta di (ii). Quindi è necessario dimostrare solamente (ii).
1)Sia (ϕ_k,ψ_k) una successione minimizzante per J. Posso assumere che per ogni x,y, $\phi_k(x)\ge x\cdot y-\psi_k(y)$. Dal momento in cui ψ_k non è identicamente $+\infty$ si può dedurre che ϕ_k è limitata dal basso da una funzione affine $x\cdot y_0+b_0$ per qualche $y_0\in Y$. Segue che $\phi_k^*(y_0)=\sup_{x\in X}[x\cdot y-\phi_k(x)]\le -b_0$. In particolare ϕ_k non è identicamente $+\infty$. Rifacendo gli stessi passaggi sostituendo ϕ_k^* a ψ si nota che ϕ_k^* è limitata dal basso da una funzione affine. Questo implica che $a_k:=\inf_{y\in Y}(\phi^*(y)+\frac{|y|^2}{2})$ è finito.
Ora prendo $(\overline{\phi_k},\overline{\psi_k})=(\phi_k^{**}+a_k,\phi_k^*-a_k)$. Si può notare che $\overline{\phi_k}=(\overline{\psi_k})^*$.
Per costruzione $\inf_{y\in Y}(\overline{\psi}_k(y)+\frac{|y|^2}{2})=0$. In più

$$\phi_k(x)+\frac{|x|^2}{2}=(\overline{\psi_k})^*(x)+\frac{|x|^2}{2}=\sup_{y\in Y}\left[x\cdot y-\overline{\psi_k}(y)+\frac{|x|^2}{2}\right]\ge\sup_{y\in Y}\left[-\overline{\psi_k}(y)-\frac{|y|^2}{2}\right]=-\inf_{y\in Y}\left[\frac{|y|^2}{2}+\overline{\psi_k}(y)\right]=0.$$

Quindi $(\overline{\phi_k},\overline{\psi_k})$ soddisfa (14)
2)Ora controllo la integrabilità di $(\overline{\phi_k},\overline{\psi_k})$. Dall'invarianza per traslazioni del problema di Kantorovich e dalle disuguaglianza (5),

$$J(\overline{\phi_k},\overline{\psi_k})=J(\phi_k^*,\phi_k^{**})\overset{(5)}{\le}J(\phi_k,\psi_k)\le+\infty \tag{16}$$

L'ultimo termine converge a causa del fatto che M_2 è finito per ipotesi.
In particolare,

$$\int_X\left(\frac{|x|^2}{2}+\overline{\phi_k}(x)\right)d\mu(x)+\int_Y\left(\frac{|y|^2}{2}+\overline{\psi_k}(y)\right)d\nu(y)\le+\infty$$

Dal momento che entrambe le funzioni integrande sono non negative, si deduce che sono integrabili, e dalle assunzioni su (μ,ν) si deduce che $\overline{\phi_k}\in L^1(d\mu),\overline{\psi_k}\in L^1(d\nu)$. Allora per definizione(M_2 è finito per ipotesi) $(\phi_k,\psi_k)\in\widetilde{\Phi}$ e per la seconda disuguaglianza in 16 costituiscono una successione minimizzante.
3) Per completare la dimostrazione del lemma, rimane da dimostrare solamente (15) . Si può notare che

$$J(\phi_k,\psi_k)+M_2=\int_X(\phi_k(x)+\frac{|x|^2}{2})d\mu(x)+\int_Y(\psi_k(y)+\frac{|y|^2}{2})d\nu(y)\ge\inf_{x\in X}(\phi_k(x)+\frac{|x|^2}{2})+\inf_{y\in Y}(\psi_k+\frac{|y|^2}{2}).$$

Dal momento che entrambe le espressioni dentro le parentesi sono non negative segue che

$$\liminf_{k\longrightarrow+\infty}\inf_{x\in X}(\phi_k(x)+\frac{|x|^2}{2})\leq\liminf_{k\longrightarrow+\infty}J(\phi_k,\psi_k)+M_2=\inf_{\tilde\Phi}J+M_2.$$

□

Dimostrazione del teorema 2. Dimostro solamente il caso in cui μ,ν hanno il supporto contenuto in insiemi compatti X,Y di $\mathbb{R}^n$. Sia (ψ_k,ϕ_k) una successione minimizzante di J su $\widetilde{\Phi}$. Sia $(\overline{\psi_k},\overline{\phi_k})$ una successione minimizzante costruita con il Lemma appena dimostrato. Scrivendo esplicitamente la coppia sopracitata

$$\overline{\phi}_k(y)=\sup_{x\in X}[x\cdot y-\phi_k(x)-a_k]$$
$$\overline{\psi}_k(x)=\sup_{y\in Y}[x\cdot y-\phi_k^*(y)+a_k]$$

deduco le stime

$$\|\overline{\psi_k}\|_{Lip(Y)}\leq\sup_{x\in X}\|x\|$$
$$\|\overline{\phi_k}\|_{Lip(X)}\leq\sup_{y\in Y}\|y\| \tag{17}$$

dove ho definito $\|\psi\|_{Lip(\cdot)}=\sup_{y_1\neq y_2}\frac{|\psi(y_1)-\psi(y_2)|}{\|y_1-y_2\|}$

Perciò $\overline{\phi_k}$, $\overline{\psi_k}$ sono uniformemente continue. D'altra parte, dal lemma precedente ho che per k grandi a sufficienza esistono $x_k\in X$ e $y_k\in Y$ tali per cui

$$-\sup_{x\in X}\frac{|x|^2}{2}\leq\overline{\phi_k}(x_k)\leq\sup_{x\in X}\frac{|x|^2}{2}+infJ+M_2+1$$
$$-\sup_{y\in Y}\frac{|y|^2}{2}\leq\overline{\psi_k}(y_k)\leq\sup_{y\in Y}\frac{|y|^2}{2}+infJ+M_2+1. \tag{18}$$

Combinando queste ultime due righe di disuguaguaglianze con la limitatezza di X, Y e con 17 si vede che $\overline{\phi_k}$ e $\overline{\psi_k}$ sono uniformemente limitate. Infatti $\overline{\phi_k}(x)=\overline{\phi_k}(x_k)-\overline{\phi_k}(x_k)+\overline{\phi_k}(x)$ è limitata per ogni x perchè gli ultimi due termini nell'uguaglianza sono limitati grazie alla lipschtzianità di $\overline{\phi_k}$, e alla limitatezza di X,Y. Ho essenzialmente esteso da ogni x_k(rispettivamente y_k) a ogni x(rispettivamente y) (18) Poichè $\overline{\phi_k}$ e $\overline{\psi_k}$ sono Lipschitziane, ovvero sono equicontinue (Si deve quindi dimostrare che $\forall\epsilon>0\;\forall\;x,z\;\in\;X\;\exists\delta>\|x-z\|>0\mid|\overline{\phi_k}(x)-\overline{\phi_k}(z)|<\epsilon$. Affinchè $|\overline{\phi_k}(x)-\overline{\phi_k}(z)|<\epsilon$ è sufficiente prendere un $\delta<\frac{\epsilon}{\sup_{y\in Y}\|y\|}$. Nel medesimo modo $\overline{\psi_k}$ rispetta la definizione di equicontinuità), e poichè sono uniformemente limitate, posso utilizzare il teorema di Ascoli-Arzelà per dedurre che se al più estraggo una sottosuccessione, le due successioni di funzioni convergono uniformemente in $C_b(X),C_b(Y)$, ai limiti continui $\overline{\phi}$, rispettivamente $\overline{\psi}$. La convergenza in norma uniforme implica in questo caso la convergenza in L^1, ovvero $J(\overline{\phi},\overline{\psi})=\lim_{k\longrightarrow+\infty}J(\overline{\phi_k},\overline{\psi_k})$, quindi $(\overline{\psi},\overline{\phi})$ è una coppia ottimale. (a priori definite solo su $X\times Y$). Ora possiamo estendere $\overline{\psi}$, $\overline{\phi}$ fuori X,Y a $+\infty$ e applicare il double convexification trick nel senso usuale(ovvero ponendo $X=Y=\mathbb{R}^n$ come nel lemma 1) per trovare un paio di funzioni convesse coniugate che soddisfino le proprietà interessate. □

Teorema 3. *[Teorema del trasporto ottimale con costo quadratico]*
Siano (μ,ν) *misure di probabilità su* $\mathbb{R}^n$ *con momenti finiti del secondo ordine nel senso di* (1). *Si consideri il problema di Monge-Kantorovich con la funzione quadratica di costo* $c(x,y)=|x-y|^2$. *Allora,*

*(i)(**Criterio di ottimalità di Knott-smith**) $\pi \in \Pi(\mu, \nu)$ è ottimale se e solo se esiste una funzione convessa semicontinua inferiormente tale per cui*

$$Supp(\pi) = Graph(\partial\phi)$$

o equivalentemente per

$$d\pi\text{-quasi tutti gli } (x, y), \quad y \in \partial\phi(x) \tag{19}$$

In più, la coppia (ϕ, ϕ^*) deve essere un minimo nel problema

$$\inf \left\{ \int_{\mathbb{R}^n} \phi d\mu + \int_{\mathbb{R}^n} \psi d\nu; \forall(x, y),\ x \cdot y \leq \phi(x) + \psi(y) \right\}$$

*(ii)(**Teorema di Brenier**) Se μ ha misura nulla sugli insiemi piccoli c'è un unico piano di trasporto ottimale che è*

$$d\pi(x, y) = d\mu(x)\delta[y = \nabla\phi(x)], \tag{20}$$

o equivalentemente

$$\pi = (Id \times \nabla\phi)\#\mu$$

dove $\nabla\phi$ è l'unico gradiente(unicamente determinato μ quasi ovunque) di una funzione convessa tale per cui ν è la misura pushforward di μ tramite $\nabla\phi$: $\nabla\phi\#\mu = \nu$. In più

$$supp(\nu) = \nabla\phi(Supp\mu)$$

(iii) Come corollario, sotto l'assunzione di (ii), $\nabla\phi$ è l'unica soluzione al problema del transporto ottimale:

$$\int_{\mathbb{R}^n} |x - \nabla\phi(x)|^2 d\mu(x) = \inf_{T\#\mu=\nu} \int_{\mathbb{R}^n} |x - T(x)|^2 d\mu(x),$$

o equivalentemente

$$\int_{\mathbb{R}^n} x \cdot \nabla\phi(x) d\mu(x) = \sup_{T\#\mu=\nu} \int_{\mathbb{R}^n} x \cdot T(x) d\mu(x).$$

(iv)Infine, se anche ν ha misura nulla sugli insiemi "piccoli", allora per $d\mu$-quasi ogni x e $d\nu$-quasi ogni y,

$$\nabla\phi^* \circ \nabla\phi(x) = x, \quad \nabla\phi \circ \nabla\phi^*(y) = y,$$

e $\nabla\phi^*$ è l'unico gradiente di una funzione convessa che trasporta (tramite pushforward) ν su μ, e anche la soluzione del problema di Monge per trasportare ν su μ con una funzione di costo quadratica.

Osservazione 5. *(i)E' facile vedere che, in generale non c'è unicità al problema di Kantorovich: è sufficiente definire (X, μ) e (Y, ν) spazi di misura tali che $X = Y = \mathbb{R}^2$ e $\mu = \delta_{(0,0)} + \delta_{(1,1)}$, $\nu = \delta_{(0,1)} + \delta_{(1,0)}$. Si deve anche notare che ci può essere una T non misurabile tale per cui vale la relazione $T\#\mu = \nu$.*
(ii)$\nabla\phi$ non è definito ovunque. Nel Teorema di Brenier questa informazione non è importante. Infatti μ è definito solo sull'insieme in cui ϕ è differenziabile. L'assunzione tale per cui μ ha misura nulla sugli insiemi piccoli è quindi fondamentale.

Dimostrazione: (1) Il teorema 1 dimostra l'esistenza di un piano di trasporto ottimale π. Dalla proposizione 2 esiste una coppia di funzioni convesse coniugate (ϕ, ϕ^*), che sono sono ottimali per il problema duale. Ora, dalla definizione del minimo del problema duale

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} (x \cdot y) d\pi(x, y) = \int_{\mathbb{R}^n} \phi d\mu + \int_{\mathbb{R}^n} \phi^* d\nu = \int_{\mathbb{R}^n \times \mathbb{R}^n} [\phi(x) + \phi^*(y)] d\pi(x, y)$$

Ovvero

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}[\phi(x)+\phi^*(y)-x\cdot y]d\pi(x,y)=0$$

Ma la funzione dentro la parentesi è non negativa per come è stato definito il problema duale in (3), ovvero deve essere uguale a 0 per dπ-quasi ogni (x,y). Ciò implica (19) per la proposizione 1.
(2) Sia invece $\pi\in\Pi(\mu,\nu)$ che soddisfa (19). Allora

$$\phi(x)+\phi^*(y)=x\cdot y.$$

Ciò implica che

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)d\pi(x,y)=\int_{\mathbb{R}^n\times\mathbb{R}^n}\phi d\mu+\int_{\mathbb{R}^n\times\mathbb{R}^n}\phi^* d\nu.$$

Quindi (ϕ,ϕ^*) sono ottimali nel problema duale. Ciò conclude la dimostrazione di (i).
(3) Ora assumiamo che μ abbia misura nulla sugli insiemi "piccoli", allora sia ϕ come nell'enunciato. Quest'ultima appartiene a $L^1(\mu)$, ovvero $\phi<+\infty$ per $d\mu$-quasi ogni $x\in X:\mu[Dom(\phi)]$=1. D'altra parte il bordo $\partial Dom(\phi)$ di un insieme $Dom(\phi)$ in cui ϕ è una funzione convessa è un insieme piccolo, che per μ ha misura nulla, detto in altro modo $\mu[Int(Dom(\phi))]=1$. Ora su $Int[Dom(\phi)]$ l'insieme in cui ϕ non è differenziabile è un insieme piccolo. Quindi per dμ-quasi ogni $x\ \in\ X$ è un punto del sottodifferenziabile di ϕ per il Teorema di Rademacher. Per μ-quasi ogni x,poichè ϕ è una funzione convessa, il sottodifferenziale nel punto x è $\{\nabla\phi(x)\}$. Richiamando il fatto che una proposizione vera per $d\mu$-quasi ogni x è vera anche per dπ quasi ogni (x,y) si ottiene che $y=\nabla\phi(x)$ per dπ-quasi ogni (x,y).
(4) E' necessario provare l'unicità di ϕ. Supponendo che $\overline{\phi}(x)$ è un'altra funzione convessa tale per cui $\nabla\overline{\phi}\#\mu=\nu$. Voglio provare che $\nabla\phi(x)=\nabla\overline{\phi}(x)$ a meno di insiemi μ-trascurabili. Da (i), (Id $\times\nabla\overline{\phi})\#\mu=\nu$ è un piano di trasporto ottimo, e di conseguenza $(\overline{\phi},\overline{\phi}^*)$ è una coppia ottimale per il problema duale, esattamente come (ϕ,ϕ^*). Perciò

$$\int_{\mathbb{R}^n}\overline{\phi}d\mu+\int_{\mathbb{R}^n}\overline{\phi}^*d\nu=\int_{\mathbb{R}^n}\phi d\mu+\int_{\mathbb{R}^n}\phi^*d\nu.$$

Sia π il piano di trasporto ottimo associato a ϕ. Dalla formula appena scritta ho che

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}\left(\overline{\phi}+\overline{\phi}^*\right)d\pi(x,y)=\int_{\mathbb{R}^n\times\mathbb{R}^n}\left(\phi+\phi^*\right)d\pi(x,y)=\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)d\pi(x,y)$$

Dal momento in cui $\pi=(Id\times\nabla\phi)\#\mu$ l'ultima equazione può essere riscritta come

$$\int_{\mathbb{R}^n}[\overline{\phi}(x)+\overline{\phi}^*(\nabla\phi(x))]d\mu(x)=\int_{\mathbb{R}^n}x\cdot\nabla\phi(x)d\mu(x).$$

Perciò

$$\int_{\mathbb{R}^n}[\overline{\phi}(x)+\overline{\phi}^*(\nabla\phi(x))-x\cdot\nabla\phi(x)]d\mu(x)=0.$$

Dal momento che l'argomento fra parentesi quadre è non negativo segue che esso sia uguale a 0 d$\mu-$quasi ovunque. Usando di nuovo la caratterizzazione del sottodifferenziale (Proposizione 1) segue che $\nabla\phi(x)\ \in\partial\overline{\phi}(x)$ per dμ-quasi ogni $x\ \in\ X$. Per il teorema di Rademacher $\overline{\phi}$ è differenziabile quasi ovunque per la misura di Lebesgue, ovvero μ-quasi ogni x. Dunque

$$\nabla\phi(x)=\nabla\overline{\phi}(x)$$

per quasi ogni x, che termina la dimostrazione di (ii).
(5)Ora dimostro che Supp(ν) $=\overline{\nabla\phi(Supp(\mu))}$. Sia $x\in Supp(\mu)$ un punto di differenziabilità di ϕ e sia $y=\nabla\phi(x)$. Dalla "continuità del gradiente" so che $\forall\epsilon>0\ \exists\delta>0$ tale per cui $\nabla\phi(B_\delta(x))\subset B_\epsilon(y)$ e in particolare

$$\nu[B_\epsilon(y)]\geq\mu[\nabla\phi^{-1}(\nabla\phi(B_\delta(x)))]\geq\mu[B_\delta(x)] \tag{21}$$

. La prima disuaguaglianza si ottiene dalla definizione di $\nu = \nabla\phi\#\mu$, mentre la seconda deriva semplicemente dalla definizione di funzione.
Se $\mu[B_\delta(x)] > 0$ allora $x \in supp(\mu)$. Ciò implica però che $\nu[B_\epsilon(y)] > 0$ per (21).
Dal momento che ϵ è arbitrariamente piccolo, deduco che $y \in Supp(\nu)$. Si conclude che

$$\nabla\phi(Supp(\mu)) \subset Supp(\nu) \tag{22}$$

(6) D'altra parte, $\nu[\nabla\phi(Supp(\mu))] \geq \mu[Supp(\mu)] = 1$. La misura del supporto di μ è uguale a 1 poichè μ è una misura di probabilità. Quindi ν è concentrata su $\nabla\phi(supp(\mu))$ e perciò per la definizione di supporto

$$supp(\nu) \subset \overline{\nabla\phi(Supp(\mu))}. \tag{23}$$

Combinando questo con 22 otteniamo che $Supp(\nu) = \overline{\nabla\phi(Supp(\mu))}$.
(7)(iii) è ora ovvio perchè so quale è la soluzione unica del problema e ho anche verificato la condizione $T\#\mu = \nu$.
(8) Se π è ottimale allora abbiamo dimostrato che $y = \nabla\phi(x)$ $d\pi(x,y)$-quasi ogni (x,y), che è equivalente a $x \in \partial\phi^*(y)$ per la caratterizzazione del sottodifferenziale. Dal momento che ϕ^* è dν-finita quasi ovunque allora è anche differenziabile quasi ovunque. Quindi, $x = \nabla\phi^*(y) = \nabla\phi^*(\nabla\phi(x))$ $d\pi(x,y)$-quasi ovunque. Segue per simmetria che $\nabla\phi(\nabla\phi^*(y)) = y$ per dν-quasi ogni y. □

2 Il Teorema di fattorizzazzione polare di Brenier

2.1 Riarrangiamenti e fattorizzazioni polari

.

Definizione 2 (riarrangiamento). *Sia $m : (W,\lambda) \to (X,\mu)$ una funzione misurabile tra due spazi di misura W e X, muniti rispettivamente delle misure λ e μ. Un'altra funzione $\widetilde{m} : (W,\lambda) \to (X,\mu)$ viene detta un riarrangiamento di m, se ha le seguenti proprietà: sia $F : X \to \mathbb{R}$ una funzione misurabile tale per cui $F \circ m \in L^1(d\lambda)$, allora $F \circ \widetilde{m} \in L^1(d\lambda)$ e*

$$\int_W (F \circ m) d\lambda = \int_W (F \circ \widetilde{m}) d\lambda. \tag{24}$$

Osservazione 6. *Se $\lambda[W] < +\infty$, allora è equivalente a richiedere che entrambi gli integrali in* (24) *coincidano per ogni funzione F misurabile e limitata.*

La definizione (2) significa che "non si può dire chi sia m o $\widetilde{m}$ guardando solamente ai valori assunti dalle due". Per esempio, m e $\widetilde{m}$ dovrebbero avere lo stesso massimo e lo stesso minimo, ma non importa in che punto. Se $X = \mathbb{R}^n$ e $F(x) = |x|^p$, si vede dalla definizione che $\|m\|_{L^p} = \|\widetilde{m}\|_{L^p}$, per ogni p, quindi le norme di Lebesgue sono invarianti sotto riarrangiamento. D'altra parte la maggior parte di riarrangiamenti di una funzione derivabile non sono derivabili; in particolare le norme di Sobolev $\|\nabla m\|_{L^p}$ non sono invarianti sotto riarrangiamenti. Il modo più semplice di costruire riarrangiamenti è attraverso mappe che preservano la misura.

Definizione 3 (Mappe che preservano la misura). *Sia (W,λ) uno spazio di misura dato. Una funzione misurabile $s : W \to W$ preserva la misura se*

$$s\#\lambda = \lambda.$$

In altre parole, per ogni insieme misurabile $A \subset W$, si ha $\lambda[s^{-1}(A)] = \lambda[A]$.

Esempio 1. *Sia Ω un sottoinsieme aperto di $\mathbb{R}^n$, munito della misura di Lebesgue n-dimensionale λ. Con un cambio di variabili si nota che un diffeomorfismo(di classe C^1) $s : \Omega \to \Omega$ preserva la misura se e solo se ha Jacobiano unitario:*

$$|\det(\nabla s)| = 1 \tag{25}$$

Perciò l'insieme dei diffeomorfismi che soddisfano (25) *è il gruppo di diffeomorfismi che preservano la misura su Ω. Un importante sottogruppo di questa classe è il gruppo di tutti i diffeomorfismi s con* $\det(\nabla s) = 1$. *Nel seguito, l'insieme di tutte le mappe che preservano la misura su uno spazio di misura (W,λ) sarà chiamato $S(W)$. Il gruppo di diffeomorfismi $s : \Omega \to \Omega$ con jacobiano unitario si denoterà come $SD(\Omega)$, e il sottogruppo di diffeomorfismi $s : \Omega \to \Omega$ tale per cui* $\det(\nabla s) = 1$ *si denoterà con $G(\Omega)$.*

Osservazione 7. *I diffeomorfismi che preservano la misura costituiscono un piccolo insieme di $S(\Omega)$. Per esempio, $SD((0,1))$ si riduce solo a $\{Id, -Id\}$, ma ci sono molti elementi in $S((0,1))$ che non sono diffeomorfismi, come $s(x) = 2x1_{\{x\leq 1/2\}} + (1-2x)1_{\{x>1/2\}}$*

Proposizione 3 (Mappe che preservano misure e riarrangiamenti)**.** *Sia $m : (W,\lambda) \to (X,\mu)$ una funzione misurabile e siano (W,λ), (X,μ) due spazi di misura. Se $s \in S(W)$ e $\widetilde{m} = m \circ s$, allora $\widetilde{m}$ è un riarrangiamento di m. Viceversa, se $\widetilde{m}$ è un riarrangiamento di m, alllora $\widetilde{m}^{-1} \circ m \in S(W)$.*

Dimostrazione. Assumo che $\widetilde{m} = m \circ s$, $s \in S(W)$. Allora, per tutte le F misurabili $F : W \to \mathbb{R}_+$,

$$\int (F \circ \widetilde{m}) d\lambda = \int (F \circ m) \circ s d\lambda = \int (F \circ m) d(s\#\lambda) = \int (F \circ m) d\lambda.$$

Quindi $\widetilde{m}$ è un riarrangiamento di m.
2) Sia $\widetilde{m}$ un riarrangiamento di m. Definisco $s = \widetilde{m}^{-1} \circ m$ e considero ogni funzione misurabile non negativa F su W. Allora

$$\int F d(s\#\lambda) = \int (F \circ s) d\lambda = \int (F \circ \widetilde{m}^{-1}) \circ m d\lambda = \int (F \circ \widetilde{m}^{-1}) \circ \widetilde{m} = \int F d\lambda.$$

Quindi s preserva la misura. □ □

Fino ad ora tutto è estremamente generale. Nel seguito, considereremo solo il caso in cui W è un sottoinsieme di $\mathbb{R}^n$. Una questione generale è se lì esiste una classe particolare di funzioni Σ, con delle buone proprietà, come ad esempio quelle tali per cui ogni funzione $m : W \to X$ ammette un riarrangiamento in Σ. Quando $W = \mathbb{R}^n$ e $X = \mathbb{R}_+$, c'è una classe molto famosa di quel tipo, cioè l'insieme di funzioni dellla forma $F(|x - x_0|)$, dove x_0 è arbitrario e F è non-negativa e nondecrescente. In questo caso considero una situazione differente, in cui $W \subset \mathbb{R}^n$ e $X \subset \mathbb{R}^n$, e sia Σ l'insieme dei gradienti delle funzioni convesse. Il risultato principale è il teorema di Brenier.

Teorema 4 (Teorema di fattorizzazione polare di Brenier)**.** *Sia Ω un sottoinsieme limitato di $\mathbb{R}^n$, con una misura di Lebesgue positiva. Sia $h : \Omega \to \mathbb{R}^n$ una mappa a valori vettoriali di L^2 che soddisfa la condizione di nondegenerazione*

$$\text{per ogni insieme piccolo } N \subset \mathbb{R}^n \quad |h^{-1}(N)| = 0 \tag{26}$$

Allora esiste un unico riarrangiamento $\nabla\phi$ di h nella classe dei gradienti di funzioni convesse in L^2, e un'unica mappa che preserva la misura $s \in S(\Omega)$, tale per cui $h = \nabla\phi \circ s$. In più, s è l'unica proiezione di h su $S(\Omega)$ in L^2.

Osservazione 8. *(i) Le norme L^2 nel teorema di fattorizzazione sono prese con la misura di Lebesgue su Ω. Quando dico che $\nabla\phi$ è il gradiente di una funzione convessa, intendo che $\nabla\phi$ è la restrizione su Ω del gradiente di una funzione convessa ϕ su $\mathbb{R}^n$. Quando dico che s è la proiezione L^2 di h su $S(\Omega)$ intendo che s minimizza $\|h - \sigma\|_{L^2(\Omega)}$ fra tutti i $\sigma \in S(\Omega)$.*
(ii)Preciso che (26) *non è la condizione necessaria per l'esistenza della fattorizzazione polare, scoperta dovuta a Burton e Douglas [4]. Questa condizione è però necessaria per l'unicità.*

2.2 Motivazioni storiche: fluidodinamica

In questa sezione spiego perchè i problemi in fluidodinamica sfociano naturalmente nella ricerca di un operatore di proiezione su un insieme di mappe che preservano la misura.

2.2.1 Le equazioni di Eulero

Le equazioni di Eulero sono tra le più vecchie e più semplici della fluidodinamica, andando indietro nel XVIII secolo, ma sono ancora ricche di misteri. Nella loro versione più semplice esse modellano un fluido incomprimibile, non viscoso in un aperto limitato $\Omega \subset \mathbb{R}^n$ ($n = 2$ o 3). L'incognita è il campo di velocità del fluido,

$v(t,x) : \mathbb{R}_+ \times \Omega \to \mathbb{R}^n$ (= spazio tangente a Ω), e la equazione di Eulero nel caso di un fluido incomprimibile è

$$\begin{cases} \frac{\partial v}{\partial t} + v \cdot \nabla v = -\nabla p, \\ \nabla \cdot v = 0. \end{cases} \tag{27}$$

Qui per definizione, $\nabla \cdot v = \sum_i (\partial v_i / \partial x_i)$ è la divergenza di v, e $v \cdot \nabla v = (v \cdot \nabla) v$ è il campo vettoriale di cui la i-esima componente è $(v \cdot \nabla) v_i = \sum_j^n v_j \frac{\partial v_i}{\partial x_j}$ All'equazione (27) deve essere aggiunta una condizione di bordo, che di solito è

$$v \text{ tangente al bordo di } \Omega.$$

. La condizione $\nabla \cdot v = 0$ significa che il fluido è incomprimibile. E' la condizione matematica affinchè il volume di un elemento di fluido non cambi durante la evoluzione temporale. Un vettore che soddisfa la condizione di incomprimibilità si dice divergence-free. L'incognita $p(t,x) : R_+ \times \Omega \to \mathbb{R}$ è la pressione, la cui evoluzione è irrilevante.

Il problema di Cauchy per l'equazione di Eulero si può formulare così: dato un certo campo di velocità v_0 su Ω tangente al bordo e divergence-free, che soddisfa certe condizioni che devono essere precisate, provare esistenza e unicità di una soluzione di (27) che soddisfa $v(0,\cdot) = v_0$.

Ovviamente questa formulazione non ha veramente senso fino a quando non si specifica che cosa significa soluzione e quali proprietà aspettarsi da quest'ultima. Il problema di Cauchy si ritiene che sia risolto in modo soddisfacente in dimensione 2 sotto la condizione che $\nabla \times v_0$ (rotore di v_0) sia limitato (Teorema di Youdovich). In dimensione $n = 3$ è un problema estremamente famoso. La seguente stima a priori(legge di conservazione dell'energia) suggerisce che uno spazio naturale dove cercare le soluzioni di un equazione di Eulero è $L^2(\Omega; \mathbb{R}^n)$.

Proposizione 4. *(Conservazione dell'energia per soluzioni liscie dell'equazione di Eulero) Sia v una soluzione liscia di* (27)*. Allora, l'energia cinetica totale $\int_\Omega |v(t,x)|^2 = \|v(t,\cdot)\|^2_{L^2(\Omega;\mathbb{R}^n)}$ è invariante rispetto al tempo.*

Cenno di dimostrazione. Prima di tutto

$$\frac{1}{2}\frac{d}{dt}\int_\Omega |v|^2 = \int_\Omega v \cdot \frac{\partial v}{\partial t} = -\int_\Omega v \cdot (v \cdot \nabla v) - \int_\Omega v \cdot \nabla p.$$

D'altra parte,

$$\begin{aligned} \int_\Omega v \cdot \nabla p &= -\int_\Omega (\nabla \cdot v) p \quad [\text{perchè } v \text{ è tangente al bordo di } \Omega] \\ &= 0 \quad [\text{perchè } \nabla \cdot v = 0], \end{aligned}$$

$$\int_\Omega v \cdot (v \cdot \nabla v) = \sum_{1 \le i,j \le n} \int_\Omega v_i v_j \partial_j (v_i) = \frac{1}{2} \sum_{ij} \int_\Omega v_j \partial_j (v_i^2) = \frac{1}{2} \int_\Omega v \cdot \nabla |v|^2 = -\frac{1}{2} \int_\Omega (\nabla \cdot v) |v|^2 = 0.$$

□

2.3 Formulazione Lagrangiana

Il punto di vista che è stato utilizzato fino ad ora è quella euleriano, in cui l'incognita è il campo di velocità. Esiste un approccio per descrivere lo stesso modello in fluidodinamica, che si focalizza sulla traiettoria delle particelle. Si ricorda che il punto di vista Lagrangiano fu introdotto da Eulero, mentre quello Euleriano deve essere attribuito a Bernoulli e D'alembert.

- Nel caso Euleriano il punto di vista è fisso su un certo punto dello spazio x, e in quella posizione si misura la velocità v ad un istante t.
- In quello Lagragiano, si identificano le varie particelle in considerazione, e si studia la traiettoria di ognuna di esse. Per esempio, decidendo di identificare una particella con la sua posizione iniziale x_0, si può scrivere

$$x = m(t, x_0)$$

la posizione al tempo t di una particella che era localizzata nella posizione x_0 al tempo 0. Si assume generalmente che la mappa $x_0 \mapsto m(t, x_0)$ definita su Ω, sia iniettiva. Per passare da una descrizione all'altra, si considera il sistema

$$\begin{cases} v(t, m(t, x_0)) = \frac{\partial m(t,x_0)}{\partial t}, \\ m(0, x_0) = x_0. \end{cases} \tag{28}$$

Se si differenzia la prima identità di (28) rispetto al tempo, si ottiene

$$\frac{\partial^2}{\partial t^2} m(t, x_0) = \left[\frac{\partial v}{\partial t} + (v \cdot \nabla v) \right] (t, m(t, x_0)) \tag{29}$$

che spiega perchè appare il termine $v \cdot \nabla v$ nell'equazione di Eulero. Si vede immediatamente che se due particelle provenienti da due posizioni differenti occupano la stessa posizione ad un tempo t (cioè $m(t, x_1) = m(t, x_2)$) non si possono più identificare le particelle con la loro posizione iniziale. Questa è una delle ragioni per cui viene imposta la iniettività di m. Inoltre, nel contesto dell'equazione di Eulero è naturale richiedere che la mappa m sia suriettiva: altrimenti, ci sarebbe dello spazio vuoto creato dentro il dominio, contraddicendo il fatto che il fluido ha una densità costante. Perciò sembra naturale cercare $(m(t, \cdot))_{t \geq 0}$ tra le famiglie di diffeomorfismi da Ω a Ω. La condizione di incomprimibilità può essere riproposta in termini di m: se v è abbastanza regolare (per esempio C^1), allora

$$\nabla \cdot v = 0 \Longleftrightarrow \det(\frac{\partial m}{\partial x_0}) = 1. \tag{30}$$

Ovviamente l'identità a destra è soddisfatta al tempo 0, dal momento che $m(0, \cdot)$ è la mappa identità ($m(0, x_0) = x_0$) allora (30) è una conseguenza della identità

$$\frac{\partial}{\partial t} \log \det \left[\frac{\partial m}{\partial x_0} \right] = (\nabla \cdot v)(t, m(t, x_0)). \tag{31}$$

La prova di (31) si deriva dalla formula

$$\mathrm{D} det\big|_A \cdot H = (\det A) tr(A^{-1} H) = tr((comA)^T H).$$

La notazione $comA$ indica la matrice dei cofattori di A, e la seconda espressione ha senso se e solo se A è invertibile. Ricapitolo le considerazioni informali dette fino ad ora: nella formulazione lagrangiana, l'equazione di Eulero diventa una equazione di evoluzione per una mappa $t \mapsto m(t, \cdot)$, con valori nel gruppo $G(\Omega)$ dei diffeomorfismi $\Omega \to \Omega$ con determinante unitario. Per riassumere il fatto che m sia appartenente a $G(\Omega)$ la chiamerò g. In particolare g preserva la misura: essa trasporta la misura di Lebesgue (ristretta ad Ω) su se stessa. L'interpretazione fisica è che il volume di un insieme di particelle è mantenuto costante durante la evoluzione temporale, che è precisamente la incompressibilità. Perciò, (28) diventa

$$v(t, g(t, x_0)) = \frac{d}{dt} g(t, x_0), \quad \text{cioè } v = \frac{\partial g}{\partial t} \circ g^{-1}.$$

Da (29), l'equazione di Eulero si traduce in una equazione per il campo $t \mapsto g(t, \cdot)$ da $\mathbb{R}_+$ in $G(\Omega)$,

$$\frac{d^2}{dt^2} g(t, x_0) = -\nabla p(t, g(t, x_0)) \tag{32}$$

2.3.1 Interpretazione di Arnold

In questa sezione spiego l'interpretazione formale dell'equazione di Eulero che fu sviluppata da Arnold e dai suoi collaboratori: l'idea è che l'equazione di Eulero è l'equazione delle geodetiche in $G(\Omega)$, dotato della struttura Riemanniana ereditata dallo spazio euclideo $L^2(\Omega)$.

Osservazione 9. *L'unico motivo per cui si utilizza $L^2(\Omega; \mathbb{R}^n)$ invece di $L^2(\Omega; \Omega)$ è che l'ultimo non è uno spazio vettoriale.* $\square$

Ora spiego l'interpretazione di Arnold. La seguente discussione è meramente formale. Si ricordi che una geodetica su uno spazio riemanniano M è una curva $\gamma(t)$ che minimizza la lunghezza

$$\sqrt{\int_{t_1}^{t_2} |g'(t)|^2 dt} \tag{33}$$

fra tutte le curve $g : [t_1, t_2] \to M$ con la condizione al bordo $g(t_1) = \gamma(t_1)$, $g(t_2) = \gamma(t_2)$, e questa proprietà di minimizzazione dovrebbe valere fino a quando t_2 è abbastanza vicino a t_1. E' equivalente dire che l'accellerazione della curva γ ha proiezione nulla sullo spazio tangente alla curva stessa. Nella interpretazione di Arnold, si considera la struttura Riemanniana su $G(\Omega)$ ereditata da $L^2(\Omega)$, e si prova che questo significa che l'accellerazione d^2g/dt^2 dovrebbe essere ortogonale allo spazio tangente $T_{g(t)}G(\Omega)$ in $L^2(\Omega; \mathbb{R}^n)$. Ora determino lo spazio tangente. Richiamo dalle riflessioni già fatte il fatto che una curva $g(t)$, che parte da $g_0 \in G$, sta in G se e solo se $\partial g/\partial t$ è tangente al bordo, e $\nabla \cdot \left[\frac{\partial g}{\partial t} \circ g^{-1}\right] = 0$. Perciò, i vettori tangenti in T_gG sono tutti i campi vettoriali h tali per cui $\nabla \cdot (h \circ g^{-1}) = 0$, o equivalentemente $h = w_0 \circ g$, con $w_0 \in D_0$, lo spazio dei campi vettoriali divergence-free. Usando il fatto che g è un diffeomorfismo che preserva la misura, si nota subito che $(T_gG)^\perp$ è lo spazio dei campi vettoriali $q_0 \circ g$, dove $q_0 \in D_0^\perp$, e $D_0^\perp$ è il sottospazio ortogonale a D_0 in L^2. E' una facile conseguenza della decomposizione di Helmoltz che, sotto una sufficiente regolarità di Ω,

$$D_0^\perp = \{-\nabla p \mid p : \Omega \to \mathbb{R}\}$$

Quindi l'equazione per le geodetiche diventa

$$\frac{d^2}{dt^2} g(t) = -\nabla p(t, g(t))$$

Questa è esattamente l'equazione di Eulero (32), nella formulazione lagrangiana.
Si noti che in questo quadro, l'integrale che appare in (33) si interpreta come l'azione della traiettoria $g(t, x)$, cioè l'integrale sul tempo della energia cinetica (a meno di un fattore 1/2), che è in funzione del campo di velocità.
Queste considerazioni formali danno una interpretazione del campo della pressione. Infatti, si potrebbe ridurre il problema a trovare una stima abbastanza regolare della pressione. Per esempio si può mostrare che, se $g(t, x_0)$ è una soluzione derivabile dell'equazione di Eulero, con un campo vettoriale limitato in $C^2(\Omega)$, uniformemente nel tempo, allora esiste un $\epsilon > 0$ tale per cui se $|t_1 - t_2| < \epsilon$,

$$\int_{t_1}^{t_2} \left(\int_\Omega |\frac{\partial g}{\partial t}(t, x_0)|^2 dx_0 \right) dt \leq \int_{t_1}^{t_2} \left(\int_\Omega |\frac{\partial \gamma}{\partial t}(t, x_0)|^2 dx_0 \right) dt$$

per un'altra curva γ con $\gamma(t_1) = g(t_1)$ e $\gamma(t_2) = g(t_2)$. Se Ω è convesso, si può scegliere $\epsilon = \pi/\sqrt{\|D^2 p\|_{L^\infty}}$, si veda [2].
D'altra parte, tali stime di regolarità costituiscono un problema aperto di notevole importanza. Se si vuole provare qualche risultato sulle soluzioni dell'equazione di Eulero con il formalismo appena descritto, è impossibile a priori applicare gli strumenti generali della geometria Riemanniana, perchè G è infinito dimensionale, e in qualche modo molto singolare.
Ora enuncio un altro risultato noto: se g_0 e g_1 sono due funzioni derivabili di $G(\Omega)$, sufficientemente vicine in in una norma di Sobolev di ordine abbastanza alto, allora esiste almeno una traiettoria che sia una geodetica in $G(\Omega)$ che connette g_0 e g_1. Ma in alcune circostanze le geodetiche possono non essere uniche (cosa abbastanza comune), o semplicemente non esistere. Shnirelman [11] è riuscito a costruire un diffeomorfismo $h \in G([0, 1]^3)$ tale per cui non c'è una geodetica che connette la identità con h; osserviamo inoltre che tale diffeomorfismo è bidimensionale (nel senso del successivo teorema 5).
Citando altri risultati, Shnirelman ha provato che quando Ω è un dominio tridimensionale allora il diametro di $G(\Omega)$ è finito [11, 10], mentre questo è falso per un dominio bidimensionale. Questa è un'altra prova del fatto che i flussi bidimensionali hanno maggiori restrizioni rispetto a di quelli tridimensionali.

2.3.2 Il problema delle geodetiche approssimate

Per le motivazioni già enunciate Brenier iniziò a studiare nozioni generalizzate delle geodetiche. In particolare, sviluppò la nozione di geodetiche approssimate su $G(\Omega)$: il problema è di trovare $g_{1/2} \in G(\Omega)$ tale che sia il minimo di

$$A(g) = \|g - g_0\|_{L^2}^2 + \|g_1 - g\|_{L^2}^2 = \frac{1}{2}\|g_0 - g_1\|_{L^2}^2 + 2\|g - \frac{g_0 + g_1}{2}\|_{L^2}^2$$

Per risolvere questo problema $g_{1/2}$ dovrebbe essere esattamente la poiezione di $h = (g_0 + g_1)/2$ su $G(\Omega)$. Anche se questa costruzione è molto naive può essere rifinita aggiungiendo più punti intermedi, alla fine ciò conduce alla costruzione di geodetiche approssimate su $G(\Omega)$. Ma come garantire che la proiezione da L^2 di un certo $h \in L^2$ su $G(\Omega)$ sia ben definita? Un teorema molto conosciuto afferma che una proiezione da L^2 in un suo sottoinsieme chiuso e convesso è ben definita. Ma si ricordi che $G(\Omega)$ non è convesso; e non è nemmeno chiuso. In realtà, si può dimostrare quanto segue: qualsiasi sia Ω sottoinsieme con bordo liscio di $\mathbb{R}^n$, $n \geq 2$, ogni mappa che preserva la misura su Ω può essere approssimata, in norma L^2, da una successione di diffeomorfismi con jacobiano unitario. Per esempio anche una mappa s con $\det(\nabla s) = -1$ può essere approssimata con elementi in $G(\Omega)$. Questi risultati si possono trovare ad esempio in Roesch [8]. Per continuare con questa idea è perciò naturale rilassare il problema della proiezione, e sostituire a $G(\Omega)$ la sua chiusura $S(\Omega)$. Allora il teorema di fattorizzazione polare di Brenier dà delle condizioni sufficienti affinchè la proiezione su $S(\Omega)$ sia ben definita. Questa è la motivazione storica del teorema di fattorizzazione polare! In [1] Brenier discute la possibilità di costruire geodetiche in senso generalizzato, non imponendo che le particelle debbano seguire tutte la stessa traiettoria; nello stesso modo in cui nella formulazione del problema del trasporto ottimale di Kantorovich, la massa trasportata può essere separata lungo il percorso. Ovviamente, sotto questa generalizzazione la proprietà della densità delle particelle di essere costante nello spazio e nel tempo(omogeneità) non è più rispettata, e l'equazione $\nabla \cdot v = 0$ deve essere sostituita da

$$\frac{\partial \rho}{\partial t} + \nabla_x \cdot (\rho v) = 0 \tag{34}$$

dove $\rho(t, x)$ è la densità delle particelle; similmente l'equazione di Eulero deve essere sostituita con

$$\frac{\partial \rho}{\partial t} + \nabla_x \cdot (\rho v \otimes v) + \rho \nabla_x p = 0. \tag{35}$$

Quindi gli oggetti costruiti da Brenier sono delle funzioni generalizzate di densità della forma $\rho(dx_0, t, x)$. Il senso di questa ρ è che "conta" la quantità delle particelle inizialmente nell'elemento di volume dx_0 attorno a x_0, che occupano la posizione x al tempo t. In questo formalismo, una generalizzata soluzione ρ si riduce ad una classica se ρ è una delta di Dirac rispetto a x_0(come il problema di Kantorovich si riduce al problema di Monge se il piano di trasporto è una massa di Dirac nel suo argomento y, formula (20)). Ora, per definire meglio ciò di cui ho scritto, cito il risultato principale di [1].

Teorema 5. *[Esistenza di geodetiche generalizzate in $S(\Omega)$] Sia $\Omega = [0, 1]^3$ e sia $T > 0$. Sia $h \in S(\Omega)$ una mappa che preserva la misura e supponiamo che h sia bidimensionale, nel senso che*

$$\forall x \in [0, 1]^3, h(x_1, x_2, x_3) = (H(x_1, x_2), x_3).$$

Allora, esiste una misura ρ dipendente dal tempo (che intuitivamente rappresenta la densità di particelle), m(che fisicamente è la densità dei momenti) definite sullo spazio di misura esteso $\Omega \times ([0, T] \times \Omega)$, che soddisfano le seguenti proprietà. Prima di tutto, $m = v\rho$, dove v è una densità in L^2 a valori in $\mathbb{R}^d$, che si pensa come una velocità locale di particelle. Inoltre, $\rho(dx_0; t, x)$ è una misura di probabilità rispetto a x_0, per quasi tutti i t e gli x e soddisfa le equazioni di Eulero generalizzate (34)-(35) *nel senso delle distribuzioni:*

$$\rho(dx_0; 0, x) = \delta_{[x_0=0]}, \quad \rho(dx_0; T, x) = \delta_{[h(x_0)=x]}.$$

Infine, la azione totale (generalizzata)

$$A = \int_0^T \left[\int_{\Omega \times \Omega} \frac{|v(x_0; t, x)|^2}{2} \rho(dx_0; t, dx) \right] dt$$

è minimale nel seguente senso: per tutti gli $\eta > 0$ esiste $\epsilon > 0$ tale che, se $g(t, x)$ è una soluzione dell'equazione di Eulero (32) *che soddisfa $\|g(T, \cdot) - h\|_{L_2} \leq \epsilon$, allora*

$$A \leq \left(\int_0^T \frac{1}{2} \int_\Omega |u(t, x)|^2 dt \right) + \eta$$

dove u è il campo di velocità associato a g.

Si noti che la scelta di una mappa bidimensionale che preserva la misura non è un assunzione che permette di semplificare molto; infatti, dalla precedente discussione sul lavoro di Shnirelman si capisce che è uno dei casi peggiori per i problemi riguardanti le geodetiche.

Bibliografia

[1] Brenier, Y. Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. *Comm. Pure Appl. Math. 52*, 4 (1999), 411-452.

[2] Brenier, Y. On the motion of an ideal incompressible fluid. In *Partial differential equations of elliptic type (Cortona, 1992)* Cambridge Univ. Press, Cambridge,1994, pp. 123-148.

[3] Brenier, Y.*Décomposition polaire et réarrangement des champs polaire de vecteurs.* R. Acad. Sci. Paris Sér. I Math., Vol. 305, No. 19. (1987) pp. 805-808.

[4] Burton, G.R., e Douglas, R.J. Rearrangements and polar factorisation of countably degenerate functions. *Proc. Roy. Soc. Edinburgh Sect. A 128*, 4 (1998), 671-681.

[5] Ebin, D.G., e Marsden, J. Groups of diffeomorphisms and the notion of an incompressible fluid. *Ann. of Math. (2) 92* (1970), 102-163.

[6] Monge, G. *Memoir sur la théorie de déblais et des remblais*, De l'Imprimerie Royale, Paris, 1781.

[7] Rockafellar, T. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1997. Reprint of the original 1970 paperbacks, Princeton Paperbacks.

[8] Roesch, M. *Contribution a la description geometrique des fluides parfait incompressibles.* Phd thesis, Univ. Paris VI,1994.

[9] Santambrogio, F. *Optimal transport for applied mathematicians.* Haim Brezis Université Pierre et Marie Curie, Paris, and Rutgers University, New Brunswick, NJ, USA, 2015. Progress in Nonlinear Differential Equations and Their Applications, volume 87.

[10] Shnirel'man, A.I. Attainable diffeomorphisms. *Geom. Funct. Anal. 3*, 3 (1993), 279-294.

[11] Shnirel'man, A.I. The geometry of the group of the diffeomorfisms and the dynamics of an incompressible fluid. *Mat. Sb. (N.S.) 128(170)*, 1 (1985), 82-109,144. English translation in *Math. USSR - Sb. 56(1987)*, 79-105.

[12] Villani, C. *Optimal transport: Old and new* Springer, New York, 2008.

[13] Villani, C. *Topics in optimal transportation.* American Mathematical Society Providence, Rhode Island, 2003. Graduate studies in mathematics, volume 58.