



UNIVERSITÀ DEGLI STUDI DI PISA
Corso di Laurea Triennale in Matematica

TESI DI LAUREA TRIENNALE
**IL TRASPORTO OTTIMO E
LA DISUGUAGLIANZA ISOPERIMETRICA**

RELATORE
Prof. Luigi AMBROSIO

CANDIDATO
Giacomo COLOMBO

ANNO ACCADEMICO 2019-2020

Indice

1	Risultati preliminari	7
1.1	Richiami di analisi funzionale	7
1.2	Richiami di teoria della misura	8
2	Il problema del trasporto ottimo	11
2.1	Il problema di Monge-Kantorovich	12
2.1.1	Esistenza dei piani ottimali	16
2.2	Il problema di Monge con costo quadratico	18
2.2.1	Mappe e piani di trasporto	18
2.2.2	Proprietà di regolarità dei piani ottimali	20
2.2.3	Il Teorema di Brenier	26
3	Il problema isoperimetrico	29
3.1	Stabilità L^p delle mappe ottimali	30
3.2	Proprietà di regolarità delle mappe ottimali	33
3.2.1	L'equazione di Monge-Ampère	35
3.2.2	Regolarità BV delle mappe ottimali	39
3.3	La disuguaglianza isoperimetrica	43
3.3.1	Cenni alle disuguaglianze isoperimetriche quantitative	46
	Bibliografia	49

Capitolo 1

Risultati preliminari

1.1 Richiami di analisi funzionale

Ricordiamo in questa sezione le definizioni di duale di uno spazio di Banach e di convergenza debole*, e richiamiamo i relativi risultati essenziali di sequenziale compattezza.

Definizione 1 (Duale topologico). Sia E uno spazio di Banach, si definisce E^* lo spazio dei funzionali lineari continui su E , ovvero

$$E^* := \{ \phi: E \rightarrow \mathbb{R} \mid \|\phi\|_{E^*} < \infty \}$$

dove si pone

$$\|\phi\|_{E^*} := \sup \{ |\phi(x)| \mid x \in E, \|x\| \leq 1 \}.$$

Si può mostrare che la norma $\|\cdot\|_{E^*}$ rende E^* uno spazio di Banach.

Definizione 2 (Convergenza debole*). Sia E uno spazio di Banach, e E^* il suo duale topologico. Si dice che una successione $(\phi_n) \subset E^*$ converge debole* a $\phi \in E^*$ (e scriveremo $\phi_n \rightharpoonup \phi$) se ci converge puntualmente, ovvero se

$$\phi_k(x) \rightarrow \phi(x) \quad \forall x \in E.$$

Per la dimostrazione del seguente classico teorema di analisi funzionale rimandiamo a [B].

Teorema 1 (Banach-Alaoglu). *Sia E uno spazio di Banach separabile. Allora data $(\phi_n) \subset B_{E^*}$ esistono una sottosuccessione ϕ_{n_k} e un elemento $\phi \in B_{E^*}$ tali che*

$$\phi_{n_k} \rightharpoonup \phi.$$

1.2 Richiami di teoria della misura

In questo paragrafo richiamiamo le definizioni e le proprietà fondamentali delle misure definite su uno spazio metrico che serviranno nei capitoli successivi.

Dato uno spazio metrico (X, d) , indichiamo con $\mathcal{B}(X)$ la σ -algebra dei boreliani su X . Dati due spazi metrici $(X, d_X), (Y, d_Y)$, una funzione $f: X \rightarrow Y$ tale che $f^{-1}(A) \in \mathcal{B}(X)$ per ogni $A \in \mathcal{B}(Y)$ è detta boreliana.

Una misura di Borel su X è una funzione $\mu: \mathcal{B}(X) \rightarrow [0, +\infty]$ σ -additiva tale che $\mu(\emptyset) = 0$. Dato che le misure considerate saranno sempre di Borel, chiameremo queste ultime semplicemente misure. Una misura tale che $\mu(X) < \infty$ è detta finita. Una misura tale che $\mu(X) = 1$ è una probabilità. Indichiamo con $\mathcal{P}(X)$ l'insieme delle probabilità su X .

Introduciamo alcune definizioni basilari delle misure che saranno utili.

Definizione 3 (Restrizione e push-forward). Sia μ una misura su X . Dato $E \in \mathcal{B}(X)$ la restrizione di μ ad E è la misura definita da

$$\mu \llcorner E(A) := \mu(E \cap A).$$

Dati uno spazio metrico Y e una funzione $f: X \rightarrow Y$ boreliana, il push-forward di μ tramite f è la misura definita da

$$f_{\#}\mu(E) := \mu(f^{-1}(E)).$$

La seguente utile proposizione permette di calcolare un integrale rispetto alla misura push-forward.

Proposizione 1 (Cambio di variabile). Siano X, Y spazi metrici e $f: X \rightarrow Y$, $\phi: Y \rightarrow \mathbb{R}$ funzioni boreliane, e sia μ una misura. Allora ϕ è integrabile rispetto a $f_{\#}\mu$ se e solo se $\phi \circ f$ lo è rispetto a μ e in tal caso vale

$$\int_Y \phi \, d f_{\#}\mu = \int_X \phi \circ f \, d\mu. \quad (1.1)$$

Dimostrazione. A meno di suddividere ϕ in parte positiva e parte negativa è sufficiente dimostrarlo per ϕ non negativa. Per ϕ semplice è la definizione; per linearità vale anche per funzioni semplici, e nel caso generale segue dal fatto che ϕ è limite di una successione crescente di funzioni semplici. \square

Definizione 4 (Supporto e μ -misurabilità). Sia μ una misura su X . Il suo supporto è il chiuso definito da

$$\text{supp } \mu := \{ x \in X \mid \mu(U) > 0 \quad \forall U \text{ intorno aperto di } x \}.$$

Diremo che μ è concentrata su un boreliano $B \subset X$ se $\mu(X \setminus B) = 0$. Se ν è una misura su X tale che μ è concentrata su B e ν su $X \setminus B$ si dice che μ e ν sono singolari e si indica con $\mu \perp \nu$.

Data una misura μ , diremo che un insieme $A \subset X$ è μ -misurabile se esistono due boreliani E_1, E_2 tali che $E_1 \subset A \subset E_2$ e $\mu(E_2 \setminus E_1) = 0$.

Diciamo quindi che μ è concentrata su un insieme A qualunque se esiste un boreliano $E \subset A$ tale che $\mu(X \setminus E) = 0$. Si noti che questo implica che A è μ -misurabile. Inoltre, μ è concentrata sul suo supporto e $\mu \perp E$ è concentrata su E , e se μ è concentrata su E allora $\mu = \mu \perp E$.

Introduciamo le misure di Radon. La loro importanza segue dalla possibilità di definire una nozione di convergenza debole in dualità con le funzioni continue che risulta coincidere con la convergenza debole* del duale degli spazi di Banach, ereditandone le proprietà di compattezza.

Definizione 5 (Misure di Radon). Dato uno spazio metrico (X, d) , una misura di Radon su X è una misura $\mu: \mathcal{B}(X) \rightarrow [0, \infty]$ tale che $\mu(K) < \infty$ per ogni $K \subset X$ compatto. Una misura di Radon con segno è $\mu = \mu^+ - \mu^-$, dove μ^+, μ^- sono misure finite.

Una misura di Radon a variazione localmente finita è una funzione d'insieme $\mu: \{B \in \mathcal{B}(K) | K \subset X \text{ compatto}\} \rightarrow \mathbb{R}$ tale che per ogni compatto $K \subset X$ la restrizione $\mu|_{\mathcal{B}(K)}$ è una misura di Radon con segno.

La massa di una misura di Radon con segno μ è data da

$$\mathbb{M}(\mu) := \sup \left\{ \sum_{i \in \mathbb{N}} |\mu(E_i)|, \{E_i\}_{i \in \mathbb{N}} \subset \mathcal{B}(X), \{E_i\} \text{ partizione di } X \right\}.$$

Si può mostrare (vedere ad esempio [AFP]) che lo spazio delle misure di Radon con segno è uno spazio vettoriale, che indichiamo con $\mathcal{M}(X)$. Indichiamo inoltre con $\mathcal{M}_+(X)$ lo spazio delle misure di Radon su X (si noti che in generale $\mathcal{M}_+(X) \not\subset \mathcal{M}(X)$). Se $\mu \in \mathcal{M}(X)$ allora si può mostrare che esiste un'unica decomposizione $\mu = \mu^+ - \mu^-$ con $\mu^+ \perp \mu^-$ misure positive finite, e in tal caso si ha $\mathbb{M}(\mu) = \mu^+(X) + \mu^-(X)$. Infine, si può mostrare che la massa è una norma su $\mathcal{M}(X)$ che lo rende uno spazio di Banach.

Definizione 6. Dato uno spazio (X, d) metrico compatto, denotiamo con $\mathcal{C}(X)$ lo spazio delle funzioni continue a valori reali, dotato della norma

$$\|f\|_\infty := \sup_{x \in X} |f(x)|.$$

Il seguente risultato è ben noto (vedere ad esempio [Ru]).

Proposizione 2. *Sia (X, d) uno spazio metrico compatto. Allora lo spazio $\mathcal{C}(X)$ dotato della norma $\|\cdot\|_\infty$ è uno spazio di Banach separabile.*

Introduciamo ora la convergenza debole tra misure.

Definizione 7 (Convergenza debole). Sia (X, d) uno spazio metrico compatto, e sia $(\mu_n) \subset \mathcal{M}(X)$ una successione di misure su X . Diciamo che μ_n converge debolmente a $\mu \in \mathcal{M}(X)$ (e scriveremo $\mu_n \rightharpoonup \mu$) se

$$\lim \int_X \phi \, d\mu_n = \int_X \phi \, d\mu \quad \forall \phi \in \mathcal{C}(X).$$

Il seguente importante Teorema permette di estendere le buone proprietà di compattezza della convergenza debole* alla convergenza debole delle misure di Radon. Alla luce di questo Teorema sarebbe più corretto usare il termine convergenza debole* anche nella Definizione 7, ma l'uso di debole è ben consolidato in letteratura. Per una dimostrazione si veda ad esempio [EG].

Teorema 2 (Riesz). *Sia (X, d) uno spazio metrico compatto, allora $(\mathcal{M}(X), \mathbb{M})$ è isometrico al duale topologico di $\mathcal{C}(X)$ tramite l'isometria suriettiva*

$$\mathcal{M}(X) \ni \mu \mapsto \left(f \mapsto \int_X f \, d\mu \right).$$

Il seguente Corollario enuncia le proprietà di compattezza della convergenza debole delle misure di Radon nella forma che sarà utilizzata nei capitoli seguenti.

Corollario 1. *Sia X uno spazio metrico compatto, e sia $(\mu_n) \subset \mathcal{P}(X)$ una successione di probabilità. Allora esistono una sottosuccessione μ_{n_k} e un elemento $\mu \in \mathcal{P}(X)$ tali che*

$$\mu_{n_k} \rightharpoonup \mu.$$

Capitolo 2

Il problema del trasporto ottimo

Nel 1781 il francese G. Monge pubblicò delle memorie intitolate "Memoires sur la théorie des déblais et des remblais", in cui si propone di trovare il modo migliore per trasportare le risorse dai siti in cui si estraggono ai posti in cui vengono utilizzate, minimizzando il costo del trasporto, ovvero minimizzando la distanza totale. Pur rientrando nell'ambito del calcolo delle variazioni, questo problema non è affrontabile con metodi classici; questa difficoltà fece sì che il problema del trasporto ottimo non venne affrontato che negli anni '40 dal premio Nobel per l'economia L. Kantorovich, che era interessato al modo migliore di distribuire beni dai luoghi di produzione ai luoghi di vendita.

A questi lavori seguirono degli studi negli anni '80 ad opera di matematici come Y. Brenier, S. T. Rachev e L. Rüschendorf. La teoria del trasporto ottimo ha ora numerose applicazioni a molteplici branche della matematica, pura e applicata, come ad esempio alle PDE, alla geometria Riemanniana e a problemi di calcolo delle variazioni.

Nel primo paragrafo di questo capitolo enunciamo il problema generale in \mathbb{R}^n e ne presentiamo alcuni aspetti fondamentali, in particolare presentiamo alcuni controesempi all'esistenza e unicità e un primo risultato di esistenza al problema di Kantorovich.

Nel secondo paragrafo ci concentriamo sul caso con costo proporzionale al quadrato della distanza; anche se questo semplifica la trattazione non toglie molto alla sua importanza: in questo caso è possibile infatti mostrare il Teorema di Brenier. Questo risultato, essenzialmente di esistenza e unicità, ha acquisito notevole importanza in larga parte grazie alle numerose applicazioni che ha in molti dei problemi sopra citati.

2.1 Il problema di Monge-Kantorovich

Una possibile formulazione moderna del problema di Monge in \mathbb{R}^n è la seguente.

(M) Siano assegnati uno spazio $X \subset \mathbb{R}^n$, due distribuzioni di probabilità $\mu, \nu \in \mathcal{P}(X)$ e una funzione costo boreliana $c: X \times X \rightarrow [0, +\infty]$. Indichiamo con $\Psi(\mu, \nu)$ la famiglia di mappe boreliane $T: X \rightarrow X$ tali che $T_{\#}\mu = \nu$. Il problema allora è di determinare

$$\min_{T \in \Psi(\mu, \nu)} \int_X c(x, T(x)) d\mu(x). \quad (\mathbf{M})$$

Le mappe $T \in \Psi(\mu, \nu)$ sono dette di mappe di trasporto; indicheremo con

$$\mathcal{C}(T) := \int_X c(x, T(x)) d\mu(x)$$

il costo di una mappa di trasporto, e con

$$\mathcal{C}(\mu, \nu) = \inf \{ \mathcal{C}(T) \mid T \in \Psi(\mu, \nu) \}$$

l'estremo inferiore considerato nel problema (M); se è chiaro dal contesto a quali misure ci si riferisce si scriverà semplicemente $\mathcal{C} = \mathcal{C}(\mu, \nu)$. Una mappa che realizza tale minimo è detta mappa ottimale.

In generale, il problema di Monge non ammette esistenza e unicità. L'esistenza può non essere garantita, oltre a ragioni dovute alla mancanza di regolarità del costo, a causa della condizione al bordo data dal vincolo sul push-forward. Questa condizione può dar luogo a fenomeni ben diversi rispetto al caso classico ed è ciò che rende il problema non affrontabile con i metodi classici del calcolo delle variazioni.

L'esempio seguente mostra che può verificarsi che non esistano mappe di trasporto tra due misure.

Esempio 1. Sia $X = \mathbb{R}^n$, e consideriamo $\nu = \mathcal{L}^n \llcorner [0, 1]^n$ e $\mu = \delta_0$. Allora non esistono mappe ammissibili, perché data una qualsiasi $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ boreliana si ha $f_{\#}\mu = \delta_{f(0)}$ che non è priva di atomi, e quindi si avrà sempre $f_{\#}\mu \neq \nu$.

Nell'esempio appena dato la misura di partenza μ è atomica; si può mostrare (vedere ad esempio [O]) che se μ è diffusa, i.e. priva di atomi, allora esistono sempre mappe di trasporto tra μ e ν .

L'esempio seguente mostra come il minimo possa non esistere anche se ci sono mappe di trasporto.

Esempio 2. Siano $X = \mathbb{R}^2$ e consideriamo la funzione costo $c(x, y) := |x - y|^2$. Posto inoltre $I := [0, 1]$, siano $\mu := \mathcal{H}^1 \llcorner \{0\} \times I, \nu := \frac{1}{2} \mathcal{H}^1 \llcorner \{-1, 1\} \times I$, dove \mathcal{H}^1 indica la misura di Hausdorff 1-dimensionale; in questo caso la si può definire tramite $\mathcal{H}^1 \llcorner (\{a\} \times I) := (f_a)_\# \mathcal{L}^1$, dove $f_a(t) := (a, t), t \in I$.

In questo caso risulta $\mathcal{C} = 1$; infatti, poiché per ogni mappa di trasporto T si ha $|T(x) - x|^2 \geq 1$ per μ -quasi ogni $x \in \mathbb{R}^2$ si ha $\mathcal{C}(T) \geq 1$. Per mostrare l'uguaglianza, si fissi un intero $N \geq 1$ e si divida l'intervallo $\{0\} \times I$ in $2N$ segmenti uguali, mentre si suddividano gli intervalli $\{\pm 1\} \times I$ in N segmenti uguali ciascuno. Si possono allora costruire delle mappe che mandano in modo lineare il $(2i - 1)$ -esimo segmento di $\{0\} \times I$ nel i -esimo segmento di $\{1\} \times I$ e il $2i$ -esimo segmento di $\{0\} \times I$ nel i -esimo segmento di $\{-1\} \times I$. È immediato verificare che in questo modo il costo relativo alle funzioni f_N così costruite è $\mathcal{C}(f_N) = 1 + O(\frac{1}{N})$, e quindi $\mathcal{C} = 1$.

Per mostrare che non ci sono mappe ottimali basta notare che si avrebbe

$$\int_{\mathbb{R}^2} ||x - T(x)|^2 - 1| d\mu(x) = 0$$

ma essendo sempre $|x - T(x)| \geq 1$ si ha che deve essere $|T(x) - 1| = 1$ μ -quasi ovunque, e questo significa che per μ -quasi ogni punto di $\text{supp } \mu$ deve essere $T(x) = x \pm (1, 0)$. Indicando con A_\pm gli insiemi in cui avviene ciascuna delle due possibilità, avremmo quindi che $T_{\#}\mu = \mathcal{H}^1 \llcorner (\{-1\} \times A_-) + \mathcal{H}^1 \llcorner (\{+1\} \times A_+) \neq \nu$.

Come nel caso dell'esistenza, la mancanza di unicità può verificarsi sia per motivi legati alla funzione costo sia per motivi legati al supporto delle misure. Prima di dare degli esempi occorre avere un criterio di ottimalità delle mappe: l'Osservazione seguente dà questo criterio in un caso particolare.

Osservazione 1 (Criterio di ottimalità). Sia $X \subset \mathbb{R}^n$ e consideriamo il problema di Monge con costo lineare $c(x, y) = |x - y|$. Allora vale

$$\mathcal{C}(\mu, \nu) \geq \sup \left\{ \int_X u d(\nu - \mu), u \in \text{Lip}_1(X) \right\}, \quad (2.1)$$

dove $u \in \text{Lip}_1(X)$ se e solo se $u: X \rightarrow \mathbb{R}$ è Lipschitziana con costante di Lipschitz non eccedente 1. Infatti, si ha

$$\begin{aligned} \int_X u d(\nu - \mu) &= \int_X u(y) d\nu(y) - \int_X u(x) d\mu(x) = \int_X (u \circ T(x) - u(x)) d\mu(x) \\ &\leq \int_X |T(x) - x| d\mu(x) = \mathcal{C}(T) \end{aligned}$$

per ogni $T \in \Psi(\mu, \nu)$.

Il seguente esempio mostra che con il costo lineare, a causa della mancanza di stretta convessità, l'unicità non è garantita.

Esempio 3. Sia $X = \mathbb{R}$, e sia $n \geq 1$ intero; consideriamo le misure di probabilità $\mu = \mathcal{L}^1 \llcorner [0, n]$, $\nu = \mathcal{L}^1 \llcorner [1, n+1]$ e la funzione costo lineare $c(x, y) = |x - y|$. Allora la mappa $T_1(s) = s + 1$ è ottimale in quanto, essendo $|T_1(s) - s| = 1$ per ogni s , si ha $\mathcal{C}(T_1) = n$ e scegliendo come funzione test Lipschitziana $u(s) = s$ nella (2.1) si ottiene che $\mathcal{C}(\mu, \nu) \geq \int_{\mathbb{R}} s d(\nu - \mu)(s) = \int_{[n, n+1]} s ds - \int_{[0, 1]} s ds = n$.

È possibile però costruire un'altra mappa ottimale; infatti considerando per $n \geq 2$

$$T_2(s) := \begin{cases} n + s & t \in [0, 1], \\ s & \text{altrimenti.} \end{cases}$$

e per $n = 1$

$$T_2(s) := 2 - s$$

si ha che $T_1 \neq T_2$ e $\mathcal{C}(T_2) = n$, pertanto anche T_2 è ottimale. Si noti come variando la costruzione appena mostrata si possano costruire infinite mappe ottimali distinte.

La mancanza di unicità può essere anche dovuta dalla strutture delle misure; l'esempio seguente illustra questo fenomeno.

Esempio 4. Siano $a, b \in \mathbb{R}^2$ punti distinti, $\nu = (\delta_a + \delta_b)/2$, e assumiamo che $\text{supp } \mu$ sia contenuto nell'asse r del segmento $[a, b]$ ¹. Allora se $c(x, y) = \varphi(|x - y|)$, con $\varphi: \mathbb{R} \rightarrow [0, \infty)$ boreliana, ogni mappa ammissibile è ottimale, infatti dato $x \in r$ vale $|x - b| = |x - a|$, pertanto

$$\mathcal{C}(T) = \int_X c(x, T(x)) d\mu(x) = \int_X \varphi(|x - a|) d\mu(x)$$

perché se $T \in \Psi$ risulta $T(x) \in \{a, b\}$ μ -quasi ovunque. Quindi, il costo $\mathcal{C}(T)$ non dipende dalla mappa T .

In vista di una dimostrazione dell'esistenza di mappe ottimali tramite il metodo diretto del calcolo delle variazioni, dagli esempi precedenti emerge che la classe delle mappe di trasporto non è la migliore in cui ambientare il problema. Seguendo la strada già tracciata in altre branche dell'analisi, il problema dell'esistenza del minimo del problema di Monge viene affrontato indebolendo la nozione di trasporto, ambientando il problema in una classe più ampia; si dimostra quindi l'esistenza nella classe più generale e poi si studiano condizioni di regolarità dei minimi per mostrare che inducono mappe di trasporto ottimali.

Nella teoria del trasporto ottimo la formulazione debole è il problema di Kantorovich (**K**), che formula il problema di Monge nella classe dei piani di trasporto.

¹ Con la notazione $[a, b]$ si intende l'insieme $\{\lambda a + (1 - \lambda)b, \text{ t.c. } \lambda \in [0, 1] \subset \mathbb{R}\}$

Definizione 8 (Piani di trasporto). Sia $X \subset \mathbb{R}^n$, e siano $\mu, \nu \in \mathcal{P}(X)$ due probabilità. Un piano di trasporto è una probabilità $\pi \in \mathcal{P}(X \times X)$ tale che

$$p_X \# \pi = \mu, \quad p_Y \# \mu = \nu, \quad (2.2)$$

dove p_X e p_Y sono la proiezione rispettivamente sulla prima e sulla seconda coordinata.

Date due probabilità μ, ν , si indicherà con $\Gamma(\mu, \nu)$ la famiglia dei piani di trasporto. Dati due boreliani $A, B \subset X$, la quantità $\pi(A \times B)$ rappresenta la massa che da A viene trasferita in B .

L'esempio seguente mostra la differenza tra piani e mappe di trasporto nel caso discreto.

Esempio 5 (Problema del Matching). Consideriamo due n -uple di punti distinti $(a_1, \dots, a_n), (b_1, \dots, b_n) \in \mathbb{R}^n$, e sia $X = \{a_i, i = 1, \dots, n\} \cup \{b_j, j = 1, \dots, n\}$. Consideriamo le misure $\mu = 1/n \sum_{i=1}^n \delta_{a_i}, \nu = 1/n \sum_{j=1}^n \delta_{b_j}$.

Una misura di probabilità su $X \times X$ è rappresentabile quindi da una matrice $\pi = (\pi_{ij})_{i,j=1}^n$ a entrate positive, tali che $\sum_{i,j=1}^n \pi_{ij} = 1$, dove π_{ij} rappresenta la massa che dal punto i è portata in j . Perché tale probabilità sia un piano di trasporto devono essere verificate le condizioni (2.2), che in questo caso sono

$$\sum_{j=1}^n \pi_{i,j} = \frac{1}{n} \quad i = 1, \dots, n; \quad \sum_{i=1}^n \pi_{i,j} = \frac{1}{n}, \quad j = 1, \dots, n.$$

Perché π sia indotto da una mappa di trasporto T (nel senso del Teorema 4) deve essere $\pi = (Id \times T) \# \mu$, quindi si deve avere $\pi_{ij} \in \{0, 1\}$. Perché π rispetti anche le condizioni sopra scritte, deve essere una matrice di permutazione. Si può quindi pensare che i piani di trasporto permettono di "disintegrare" le particelle di massa e distribuirle dove è meglio, mentre le mappe di trasporto spostano la massa senza poterla ulteriormente suddividere.

Il problema di Kantorovich può essere formulato come segue.

(K) Siano $X \subset \mathbb{R}^n$, e siano assegnate $\mu, \nu \in \mathcal{P}(X)$ due misure di probabilità e una funzione costo boreliana $c: X \times X \rightarrow [0, \infty]$. Il problema è di determinare

$$\min_{\pi \in \Gamma(\mu, \nu)} \int_{X \times X} c(x, y) d\pi(x, y). \quad (\mathbf{K})$$

Dato un piano $\pi \in \Gamma(\mu, \nu)$, si indicherà con

$$\mathcal{L}(\pi) = \int_{X \times X} c(x, y) d\pi(x, y)$$

il suo costo di trasporto, e con

$$\mathcal{C}_K(\mu, \nu) = \inf \{ \mathcal{C}(\pi) \mid \pi \in \Gamma(\mu, \nu) \}$$

l'estremo inferiore nel problema **(K)**. Un piano che realizza tale estremo inferiore è detto piano ottimale. Se dal contesto è chiaro se ci si riferisce all'estremo inferiore del problema di Monge piuttosto che del problema di Kantorovich si scriverà semplicemente \mathcal{C} , altrimenti si scriverà \mathcal{C}_M e \mathcal{C}_K per indicare l'estremo inferiore nel primo e nel secondo problema rispettivamente.

La formulazione di Kantorovich va intesa come un rilassamento del problema di Monge. Infatti, se $T: X \rightarrow X$ è una mappa boreliana di trasporto, allora la probabilità $\pi := (Id \times T)_{\#}\mu$ è tale che $\mathcal{C}(\pi) = \mathcal{C}(T)$ per la (1.1), pertanto si ha $\mathcal{C}_M(\mu, \nu) \geq \mathcal{C}_K(\mu, \nu)$. L'uguaglianza non è sempre valida semplicemente perché potrebbero non esistere mappe di trasporto, ma si può mostrare (vedere [P]) che sotto ipotesi minimali sulle misure e sul costo si ha anche l'uguaglianza $\mathcal{C}_M(\mu, \nu) = \mathcal{C}_K(\mu, \nu)$.

Il primo vantaggio della formulazione debole è che la famiglia di piani di trasporto non è mai vuota, perché anche senza assunzioni aggiuntive sulle misure si ha che $\mu \times \nu \in \Gamma(\mu, \nu)$, mentre l'esempio 1 mostra come questo non sia vero in generale per le mappe di trasporto. Inoltre, un'applicazione immediata del metodo diretto del calcolo delle variazioni mostra che il problema **(K)** ammette, sotto deboli assunzioni sul costo, l'esistenza di un minimo, risultato dimostrato nel Teorema 3.

2.1.1 Esistenza dei piani ottimali

In questa sezione si dimostrerà l'esistenza di piani ottimali nel problema **(K)**. La linea della dimostrazione utilizza il fatto che un funzionale semicontinuo inferiormente in uno spazio sequenzialmente compatto ammette minimo. Premettiamo due Lemmi.

Lemma 1. *Sia X uno spazio metrico compatto, e sia $(\mu_n) \subset \mathcal{P}(X)$ una successione tale che $\mu_n \rightarrow \mu \in \mathcal{P}(X)$. Allora se $\phi: X \rightarrow \mathbb{R}$ è semicontinua inferiormente si ha*

$$\int_X \phi \, d\mu \leq \liminf_n \int_X \phi \, d\mu_n, \quad (2.3)$$

e se $A \subset X$ è un aperto vale

$$\mu(A) \leq \liminf_n \mu_n(A). \quad (2.4)$$

Dimostrazione. La (2.4) segue dalla (2.3) prendendo $\phi = \mathbb{1}_A$, che essendo A aperto è semicontinua inferiormente.

Mostriamo quindi la (2.3). Preliminarmente mostriamo che esiste una successione di funzioni continue $c_k: X \times X \rightarrow [0, +\infty]$ che converge puntualmente crescendo a c . Consideriamo le funzioni

$$c_k(x, y) := \inf \{ c(r, s) + k(|x - r| + |y - s|) \mid (r, s) \in X \}.$$

Allora c_k è k -Lipschitziana perché estremo inferiore di funzioni k -Lipschitziane, inoltre dalla definizione segue $0 \leq c_k \leq c_{k+1} \leq c$.

Siano ora $x, y \in X$ fissati, e consideriamo $(r'_k, s'_k) \in X \times X$ tali che

$$c(r'_k, s'_k) + k(|x - r'_k| + |y - s'_k|) < c_k(x, y) + \frac{1}{k}.$$

Poiché senza perdita di generalità $c_k(x, y)$ è limitata deve essere che $r'_k \rightarrow x$ e $s'_k \rightarrow y$, pertanto

$$\lim_k c_k(x, y) \geq \liminf_k c(r'_k, s'_k) \geq c(x, y),$$

come voluto.

Sia quindi $(\mu_n) \subset \mathcal{P}(X)$ tale che $\mu_n \rightarrow \mu \in \mathcal{P}(X)$, e siano c_k come sopra. Poiché tali funzioni sono continue si ottiene

$$\liminf_n \int_X \phi \, d\mu_n \geq \lim_n \int_X c_k \, d\mu_n = \int_X c_k \, d\mu,$$

da cui, prendendo il limite su k , si ottiene per convergenza monotona

$$\liminf_n \int_X \phi \, d\mu_n \geq \int_X \phi \, d\mu. \quad \square$$

Lemma 2. *Sia X uno spazio metrico compatto, e sia $(\mu_n) \subset \mathcal{P}(X \times X)$ una successione di probabilità tale che $\mu_n \rightarrow \mu \in \mathcal{P}(X \times X)$. Allora*

$$(p_X)_\# \mu_n \rightarrow (p_X)_\# \mu \quad e \quad (p_Y)_\# \mu_n \rightarrow (p_Y)_\# \mu, \quad (2.5)$$

dove p_X e p_Y indicano la proiezione rispettivamente sulla prima e sulla seconda coordinata.

Dimostrazione. Se $f \in \mathcal{C}(X)$ allora la mappa $((x, y) \mapsto f(x)) \in \mathcal{C}(X \times X)$, pertanto per la (1.1) si ottiene

$$\begin{aligned} \lim_n \int_X f(x) \, d(p_X)_\# \mu_n(x) &= \lim_n \int_{X \times X} f(x) \, d\mu_n(x, y) \\ &= \int_{X \times X} f(x) \, d\mu(x, y) = \int_X f(x) \, d(p_X)_\# \mu(x), \end{aligned}$$

e analogamente per p_Y . □

Si può mostrare ora l'esistenza di piani ottimali.

Teorema 3 (Esistenza di piani ottimali). *Sia $X \subset \mathbb{R}^n$ compatto, se il costo $c: X \times X \rightarrow [0, +\infty]$ è semicontinuo inferiormente allora il problema (K) ammette minimo.*

Dimostrazione. Basta mostrare che, dotata della convergenza debole di misure, la classe $\Gamma(\mu, \nu)$ è sequenzialmente compatta, e che rispetto a questa convergenza il funzionale costo $\mathcal{C}(\pi)$ è semicontinuo inferiormente. Ma la semicontinuit  inferiore segue dal Lemma 1, mentre per il Corollario 1 si ha che $\mathcal{P}(X)$   sequenzialmente compatto, quindi   sufficiente mostrare che $\Gamma(\mu, \nu)$   chiuso in $\mathcal{P}(X)$, ovvero che se $(\pi_n) \subset \Gamma(\mu, \nu)$   tale che $\pi_n \rightharpoonup \gamma$ allora γ soddisfa le condizioni (2.2). Ma questo segue dal Lemma 2, perch  tali marginali sono costanti. \square

2.2 Il problema di Monge con costo quadratico

Il resto del capitolo   volto a dimostrare l'esistenza di una mappa ottimale sotto alcune ipotesi minimali sulle misure. Non lo si far  nel massimo della generalit  perch  va oltre i limiti di questa trattazione. Il problema che ci si propone di risolvere, a noi sufficiente,   il seguente.

(MQ) Siano assegnati uno spazio $X \subset \mathbb{R}^n$ compatto e convesso e due distribuzioni di probabilit  $\mu, \nu \in \mathcal{P}(X)$. Il problema allora   di determinare

$$\min_{T \in \Psi(\mu, \nu)} \int_X |T(x) - x|^2 d\mu(x). \quad (\text{MQ})$$

Stiamo quindi considerando il problema di Monge con costo quadratico in un sottoinsieme $X \subset \mathbb{R}^n$ convesso e compatto; queste assunzioni sono fatte con lo scopo di semplificare la trattazione e sono ben lungi dall'essere ottimali, ma la dimostrazione coinvolge molte delle idee fondamentali delle dimostrazioni dei casi pi  generali e il risultato principale ottenuto, il Teorema di Brenier,   comunque un teorema centrale della teoria del trasporto ottimale in quanto si presta a molte importanti applicazioni, una delle quali   presentata nel Capitolo 3.

2.2.1 Mappe e piani di trasporto

Se T   una mappa di trasporto tra μ e ν , la probabilit  $\pi_T := (Id \times T)_{\#}\mu$   un piano di trasporto indotto da T , perch  $(p_X)_{\#}\pi_T = (p_X \circ (Id \times T))_{\#}\mu = \mu$ e $(p_Y)_{\#}\pi_T = (p_Y \circ (Id \times T))_{\#}\mu = T_{\#}\mu = \nu$. Viceversa, per dedurre l'esistenza di una mappa di trasporto dalla regolarit  di un piano ottimale si   spinti a cercare

condizioni sufficienti per l'esistenza di una mappa di trasporto che induca il piano ottimale. La risposta è data dal seguente Teorema; ricordiamo che se $\gamma \in \mathcal{M}_+(X)$, $E \subset X$ è γ -misurabile se esistono due boreliani $A \subset E \subset B$ tali che $\gamma(B \setminus A) = 0$.

Teorema 4. *Sia $\pi \in \Gamma(\mu, \nu)$ un piano di trasporto. Allora esiste una mappa boreliana di trasporto $T: X \rightarrow X$ tale che*

$$\pi = (Id \times T)_{\#}\mu$$

se e solo se π è concentrato in un grafico G_T π -misurabile di una funzione T .

Dimostrazione. Si è già visto che se T è una mappa di trasporto allora $\pi_T := (Id \times T)_{\#}\mu \in \Gamma(\mu, \nu)$.

Per la π_T -misurabilità del grafico di T , fissiamo un $\varepsilon > 0$. Per la proprietà di Lusin, esiste un compatto $K_\varepsilon \subset X$ tale che $\mu(X \setminus K_\varepsilon) < \varepsilon$ e T ristretta a K_ε è continua. Allora gli insiemi $G_\varepsilon := K_\varepsilon \times X \cap G_T$ sono boreliani per continuità di T , e si ha $\pi_T(X \setminus G_\varepsilon) = \mu(X \setminus K_\varepsilon) < \varepsilon$, quindi il grafico G_T è π_T -misurabile. Infine, per la (1.1) abbiamo che se $A \cap G_T = \emptyset$ allora

$$\pi_T(A) = \int \mathbb{1}_A d(Id \times T)_{\#}\mu = \int_{(Id \times T)^{-1}(A)} \mathbb{1} d\mu = 0,$$

quindi π_T è concentrato su G_T .

Mostriamo l'altra implicazione. L'idea è di confrontare il piano π con il piano indotto da π_T e di vedere che le condizioni sui marginali forzano l'uguaglianza di questi piani. Mostriamo quindi che T è μ -misurabile. Infatti, per definizione esiste un boreliano $B \subset G_T$ tale che $\pi(G_T \setminus B) = 0$; per regolarità di π esiste una successione crescente di compatti $C_h \subset C_{h+1} \subset G_T$ tali che $\pi(G_T \setminus \cup_h C_h) = 0$. Indicando con K_h la proiezione di C su X abbiamo che anche K_h è compatto. Poiché C_h è compatto, si ha che $T|_{C_h}$ è continua; essendo $\mu(X \setminus \cup_h K_h) \leq \pi(G_T \setminus \cup_h C_h) = 0$ abbiamo che T è μ -misurabile, pertanto è ben definito $\pi_T := (Id \times T)_{\#}\mu$.

Per dedurre che $\pi_T = \pi$ basta verificare che vale

$$\pi_k := \pi \llcorner C_k = \pi_T \llcorner C_k =: \gamma_k,$$

perché si ottiene l'uguaglianza voluta passando al limite per $k \rightarrow \infty$, ed è sufficiente verificarlo sugli insiemi $A \times B$, con $A, B \subset X$ compatti. Poiché entrambe le misure sono concentrate su C_k si ha

$$\gamma_k(A \times B) = \gamma_k((A \times B) \cap C_k) = \gamma_k(p_X((A \times B) \cap C_k) \times X) = \mu(p_X((A \times B) \cap C_k))$$

e analogamente per π_k , quindi $\gamma_k = \pi_k$. Infine, dalla relazione $T_{\#}\mu = (p_Y)_{\#}\pi_T = (p_Y)_{\#}\pi = \nu$ si ha che T è una mappa di trasporto. \square

2.2.2 Proprietà di regolarità dei piani ottimali

L'intenzione a questo punto è quindi di trovare proprietà di regolarità del supporto dei piani ottimali. Si noti come per molti dei risultati in questa sezione esiste un risultato analogo anche per altri costi. Per motivare la Definizione 9 nel caso generale cominciamo con un esempio di condizione necessaria e sufficiente in un caso discreto.

Esempio 6. Consideriamo l'Esempio 5. Nel caso di costo quadratico, dato un piano $\pi = (\pi_{ij}) \in \Gamma$ si ha

$$\mathcal{C}(\pi) = \sum_{i,j=1}^n |x_i - y_j|^2 \pi_{ij}.$$

In questo caso particolare il problema di Monge è quindi un problema di ottimizzazione di un funzionale affine sull'insieme convesso

$$D = \left\{ \pi = (\pi_{ij}) \mid \pi_{ij} \geq 0 \quad \text{e} \quad \sum_{i=1}^n \pi_{ij} = \sum_{j=1}^n \pi_{ij} = \frac{1}{n} \right\}$$

I minimi di questo problema sono quindi assunti anche tra i punti estremali dell'insieme convesso; per il teorema di Birkhoff, tali punti estremali corrispondono alle matrici di permutazione, ovvero esattamente ai piani indotti dalle mappe di trasporto. In questo caso quindi esiste un piano π_T ottimale indotto da una mappa T . Possiamo supporre, a meno di rinumerare i punti, che $T(x_i) = y_i$.

Un'altra mappa di trasporto S può essere ottenuta semplicemente permutando le immagini y_i , ovvero esiste una permutazione $\sigma \in S_n$ tale che $S = \sigma \circ T$. La condizione $\mathcal{C}(S) \geq \mathcal{C}(T)$ di ottimalità per T può quindi essere scritta come

$$\sum_{i=1}^n |x_i - y_i|^2 \leq \sum_{i=1}^n |x_i - y_{\sigma(i)}|^2 \quad \forall \sigma \in S_n, \quad (2.6)$$

che è una condizione di monotonia del supporto di π_T equivalente alla sua ottimalità.

Siamo quindi motivati a dare la seguente definizione.

Definizione 9. Un sottoinsieme $S \subset \mathbb{R}^n \times \mathbb{R}^n$ è detto *ciclicamente monotono* se per ogni $k \geq 1$ e per ogni k -upla $((x_1, y_1), \dots, (x_k, y_k)) \subset S$ si ha

$$\sum_{i=1}^k |x_i - y_i|^2 \leq \sum_{i=1}^k |x_i - y_{i+1}|^2, \quad (2.7)$$

con $y_{k+1} = y_1$.

Anche se la definizione appena data può sembrare più debole della condizione di monotonia trovata prima, in realtà è equivalente. Infatti per mostrare l'equivalenza ci si può limitare a permutazioni formate da un solo ciclo, perché si può scomporre una permutazione σ come prodotto di cicli e punti fissi e quindi sommare le disuguaglianze valide per ognuno di quei cicli.

Prendendo quindi $z_i := x_{\sigma^i(1)}$, $w_i := y_{\sigma^i(1)}$ e applicando a queste k -uple la (2.7) si ottiene,

$$\sum_{i=1}^k |x_{\sigma^i(1)} - y_{\sigma^i(1)}|^2 \leq \sum_{i=1}^k |x_{\sigma^{i+1}(1)} - y_{\sigma^{i+1}(1)}|^2,$$

ed essendo σ un ciclo ogni membro di questa disuguaglianza è ottenuto permutando il corrispondente membro della (2.6).

Analogamente si dimostra che nella Definizione 9 la (2.7) è verificata se e solo se, con le stesse notazioni, si ha

$$\sum_{i=1}^k |x_i - y_i|^2 \leq \sum_{i=1}^k |x_{i+1} - y_i|^2,$$

che a sua volta è equivalente, espandendo i quadrati, a

$$\sum_{i=1}^k y_i \cdot (x_{i+1} - x_i) \leq 0, \quad (2.8)$$

dove $x_1 = x_{k+1}$.

Inoltre, ponendo $k = 2$ nella (2.8) si ottiene

$$(y_2 - y_1) \cdot (x_2 - x_1) \geq 0,$$

che è la proprietà di monotonia propria dei gradienti delle funzioni convesse. Questa osservazione motiva lo studio di tali funzioni.

Definizione 10 (Funzioni convesse). Una funzione $u: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ è una funzione convessa se è semicontinua inferiormente, non vale identicamente $+\infty$ e si ha

$$\lambda u(x) + (1 - \lambda)u(y) \geq u(\lambda x + (1 - \lambda)y) \quad \forall x, y \in \mathbb{R}^n, \forall \lambda \in [0, 1]. \quad (2.9)$$

Il *dominio* di u è il sottoinsieme di \mathbb{R}^n definito da

$$\text{dom } u := \text{Int} \{ x \in \mathbb{R}^n \mid u(x) < +\infty \},$$

dove $\text{Int}(A)$ denota la parte interna di A .

Il *sottodifferenziale* di u in $x \in \mathbb{R}^n$ è definito come

$$\partial u(x) := \{ y \in \mathbb{R}^n \mid u(x) + z \cdot y \leq u(x+z) \forall z \in \mathbb{R}^n \}.$$

In seguito, con un abuso di notazione, si indicherà con ∂u il grafico del sottodifferenziale di u .

La funzione *coniugata* di u è la funzione $u^*: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ definita da

$$u^*(y) := \sup \{ x \cdot y - u(x) \}.$$

È importante notare come le funzioni sopra definite convesse sono genericamente chiamate funzioni convesse proprie e semicontinue inferiormente. Inoltre, il dominio di u è solitamente inteso senza prendere la parte interna. Queste differenze sono motivate dal fatto che non siamo interessati a studiare le funzioni convesse in generale ma solo quelle proprietà che consentiranno di provare il teorema 6 di regolarità dei piani ottimali.

Riassumiamo ora alcune proprietà ben note delle funzioni convesse che serviranno in seguito. Per la dimostrazione rimandiamo a [Ro] e [ET].

Proposizione 3 (Proprietà delle funzioni convesse). *Sia $u: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una funzione convessa. Allora*

- i) *il dominio $\text{dom } u$ è un sottoinsieme convesso di \mathbb{R}^n ;*
- ii) *u è Lipschitziana in ogni compatto $K \subset \text{dom } u$;*
- iii) *per ogni $x \in \text{dom } u$ il sottodifferenziale di u in x è non vuoto, inoltre è costituito da un solo punto se e solo se u è differenziabile in x e in tal caso si ha $\partial u(x) = \{\nabla u(x)\}$;*
- iv) *la funzione coniugata u^* è una funzione convessa. Vale inoltre la disuguaglianza di Fenchel*

$$\langle x, y \rangle \leq u(x) + u^*(y) \quad \forall x, y \in \mathbb{R}^n, \quad (2.10)$$

dove vale l'uguaglianza se e solo se $y \in \partial u(x)$.

Il seguente teorema è un risultato classico di analisi convessa che caratterizza i sottoinsiemi di \mathbb{R}^{2n} ciclicamente monotoni. Data la sua rilevanza per la teoria del trasporto ottimale ne diamo una dimostrazione completa.

Teorema 5 (Rockafellar). *Sia $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$. Allora esiste una funzione convessa $u: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ convessa tale che $\Gamma \subset \partial u$ se e solo se Γ è ciclicamente monotono.*

Dimostrazione. Sia u come nell'enunciato, allora presi $((x_1, y_1), \dots, (x_k, y_k)) \in \partial u$ abbiamo per definizione che

$$u(x_{i+1}) - u(x_i) \geq y_i \cdot (x_{i+1} - x_i).$$

Sommando per $i = 1, \dots, k$ si ottiene la (2.8), ovvero ∂u è ciclicamente monotono, e quindi lo è ogni suo sottoinsieme.

D'altro canto, dato Γ come nell'enunciato, si fissi $(x_0, y_0) \in \Gamma$ e si consideri la funzione

$$u(x) := \sup \{ y_k \cdot (x - x_k) + \dots + y_0 \cdot (x_1 - x_0), (x_i, y_i) \in \Gamma, i = 1, \dots, k \}.$$

Essendo l'estremo superiore di funzioni affini, u è semicontinua inferiormente e soddisfa la disuguaglianza (2.9), inoltre prendendo $x = x_0 = x_1, y_1 = y_0$ si ha $u(x_0) \geq 0$. Poiché Γ è ciclicamente monotono allora per ogni scelta di $(x_1, y_1), \dots, (x_k, y_k) \in \Gamma$ si ha $y_k \cdot (x_0 - x_k) + \dots + y_0 \cdot (x_1 - x_0) \leq 0$ pertanto $u(x_0) = 0$. Quindi u non assume mai valore $-\infty$ e non è identicamente $+\infty$, e quindi u è convessa nel senso della definizione 10.

Basta quindi mostrare che $\Gamma \subset \partial u$, ovvero scelti $(x, y) \in \Gamma$ si ha

$$u(x+z) \geq u(x) + z \cdot y.$$

Innanzitutto $u(x) < \infty$, infatti scelti m punti $(x_1, y_1), \dots, (x_k, y_k) \in \Gamma$ per la (2.8) applicata a $(x_1, y_1), \dots, (x_k, y_k), (x, y)$ si ha che $u(x) \leq (x - x_0) \cdot y$.

Infine, per ogni $\varepsilon > 0$ esiste una k -upla $(x_1, y_1), \dots, (x_k, y_k)$ tale che

$$u(x) \leq y_k \cdot (x - x_k) + \dots + y_0 \cdot (x_1 - x_0) + \varepsilon.$$

Pertanto, usando nella definizione di $u(x+z)$ la $(k+1)$ -upla $(x_1, y_1), \dots, (x_k, y_k), (x, y)$ si ottiene

$$u(z+x) + \varepsilon \geq y \cdot z + y_k \cdot (x - x_k) + \dots + y_0 \cdot (x_1 - x_0) + \varepsilon \geq u(x) + z \cdot y,$$

e per l'arbitrarietà di ε si conclude. \square

Osservazione 2. Si noti che se Γ è limitato, i.e. $\Gamma \subset B_R$, la funzione convessa u del Teorema precedente può essere presa finita su tutto \mathbb{R}^n . Infatti, vedendo la dimostrazione, si ha che u è definita come estremo superiore di funzioni R -lipschitziane, ed è quindi R -lipschitziana. In particolare $|u(x)| \leq R|x - x_0|$.

Teorema 6 (Regolarità dei piani ottimali). *Siano $\mu, \nu \in \mathcal{P}(X)$ tali che valga*

$$\int_X |x|^2 d\mu(x) + \int_X |y|^2 d\nu(y) < +\infty, \quad (2.11)$$

e sia $\pi \in \Gamma(\mu, \nu)$ un piano di trasporto. Allora π è ottimale se e solo se il suo supporto è ciclicamente monotono.

Dimostrazione. Necessità. Sia π un piano ottimale, e supponiamo per assurdo che non sia ciclicamente monotono. Allora esiste un intero positivo k e due k -uple di punti $(x_1, \dots, x_k), (y_1, \dots, y_k) \in X^2$ tali che $(x_i, y_i) \in \text{supp } \pi, i = 1, \dots, k$ tali che

$$\sum_{i=1}^k |x_i - y_i|^2 > \sum_{i=1}^k |x_i - y_{i+1}|^2.$$

L'idea allora è che trasferire massa da x_i a y_{i+1} è più conveniente che spostarla da x_i a y_i . Per rendere rigorosa questa intuizione, vogliamo costruire una probabilità $\eta \in \mathcal{M}(X \times X)$ tale che:

- a) $\pi + \eta \in M_+(X \times X)$;
- b) i marginali di η siano nulli;
- c) $\int_{X \times X} |x - y|^2 d\eta(x, y) < 0$.

In questo caso, infatti, le condizioni a) e b) assicurano che $\bar{\pi} := \pi + \eta$ sia un piano di trasporto ammissibile, mentre la c) permette di concludere che $\mathcal{C}(\bar{\pi}) < \mathcal{C}(\pi)$, che è assurdo per ottimalità di π .

Consideriamo quindi degli intorni U_i, V_i di x_i, y_i rispettivamente tali che $m_i := \pi(V_i) > 0$ e

$$\sum_{i=1}^k |u_i - v_i|^2 > \sum_{i=1}^k |u_i - v_{i+1}|^2 \quad \forall u_i \in U_i, v_i \in V_i, \quad (2.12)$$

che esistono per continuità del costo e per definizione di supporto.

Sia (Ω, \mathbb{P}) lo spazio di probabilità definito da

$$\Omega := \prod_{i=1}^k (U_i \times V_i), \quad \mathbb{P} := \prod_{i=1}^k \frac{\pi \llcorner (U_i \times V_i)}{m_i},$$

e indichiamo con $\phi_{i,j}: \Omega \rightarrow X \times X$ la mappa che a $(x_i, y_i)_{i=1}^k \in \Omega$ associa $(x_i, y_j) \in X \times X$. La misura

$$\eta := \frac{\min_i m_i}{k} \sum_{i=1}^k \phi_{i,i+1\#} \mathbb{P} - \phi_{i,i\#} \mathbb{P}$$

ha le proprietà richieste.

Infatti, la a) è soddisfatta perché per ogni boreliano $A \subset X \times X$ vale

$$\pi(A) \geq \pi((U_i \times V_i) \cap (A)) \geq \frac{m_i}{m_i} \pi \llcorner (U_i \times V_i)(A) = \phi_{i,i\#} \mathbb{P}(A).$$

Per verificare la b) basta mostrare, per additività, che

$$\sum_{i=1}^k (\pi_h \circ \phi_{i,i})_{\#} \mathbb{P} - (\pi_h \circ \phi_{i,i+1})_{\#} \mathbb{P} = 0, \quad h = 1, 2, \quad (2.13)$$

ma per ogni coppia di $1 \leq i, j \leq k$ si ha

$$(\pi_1 \circ \phi_{i,j})_{\#} \mathbb{P} = \frac{1}{m_i} \gamma \llcorner (U_i \times V_i), \quad (\pi_2 \circ \phi_{i,j})_{\#} \mathbb{P} = \frac{1}{m_j} \gamma \llcorner (U_j \times V_j),$$

pertanto sommando su i le precedenti relazioni si ottiene la (2.13).

Infine, per mostrare la c) basta verificare, per linearità, che

$$\sum_{i=1}^k \int_{X \times X} |x - y|^2 d(\phi_{i,i})_{\#} \mathbb{P} > \sum_{i=1}^k \int_{X \times X} |x - y|^2 d(\phi_{i,i+1})_{\#} \mathbb{P}.$$

Per la (1.1) la disuguaglianza precedente è equivalente a

$$\int_{\Omega} \sum_{i=1}^k |x_i - y_i|^2 d\mathbb{P}((x_l, y_l)_{l=1}^k) > \int_{\Omega} \sum_{i=1}^k |x_i - y_{i+1}|^2 d\mathbb{P}((x_l, y_l)_{l=1}^k),$$

che è vera per la (2.12).

Sufficienza. Sia π un piano di trasporto tale che $\text{supp } \pi$ sia ciclicamente monotono. Dal Teorema 5 sappiamo che esiste una funzione convessa u tale che $\text{supp } \pi \subset \partial u$. Sia quindi γ un qualsiasi piano di trasporto. Si ha che

$$\begin{aligned} \mathcal{E}(\pi) &= \int_{X \times X} |x - y|^2 d\pi(x, y) \\ &= \int_X |x|^2 d\mu(x) + \int_X |y|^2 d\nu(y) - 2 \int_{X \times X} x \cdot y d\pi(x, y) \\ &= \mathcal{E}(\gamma) + \int_{X \times X} x \cdot y d(\gamma - \pi)(x, y), \end{aligned}$$

pertanto per mostrare l'ottimalità di η basta mostrare che l'ultimo termine è minore o uguale di 0.

Poiché $|x \cdot y| \leq \frac{1}{2}(|x|^2 + |y|^2)$, dalla (2.10) si ottiene $u^*(y) \geq -\frac{|x|^2 + |y|^2}{2} - u(x)$. Scegliendo x tale che $u(x) < +\infty$ si ha che il membro a destra è sommabile rispetto a ν , e quindi lo è anche la parte negativa di u^* . Analogamente si ottiene che $u^- \in L^1(\mu)$. Pertanto per ogni piano di trasporto γ si ha $\int_{X \times X} (u + u^*) d\gamma = \int_X u d\mu + \int_X u^* d\nu$.

Ricordando quindi la disuguaglianza di Fenchel si ottiene

$$\begin{aligned}
 \int_{X \times X} x \cdot y \, d\pi(x, y) &= \int_{X \times X} (u(x) + u^*(y)) \, d\pi(x, y) \\
 &= \int_X u(x) \, d\mu(x) + \int_X u^*(y) \, d\nu(y) \\
 &= \int_{X \times X} (u(x) + u^*(y)) \, d\gamma(x, y) \\
 &\geq \int_{X \times X} x \cdot y \, d\gamma(x, y),
 \end{aligned}$$

dove si è usata la caratterizzazione dei casi di uguaglianza della (2.10). Essendo $\gamma \in \Gamma(\mu, \nu)$ arbitrario si ottiene l'ottimalità di π . \square

2.2.3 Il Teorema di Brenier

Grazie al Teorema di regolarità dei piani ottimali e al Teorema di Rockafellar è possibile intravedere una candidata mappa ottimale per il problema (MQ). Infatti, se la funzione u data dal Teorema 5 fosse differenziabile in ogni punto allora il grafico del sottodifferenziale coinciderebbe col grafico del gradiente di u , che per il Teorema 6 sarebbe quindi la mappa ottimale.

Il controesempio all'esistenza, mostrato nell'Esempio 2, si basa sul fatto che le funzioni convesse non sono sempre differenziabili. Tuttavia, è possibile dimostrare che le funzioni convesse sono differenziabili \mathcal{L}^n -quasi ovunque. Il seguente risultato, infatti, è ben noto (rimandiamo a [T] per la dimostrazione).

Teorema 7 (Rademacher). *Sia $f: \mathbb{R}^n \rightarrow \mathbb{R}$ lipschitziana. Allora f è differenziabile \mathcal{L}^n -quasi ovunque.*

Poiché per la proposizione 3 le funzioni convesse sono localmente lipschitziane, ed essendo la differenziabilità una proprietà locale, si ottiene immediatamente che anche le funzioni convesse sono differenziabili \mathcal{L}^n -quasi ovunque.

Teorema 8 (Brenier). *Sia $X \subset \mathbb{R}^n$ compatto e siano $\mu, \nu \in \mathcal{P}(X)$ due misure di probabilità su X . Supponiamo che $\mu \ll \mathcal{L}^n$ e che μ, ν soddisfino la (2.11). Allora*

- i) esiste un unico piano ottimale di trasporto π per il costo quadratico. Inoltre, tale piano di trasporto è indotto da una mappa T , soluzione di (MQ), ed è tale che esiste una funzione convessa $u: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ differenziabile μ -quasi ovunque tale che $T = \nabla u$;*
- ii) viceversa, se $u: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ è una funzione convessa tale che $\nabla u \in L^2(\mu)$, allora la mappa $T = \nabla u \in \Psi(\mu, T_{\#}\mu)$ è ottimale.*

Dimostrazione. i) Per continuità del costo si può applicare il Teorema 3, da cui si deduce l'esistenza di un piano ottimale π . Poiché le misure soddisfano la (2.11) si può applicare il teorema 6 e si ha quindi che $\text{supp } \pi$ è ciclicamente monotono. Per il teorema 5 esiste una funzione convessa u tale che $\text{supp } \pi \subset \partial u$. Sia Σ l'insieme dei punti di non differenziabilità di u ; grazie al teorema 7 si ha $\mathcal{L}^n(\Sigma) = 0$ e per assoluta continuità di μ si ottiene $\pi(\Sigma \times X) = \mu(\Sigma) = 0$. Pertanto, π è concentrato sull'insieme $\partial u \setminus (\Sigma \times X)$ che è il grafico di $T = \nabla u$, e per il teorema 4 T è una mappa ottimale.

Abbiamo così mostrato che ogni piano ottimale è indotto da una mappa di trasporto ottimale. Per provare l'unicità basta notare che se π_1, π_2 sono piani ottimali indotti da T_1, T_2 rispettivamente, allora anche $\tilde{\pi} := \frac{\pi_1}{2} + \frac{\pi_2}{2}$ è un piano ottimale, ma per il teorema 4 è indotto da una mappa se e solo se $T_1 = T_2$ μ -quasi ovunque, ovvero se e solo se $\pi_1 = \pi_2$.

ii) Sia $\nu = T_{\#}\mu$. Il piano $\pi_T := (Id \times T)_{\#}\mu$ indotto da T è un piano di trasporto tra μ e ν ; poiché $T = \nabla u$, con u convessa, allora il grafico G_T è contenuto nel sottodifferenziale di u ed è pertanto ciclicamente monotono. Poiché $\nabla u \in L^2(\mu)$ si ha che μ, ν soddisfano la (2.11) e pertanto per il teorema 6 π_T è ottimale, e quindi anche T .

□

Capitolo 3

Il problema isoperimetrico

Il problema isoperimetrico, probabilmente uno dei problemi più antichi della matematica, chiede di determinare, tra tutti i sottoinsiemi del piano con area fissata, quelli che minimizzano il perimetro. Anche se la soluzione, il cerchio, era nota sin dall'antichità (celebre a questo proposito l'episodio su Didone citato nell'Eneide) una dimostrazione completa è stata data solo nel XIX secolo, seguita da ulteriori generalizzazioni in dimensione maggiore e classi più ampie di insiemi, fino al risultato di De Giorgi del 1958 ([G]) di ottimalità della sfera nella classe di insiemi di perimetro finito.

Si noti come parte integrante del problema è la sua formulazione, in quanto la definizione di perimetro non è immediata se la classe in cui si ambienta il problema ammette insiemi irregolari. Per semplicità enunciamo e risolviamo il problema isoperimetrico nella classe degli insiemi aperti limitati di classe C^1 (si veda la Definizione 15), per i quali la definizione di perimetro è la misura $(n - 1)$ -dimensionale del bordo.

Il problema isoperimetrico può essere quindi enunciato nel modo seguente.

(P) Sia $n \geq 2$. Indicando con $E \subset \mathbb{R}^n$ un aperto limitato con bordo C^1 e con σ_{n-1} la misura $(n - 1)$ -dimensionale, il problema è di determinare

$$\min_{\mathcal{L}^n(E)=1} \sigma_{n-1}(\partial E). \quad (\mathbf{P})$$

L'ottimalità della sfera per il problema isoperimetrico implica, grazie all'invarianza del problema per riscalamento, la validità della disuguaglianza isoperimetrica, per la quale per ogni aperto $E \subset \mathbb{R}^n$ di classe C^1 con $\mathcal{L}^n(E) < \infty$ vale

$$\mathcal{L}^n(E) \leq C_n \sigma_{n-1}(\partial E)^{\frac{n}{n-1}}, \quad (3.1)$$

dove $C_n = n^{n'} |B_1|^{1+n'}$, $n' = n/(1 - n)$.

Le prime soluzioni del problema isoperimetrico utilizzano tecniche geometriche come la simmetrizzazione di Steiner o la disuguaglianza di Brunn-Minkowski. In [MS] Gromov propose una soluzione del problema isoperimetrico con l'utilizzo di una mappa di trasporto (la mappa di Knoethe) che è tanto flessibile quanto le precedenti e, anzi, si può affinare per dimostrare generalizzazioni del problema come la disuguaglianza isoperimetrica anisotropa o come la disuguaglianza isoperimetrica quantitativa.

La soluzione di Gromov è adattabile alla mappa di Brenier e, almeno formalmente, si può presentare in poche righe. Infatti, dato $E \subset \mathbb{R}^n$ tale che $\mathcal{L}^n(E) = 1$, si ponga $\mu = \mathcal{L}^n \llcorner E$, $\nu = \mathcal{L}^n \llcorner B$, con B la palla centrata nell'origine tale che $\mathcal{L}^n(B) = 1$. Indicando con T la mappa di Brenier da μ a ν abbiamo che $|T| \leq 1$ μ -quasi ovunque, e poiché la relazione $T_{\#}\mu = \nu$ implica che T conserva i volumi dalla formula di cambio di variabile otteniamo $|\det \nabla T| = 1$, che per la disuguaglianza tra medie applicata agli autovalori di ∇T (che è simmetrica semidefinita positiva) diventa $\operatorname{div} T \geq n$. Pertanto, almeno formalmente, si ha

$$\sigma_{n-1}(B) = n\mathcal{L}^n(E) \leq \int_E \operatorname{div} T = \int_{\partial E} \nu_E \cdot T \leq \sigma_{n-1}(E). \quad (3.2)$$

Lo scopo di questo Capitolo è quindi di mostrare proprietà di regolarità delle mappe di trasporto che permettano di rendere rigorosi questi passaggi formali. Sarà possibile inoltre mostrare che l'unica soluzione, nella classe di insiemi considerati, è la palla.

3.1 Stabilità L^p delle mappe ottimali

Lo scopo di questa sezione è di dimostrare il Teorema 9, che mostra come le mappe ottimali siano stabili in L^p rispetto alla convergenza debole delle misure immagine. Facendo tesoro della linea dimostrativa del Teorema di Brenier, anche in questo caso si mostra dapprima la stabilità dei piani di trasporto per poi dedurne la stabilità delle mappe.

Proposizione 4 (Stabilità dei piani ottimali). *Sia $X \subset \mathbb{R}^n$ compatto, e siano $(\mu_k), (\nu_k) \subset \mathcal{P}(X)$ che convergono debolmente a $\mu, \nu \in \mathcal{P}(X)$ rispettivamente, e supponiamo inoltre che μ, ν soddisfino la (2.11). Sia quindi scelto per ogni intero $k \geq 1$ un piano $\pi_k \in \Gamma(\mu_k, \nu_k)$ ottimale. Allora la successione (π_k) ammette sottosuccessioni debolmente convergenti, inoltre ogni punto limite è ottimale in $\Gamma(\mu, \nu)$.*

Dimostrazione. Essendo π_k una successione di probabilità, per il Corollario 1 ammette una successione debolmente convergente, e per il Lemma 2 ogni punto di accumulazione è un piano in $\Gamma(\mu, \nu)$.

Supponiamo quindi che $\pi_n \rightharpoonup \pi \in \Gamma(\mu, \nu)$. Poiché μ, ν soddisfano la (2.11) per mostrare l'ottimalità di π è sufficiente mostrare che $\text{supp } \pi$ è ciclicamente monotono.

Mostriamo preliminarmente che se $(x, y) \in \text{supp } \pi$ allora esiste una successione $(x_k, y_k) \in \text{supp } \pi_k$ tale che $(x_k, y_k) \rightarrow (x, y)$. Infatti, se per assurdo non fosse vero, esisterebbe un aperto $A \ni (x, y)$ tale che $\pi(A) > 0$ ma per cui esistono infiniti piani π_{n_k} tali che $A \cap \text{supp } \pi_{n_k} = \emptyset$, ovvero $\pi_{n_k}(A) = 0$, il che è assurdo per la (2.4).

Poiché da $|x|^2 \in \mathcal{C}(X)$ si ha

$$\lim_k \int_X |x|^2 d\mu_k(x) + \int_X |y|^2 d\nu_k(y) = \int_X |x|^2 d\mu(x) + \int_X |y|^2 d\nu(y) < +\infty,$$

si può assumere senza perdita di generalità che anche μ_k, ν_k soddisfino la (2.11). Per il Teorema 6 si ha che $\text{supp } \pi_n$ è ciclicamente monotono. Pertanto, preso un intero $h \geq 1$ e due h -uple $(x_1, \dots, x_h), (y_1, \dots, y_h)$ tali che $(x_i, y_i) \in \text{supp } \pi$, $i = 1, \dots, h$ esistono delle successioni $x_{k,i}, y_{k,i}$ tali che $\text{supp } \pi_k \ni (x_{k,i}, y_{k,i}) \rightarrow (x_i, y_i)$ per $k \rightarrow \infty$. Poiché il supporto di π_k è ciclicamente monotono abbiamo che per ogni k vale

$$\sum_{i=1}^h |x_{k,i} - y_{k,i}|^2 \leq \sum_{i=1}^h |x_{k,i} - y_{k,i+1}|^2$$

con $y_{k,i+1} = y_{k,1}$, e pertanto passando al limite su k si ottiene la (2.7), come voluto. \square

Dedurre la convergenza in L^p delle mappe ottimali dalla convergenza debole dei piani che inducono è un passaggio delicato che passa attraverso la convergenza in misura delle mappe. Tale convergenza va intesa come intermedia tra le due, infatti se da un lato la convergenza in misura è legata alla convergenza in L^p dal Lemma 3, dall'altro è intimamente legata alla convergenza debole dalla Proposizione 5.

Definizione 11 (Convergenza in misura). Sia $X \subset \mathbb{R}^n$, sia $\mu \in \mathcal{P}(X)$ e siano $f_k, f: X \rightarrow X$ boreliane. Si dice che f_k tende a f in misura se per ogni $\varepsilon > 0$

$$\lim_k \mu(\{|f_k - f| > \varepsilon\}) = 0.$$

Il seguente lemma lega la convergenza in misura alla convergenza in L^p .

Lemma 3. Sia $X \subset \mathbb{R}^n$, sia $\mu \in \mathcal{P}(X)$ e siano $f_k, f: X \rightarrow X$ boreliane. Allora

$$f_k \rightarrow f \text{ in misura} \iff \int_X \min\{1, |f_k - f|\} d\mu \rightarrow 0.$$

Inoltre, se $M := \sup_k \|f_k\|_\infty < \infty$ e $f_k \rightarrow f$ in misura, allora $f_k \rightarrow f$ in L^p per ogni $p \in [1, \infty)$.

Dimostrazione. Poniamo per comodità $\phi_k := \min\{1, |f_k - f|\}$.

Supponiamo che $f_k \rightarrow f$ in misura. Per il teorema di Fubini-Tonelli applicato alla funzione $\mathbb{1}_{\{\phi_k > t\}}(x) \in L^1(X \times [0, 1], \mu \times (\mathcal{L}^1 \llcorner [0, 1]))$ abbiamo che

$$\int_X \phi_k \, d\mu = \int_0^1 \mu(\{|f - f_k| > s\}) \, ds.$$

Poiché per ipotesi la successione di funzioni $s \mapsto \mu(|f - f_k| > s)$ tende puntualmente a zero, per convergenza dominata (si ricorda che μ è una probabilità) si ha che $\int_X \phi_k \, d\mu \xrightarrow{k} 0$, come voluto.

D'altro canto, se $\int_X \phi_k \, d\mu \rightarrow 0$ per la disuguaglianza di Markov si ha, se $\varepsilon < 1$,

$$\mu(|f_k - f| > \varepsilon) = \mu(\phi_k > \varepsilon) \leq \frac{1}{\varepsilon} \int_X \phi_k \, d\mu \xrightarrow{k} 0.$$

Poiché se $\varepsilon' > \varepsilon$ allora $\mu(|f_k - f| > \varepsilon') \leq \mu(|f_k - f| > \varepsilon)$ si ha che $f_k \rightarrow f$ in misura.

Supponiamo ora che $\|f_k\|_\infty \leq M < \infty$. Posto $\ell := \limsup_k \|f_k - f\|_p$ è sufficiente mostrare $\ell = 0$, e per farlo sia f_{k_j} una sottosuccessione tale che $\|f_{k_j} - f\|_p \rightarrow \ell$. Poiché si ha $\phi_{k_j} \rightarrow 0$ in L^1 , a meno di sottosuccessione si ha $|f - f_{k_j}| \rightarrow 0$ μ -quasi ovunque, ovvero $f_{k_j} \rightarrow f$ μ -quasi ovunque. Pertanto si ha $|f| \leq M$ μ -quasi ovunque e quindi $|f - f_{k_j}|^p \leq (2M)^p$. Per convergenza dominata si ha allora $\ell = \lim_j \int_X |f - f_{k_j}|^p \, d\mu = 0$. \square

La seguente proposizione lega la convergenza in misura delle mappe alla convergenza debole dei piani indotti.

Proposizione 5. *Siano $X \subset \mathbb{R}^n$ compatto, $\mu \in \mathcal{P}(X)$ e $f_k, f: X \rightarrow X$ boreliane. Allora*

$$f_k \rightarrow f \text{ in misura} \iff (Id \times f_k)_\# \mu \rightharpoonup (Id \times f)_\# \mu.$$

Dimostrazione. Supponiamo che $f_k \rightarrow f$ in misura, e supponiamo per assurdo che non valga $(Id \times f_k)_\# \mu \rightharpoonup (Id \times f)_\# \mu$. Allora esistono $\varepsilon > 0$, $\phi \in \mathcal{C}(X)$ e una sottosuccessione f_{n_k} tali che

$$\left| \int_{X \times X} \phi \, d(Id \times f_{n_k})_\# \mu - \int_{X \times X} \phi \, d(Id \times f)_\# \mu \right| \geq \varepsilon, \quad (3.3)$$

che per la (1.1) è equivalente a

$$\left| \int_X (\phi(x, f_{n_k}(x)) - \phi(x, f(x))) \, d\mu(x) \right| > \varepsilon.$$

Detto ω il modulo di continuità di ϕ e scelto δ tale che $\omega(\delta) < \varepsilon/2$, si ha

$$\left| \int_X [\phi(x, f_{n_k}(x)) - \phi(x, f(x))] d\mu(x) \right| \leq \|\phi\|_\infty \mu(|f - f_{n_k}| \geq \delta) + \frac{\varepsilon}{2} \xrightarrow[k]{} \frac{\varepsilon}{2},$$

che contraddice la (3.3).

Supponiamo ora che $(Id \times f_k)_{\#}\mu \rightarrow (Id \times f)_{\#}\mu$. Fissato $\varepsilon > 0$ per la proprietà di Lusin esiste una funzione continua $\bar{f}: X \rightarrow X$ tale che $\mu(f \neq \bar{f}) < \varepsilon$. Considerando la funzione $\phi \in \mathcal{C}(X \times X)$ definita da

$$\phi(x, y) = \min \{1, |\bar{f}(x) - y|\},$$

per ipotesi abbiamo che $\int_{X \times X} \phi d(Id \times f_k)_{\#}\mu \rightarrow \int_{X \times X} \phi d(Id \times f)_{\#}\mu$, che per la (1.1) implica

$$\lim_k \int_X \min \{1, |f_k(x) - \bar{f}(x)|\} d\mu(x) = \int_X \min \{1, |f - \bar{f}|(x)\} d\mu(x) < \varepsilon,$$

pertanto definitivamente in k abbiamo

$$\int_X \min \{1, |f_k - f|\} d\mu \leq \int_X \min \{1, |f_k - \bar{f}|\} d\mu + \int_X \min \{1, |f - \bar{f}|\} d\mu < 2\varepsilon,$$

da cui, essendo ε arbitrario, si conclude per il Lemma 3. \square

Dati $X \subset \mathbb{R}^n$ e due probabilità $\mu, \nu \in \mathcal{P}(X)$ tali che $\mu \ll \mathcal{L}^n$ e che soddisfano la (2.11), indichiamo con T_μ^ν la mappa di trasporto ottimale per il problema (MQ).

Teorema 9 (Stabilità L^p delle mappe ottimali). *Sia $X \subset \mathbb{R}^n$ compatto, e siano $\mu, \nu \in \mathcal{P}(X)$ con $\mu \ll \mathcal{L}^n$ tali che valga (2.11). Consideriamo una successione $\nu_k \subset \mathcal{P}(X)$ tale che $\nu_k \rightarrow \nu$. Allora $T_\mu^{\nu_k} \rightarrow T_\mu^\nu$ in $L^p(X, \mu)$ per ogni $p \in [1, \infty)$.*

Dimostrazione. Per il Lemma 3 basta mostrare che $T_\mu^{\nu_k}$ sia equilimitata in $L^\infty(\mu)$ e che $T_\mu^{\nu_k} \rightarrow T_\mu^\nu$ in misura. Per la Proposizione 4 si ha $(Id \times T_\mu^{\nu_k})_{\#}\mu \rightarrow (Id \times T_\mu^\nu)_{\#}\mu$, e per il Teorema 5 si ha che $T_\mu^{\nu_k} \rightarrow T_\mu^\nu$ in misura. Inoltre, poiché X è compatto esiste una costante $R > 0$ tale che $X \subset B_R$, e poiché $\text{supp } \nu_k \subset X \subset B_R$ abbiamo che $T_\mu^{\nu_k} \in B_R$ μ -quasi ovunque, ovvero la successione $T_\mu^{\nu_k}$ è limitata in $L^\infty(\mu)$. \square

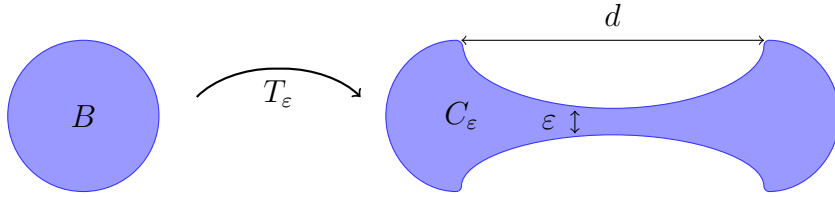
3.2 Proprietà di regolarità delle mappe ottimali

Sono possibili più strade per rendere rigorosa la soluzione euristica del problema isoperimetrico presentata in (3.2). Una possibilità consiste nel mostrare proprietà di regolarità della mappa ottimale per poter provare che i passaggi considerati sono effettivamente rigorosi. Per far questo si mostrerebbe innanzitutto che è verificata

in un qualche senso l'equazione $|\det \nabla T| = 1$, ed avendo così ottenuto che la mappa ottimale soddisfa un'equazione alle derivate parziali, si potrebbe sperare di ottenere le proprietà di regolarità desiderate; questo è in effetti possibile, ma la teoria di regolarità dell'equazione di Monge-Ampère (3.7) non è banale. L'Esempio seguente mostra come in generale non sia possibile sperare nemmeno nella continuità.

Esempio 7 (Caffarelli). Sia $B \subset \mathbb{R}^2$ la palla centrata nell'origine di raggio 1, e poniamo $B_+ := B \cap \{(x, y) \in \mathbb{R}^2 \mid x \geq 0\}$ e $B_- := B \cap \{(x, y) \in \mathbb{R}^2 \mid x \leq 0\}$. Considerando le misure $\mu := \omega_n^{-1} \mathcal{L}^n \llcorner B$ e $\nu := \omega_n^{-1} \mathcal{L}^n \llcorner C$ dove $C := (B_+ + (d, 0)) \cup (B_- + (-d, 0))$ con $d > 0$, abbiamo per il Teorema 8 che la mappa $u(x, y) = \frac{1}{2}(x^2 + y^2) + |x|$ è il potenziale convesso della mappa ottimale, ed è chiaro che il suo gradiente non è continuo.

In questo caso però la mancanza di continuità è data da ragioni topologiche. Per mostrare che è legata alla natura dell'equazione ragioniamo perturbando l'insieme di arrivo e sfruttando la stabilità e la monotonia delle mappe ottimali. Consideriamo quindi degli insiemi C_ε connessi tali che $|C_\varepsilon \Delta C| \rightarrow 0$. Ponendo $\nu_\varepsilon = \frac{1}{|C_\varepsilon|} \mathcal{L}^n \llcorner C_\varepsilon$ abbiamo che le mappe $T_\mu^{\nu_\varepsilon}$ sono discontinue per ε sufficientemente piccolo.



Per provare quest'ultima affermazione, notiamo innanzitutto che essendo nelle ipotesi del Teorema 9 abbiamo che $T_\varepsilon := T_\mu^{\nu_\varepsilon} \rightarrow T = \nabla u$ in L^p per ogni $p \in [1, \infty)$. Dato $\delta > 0$ consideriamo delle palle $D_+ = B((0, 1 - \delta), \delta)$, $D_- = B((0, \delta - 1), \delta)$. Poiché le mappe T_ε convergono a T allora se ε è sufficientemente piccolo esistono punti $w \in D_- \cap \{x > 0\}$ tali che $1 - (T_\varepsilon(w) - w) \cdot \hat{x} \ll 1$ ed esistono $w \in D_- \cap \{x < 0\}$ tali che $1 + (T_\varepsilon(w) - w) \cdot \hat{x} \ll 1$. Se T_ε fosse continua si potrebbe scegliere allora un punto $z_- \in D_-$ tale che $\frac{1}{2} - (T_\varepsilon(z_-) - z_-) \cdot \hat{x} \ll 1$. Analogamente, posso scegliere $z_+ \in D_+$ tale che $1 + (T_\varepsilon(z_+) - z_+) \cdot \hat{x} \ll 1$.

Se δ è sufficientemente piccolo abbiamo che $1 - (z_+ - z_-) \cdot \hat{y} \ll 1$ mentre $|T_\varepsilon(z_+) - T_\varepsilon(z_-)| \cos \frac{\pi}{8} + (T_\varepsilon(z_+) - T_\varepsilon(z_-)) \cdot (\cos \frac{\pi}{8} \hat{x} + \sin \frac{\pi}{8} \hat{y}) \ll 1$, quindi si avrebbe $\langle T_\varepsilon(z_+) - T_\varepsilon(z_-), z_+ - z_- \rangle < 0$ che contraddice la monotonia delle mappe di trasporto.

La strada scelta in questa trattazione è quindi differente, ed è un misto tra la precedente e la regolarizzazione per convoluzione delle mappe. Evitando di studiare come si comporta il push-forward rispetto alla convoluzione, nella sezione

seguinte mostriamo che la mappa ottimale soddisfa l'equazione (3.7) \mathcal{L}^n -quasi ovunque, mentre nella sezione 3.2.2 mostriamo, usando solamente il fatto che la mappa ottimale ha potenziale convesso, proprietà di regolarità della mappa ottimale che permettono di passare al limite le disuguaglianze ottenute con le funzioni approssimanti.

3.2.1 L'equazione di Monge-Ampère

La formula di cambio di variabili vale per le funzioni lipschitziane nella forma seguente (si veda [AFP]).

Proposizione 6 (Formula dell'area per funzioni lipschitziane). *Sia $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una funzione lipschitziana. Allora dato un insieme $E \subset \mathbb{R}^n$ \mathcal{L}^n -misurabile la mappa $y \mapsto \mathcal{H}^0(E \cap f^{-1}(y))$ è \mathcal{L}^n -misurabile e vale*

$$\int_{\mathbb{R}^n} \mathcal{H}^0(E \cap f^{-1}(y)) \, d\mathcal{L}^n(y) = \int_E |\det \nabla f(x)| \, d\mathcal{L}^n(x). \quad (3.4)$$

Questo risultato mette in luce il fatto che il differenziale di una funzione permette di controllare quanto la funzione alteri i volumi anche in ipotesi di regolarità più deboli che C^1 . Tuttavia, in generale le mappe ottimali sono definite su un insieme boreliano; è quindi necessario dare una definizione di differenziale nel caso in cui il dominio sia meno regolare di un aperto.

Definizione 12. Sia $D \subset \mathbb{R}^n$ un boreliano. Una funzione boreliana $f: D \rightarrow \mathbb{R}^m$ si dice differenziabile in $x \in D$ se esiste una funzione lineare $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ tale che

$$f(y) = f(x) + L \cdot (y - x) + o(|y - x|), \quad D \ni y \rightarrow x. \quad (3.5)$$

A priori questa definizione non è buona perché il differenziale in un punto potrebbe non essere unico, e quindi potrebbe non essere ben definita la mappa ∇f . Tuttavia è possibile mostrare che il differenziale è unico per \mathcal{L}^n -quasi ogni $x \in D$, e quindi la mappa ∇f è ben definita \mathcal{L}^n -quasi ovunque. Per i nostri scopi è pienamente sufficiente perché questa ambiguità non ha rilevanza per i risultati che vogliamo mostrare.

Infatti denotando con

$$\Theta^*(D, x) := \limsup_{r \rightarrow 0^+} \frac{\mathcal{L}^n(D \cap B(x, r))}{\mathcal{L}^n(B(x, r))}$$

la densità superiore di D in x si può mostrare (si veda ad esempio [EG]) che per \mathcal{L}^n -quasi ogni $x \in D$ si ha

$$\Theta^*(D, x) = 1,$$

pertanto il seguente lemma permette di concludere.

Lemma 4. *Siano $D \subset \mathbb{R}^n$ boreliano e $f: D \rightarrow \mathbb{R}^m$ boreliani, e supponiamo che f sia differenziabile in $x \in D$. Allora se il differenziale di f in x non è unico si deve avere*

$$\Theta^*(D, x) = 0.$$

Dimostrazione. Premettiamo una notazione: sia $v \in \mathbb{R}^n$ fissato; dati $x \in D$ e $\varepsilon, \delta > 0$ poniamo con $R_\delta^\varepsilon(x)$ il rettangolo centrato in x con lato parallelo a v lungo $2\varepsilon\delta$ e lati perpendicolari a v lunghi 2δ .

Si noti che basta mostrare l'ultima uguaglianza e che è sufficiente mostrarla per $m = 1$, perché il caso generale segue considerando una componente. Supponiamo ora che esistano $L \neq M \in \mathbb{R}^n$ che soddisfano (3.5), e fissiamo $\varepsilon > 0$. Poiché per differenza si ha $(L - M) \cdot (y - x) = o(|y - x|)$ per $D \ni y \rightarrow x$, allora esiste $\delta > 0$ tale che $|(L - M) \cdot (y - x)| < \varepsilon |x - y|$, per ogni $y \in B(x, \delta) \cap D$, quindi si avrebbe, con $v = L - M$, che $D \cap B(x, s) \subset R_s^\varepsilon$ per ogni $s < \delta$, quindi $\Theta^*(D, x) < c_n \varepsilon$ con c_n dipendente solo da n . Per l'arbitrarietà di ε si conclude. \square

La strategia per dimostrare che la mappa ottimale soddisfa la (3.7) è di mostrare che le funzioni convesse sono differenziabili due volte quasi ovunque e che è possibile approssimare le funzioni differenziabili quasi ovunque con funzioni lipschitziane (nel senso di Lusin) in modo da poter passare al limite la formula dell'area all'interno del dominio di differenziabilità. La dimostrazione è suddivisa in alcuni lemmi.

Il seguente classico teorema (si veda [EG]) è un risultato analogo al Teorema 7 per il second'ordine delle funzioni convesse che implica la differenziabilità \mathcal{L}^n -quasi ovunque della mappa di Brenier.

Teorema 10 (Aleksandrov). *Siano $\Omega \subset \mathbb{R}^n$ un aperto convesso e $u: \Omega \rightarrow \mathbb{R}$ una funzione convessa. Allora per \mathcal{L}^n -quasi ogni $x \in \Omega$ esiste una matrice $n \times n$ simmetrica e semidefinita positiva $\nabla^2 u(x)$ tale che*

$$u(y) = u(x) + \nabla u(x) \cdot (y - x) + \frac{1}{2}(y - x) \cdot \nabla^2 u(x)(y - x) + o(|y - x|^2), \quad y \rightarrow x.$$

Inoltre, la mappa $T := \nabla u$ è differenziabile \mathcal{L}^n -quasi ovunque e si ha $\nabla T = \nabla^2 u$ \mathcal{L}^n -quasi ovunque.

Lemma 5. *Sia $E \subset \mathbb{R}^n$ un boreliano e sia $f: E \rightarrow \mathbb{R}^n$ boreliana. Detto D l'insieme dei punti di differenziabilità di f , esistono una successione di boreliani $C_k \subset D$ e una successione di funzioni lipschitziane $f_k: \mathbb{R}^n \rightarrow \mathbb{R}^n$ tali che $C_k \nearrow D$ e $f_k = f$ su C_k .*

Dimostrazione. Poniamo

$$C_k := \{x \in D \mid |f(x)| \leq k, |\nabla f(x)| \leq k, \\ |f(y) - f(x)| \leq (1 + |\nabla f(x)|) |y - x| \quad \forall y \in B_{1/k}(x)\}.$$

Mostro prima che $C_k \nearrow D$. Sicuramente $C_k \subset C_{k+1}$, e fissato $x \in D$ sia δ tale che

$$|f(y) - f(x) - \nabla f(x) \cdot (y - x)| \leq |y - x| \quad \forall y \in B_\delta(x),$$

allora ponendo $k := \max \{|f(x)|, |\nabla f(x)|, \lceil 1/\delta \rceil\}$ si ha che $x \in C_k$, perché valgono chiaramente le prime due proprietà, e inoltre se $y \in B_{1/k}(x)$ allora per definizione $y \in B_\delta(x)$ e quindi vale $|y - x| - |\nabla f(x)| |x - y| \leq |x - y|$ da cui l'ultima richiesta.

Mostro ora che $f_k := f|_{C_k}$ è L_k -lipschitziana, con $L_k := \max \{k + 1, 2k^2\}$. Siano $x, y \in C_k$, allora se $|x - y| < 1/k$ per definizione si ha $|f(x) - f(y)| \leq (k+1) |x - y| \leq L_k |x - y|$, mentre altrimenti se $|x - y| \geq 1/k$ si ha $|f(x) - f(y)| \leq |f(x)| + |f(y)| \leq 2k \leq 2k^2 |x - y| \leq L_k |x - y|$.

A questo punto è possibile estendere le funzioni $f|_{C_k}$ a $f_k: \mathbb{R}^n \rightarrow \mathbb{R}^n$ mantenendo la lipschitzianità. Infatti si può ragionare componente per componente, e dati un boreliano $A \subset \mathbb{R}^n$ boreliano e $g: A \rightarrow \mathbb{R}$ lipschitziana, la si può estendere ponendo

$$f_k^i(x) := \inf \{g(y) + L_k |x - y| \mid y \in A\}, \quad L_k = \text{Lip}(g).$$

La f_k^i è lipschitziana perché estremo inferiore di funzioni equilipschitziane, inoltre essendo g L_k -lipschitziana allora $g(y) + L_k |x - y| \geq g(x)$ e quindi $y = x$ è un minimo quando $x \in A$, pertanto $f_k^i(x) = g(x)$. \square

Lemma 6. *Siano $E \subset \mathbb{R}^n$ e $f: E \rightarrow \mathbb{R}^n$ boreliani, allora denotando con $\Sigma := \{x \in E \mid \exists \nabla f(x) \text{ e } \det \nabla f(x) = 0\}$ si ha $\mathcal{L}^n(\Sigma) = 0$.*

Inoltre, supponiamo che esista un insieme $B \subset E$ tale che $f|_B$ sia differenziabile e iniettiva. Allora l'insieme $f(B)$ è misurabile e data una funzione $\varphi: E \rightarrow [0, \infty]$ vale

$$\int_{f(B)} \varphi(y) d\mathcal{L}^n(y) = \int_B \varphi \circ f(x) |\det \nabla f(x)| d\mathcal{L}^n(x). \quad (3.6)$$

Dimostrazione. Consideriamo gli insiemi C_k e le funzioni f_k dati dal Lemma 5 applicato a $f|_B$. Allora $\nabla f_k(x) = \nabla f(x)$ per \mathcal{L}^n -quasi ogni $x \in C_k$; infatti, essendo f_k lipschitziana allora i differenziali esistono \mathcal{L}^n -quasi ovunque in C_k , ed essendo f_k un'estensione di $f|_{C_k}$ ed essendo il differenziale unico per quasi ogni punto allora i due differenziali devono coincidere \mathcal{L}^n -quasi ovunque in C_k .

Mostriamo la prima parte dell'enunciato: posto $\Sigma_k := \Sigma \cap C_k$, applicando la (3.4) agli insiemi Σ_k si ottiene $\mathcal{L}^n(f(\Sigma_k)) = 0$, e poiché $\Sigma_k \nearrow \Sigma$ si conclude per continuità della misura.

Per la misurabilità di $f(B)$ poniamo $B_k := B \cap C_k$. Per iniettività di $f|_B$ e per la Proposizione 6 usata con f_k e B_k si ha che B_k è misurabile, ma poiché $B_k \nearrow B$ allora $f(B_k) \nearrow f(B)$ e quindi anche $f(B)$ lo è.

Per l'ultima parte consideriamo prima il caso in cui $\varphi = \mathbb{1}_D$. Poiché per le funzioni f_k vale la (3.4), applicandola a $B_k \cap f^{-1}(D)$ e ricordando l'iniettività di

$f|_B$ si ottiene

$$\int_{f(B_k)} \mathbb{1}_D(y) \, d\mathcal{L}^n(y) = \int_{B_k} \mathbb{1}_D \circ f(x) |\det \nabla f(x)| \, d\mathcal{L}^n(x),$$

da cui, poiché $B_k \nearrow B$, per convergenza monotona si ottiene la (3.6). Se φ è semplice si ragiona per linearità, mentre nel caso generale si approssima con funzioni semplici e si conclude per convergenza monotona. \square

Nel lemma precedente si è integrato solamente su un insieme in cui la funzione era differenziabile. La ragione per cui nella Proposizione 6 si considerano solo funzioni lipschitziane è il fatto che l'immagine di insiemi \mathcal{L}^n -trascurabili resta \mathcal{L}^n -trascurabile, cosa che non è garantita invece per f generica (neppure per f continua senza maggiore regolarità).

Questa difficoltà non si pone però se si vuole calcolare la densità della misura push-forward della mappa ottimale, cosa che permette di ottenere il seguente risultato.

Teorema 11. *Sia $X \subset \mathbb{R}^n$ limitato, e siano $\mu, \nu \in \mathcal{P}(X)$ tali che $\mu = \rho \mathcal{L}^n$, $\nu = \eta \mathcal{L}^n$. Sia $T_\mu^\nu = \nabla \phi$ la mappa ottimale tra μ e ν . Allora*

$$\eta(\nabla \phi) \det \nabla^2 \phi = \rho \quad \mathcal{L}^n\text{-quasi ovunque.} \quad (3.7)$$

Dimostrazione. Denotiamo con $E \subset \text{dom } u$ l'insieme dei punti di differenziabilità di ϕ , con $D \subset E$ l'insieme dei punti di differenziabilità di T e con $\Sigma := \{x \in D \mid \det \nabla T(x) = 0\}$. Ponendo $B := D \setminus \Sigma$ abbiamo che $T|_B$ è differenziabile, e mostriamo ora che è anche iniettiva.

Innanzitutto, se f è una funzione convessa definita sulla retta reale tale che $f''(0) > 0$ allora si ha che esiste $\varepsilon > 0$ tale che $f'(1) - f'(0) > \varepsilon f''(0)$, perché per definizione di limite esiste un $\varepsilon > 0$ tale che $f'(t) > f'(0) + t f''(0)/2$ per $0 < t < 2\varepsilon$, e per monotonia della derivata si conclude.

Pertanto, considerando la funzione $f(t) := \phi(x + t(y - x))$ con $x \notin \Sigma$ abbiamo che $f''(0) = \langle y - x, \nabla^2 \phi \cdot (y - x) \rangle > 0$ per non degenerazione della forma bilineare $\nabla^2 \phi$, e quindi $\langle T(y) - T(x), y - x \rangle > 0$, da cui l'iniettività di T ristretta a $D \setminus \Sigma$.

Mostriamo ora che, data $\varphi: \mathbb{R}^n \rightarrow [0, \infty]$, vale

$$\int_{\mathbb{R}^n} \varphi \, d\nu = \int_{\mathbb{R}^n} \varphi \circ T \eta \circ T |\det \nabla T| \, d\mathcal{L}^n. \quad (3.8)$$

Per farlo mostriamo prima che $\nu(T(B)^c) = \mu(B^c) = 0$. Infatti si ha che $\nu(T(B)^c) = \mu(T^{-1}(T(B)^c)) \leq \mu(B^c)$, quindi basta che $\mu(B^c) = 0$. Per il Teorema 10 si ha che $\mu(E \setminus D) = \mathcal{L}^n(E \setminus D) = 0$, e mostriamo ora che $\mathcal{L}^n(\Sigma) = 0$. Notiamo che per

la prima parte del Lemma 6 abbiamo $\mathcal{L}^n(T(\Sigma)) = 0$. Poiché $\nu \ll \mathcal{L}^n$ allora deve essere $\nu(T(\Sigma)) = 0$, ma per definizione $0 = \nu(T(\Sigma)) = \mu(T^{-1}(T(\Sigma))) \geq \mu(\Sigma)$.

Pertanto la (3.8) si riduce a mostrare che

$$\int_{T(B)} \varphi \eta \, d\mathcal{L}^n = \int_B \varphi \circ T \eta \circ T |\det \nabla T| \, d\mathcal{L}^n,$$

e questo segue dal Lemma 6.

Inoltre, data $\varphi: \mathbb{R}^n \rightarrow [0, \infty]$, per la (1.1) la condizione $T_{\#}\mu = \nu$ si può scrivere come

$$\int_{\mathbb{R}^n} \varphi \, d\nu = \int_{\mathbb{R}^n} \varphi \circ T \rho \, d\mathcal{L}^n,$$

che insieme alla (3.8) implica

$$\int_{\mathbb{R}^n} \varphi \circ T \eta \circ T |\det \nabla T| \, d\mathcal{L}^n = \int_{\mathbb{R}^n} \varphi \circ T \rho \, d\mathcal{L}^n. \quad (3.9)$$

Mostriamo infine che questa relazione implica la tesi. Infatti poiché $\mu(B^c) = \nu(T(B)^c) = 0$ è sufficiente mostrare che dato un boreliano $F \subset B$ vale

$$\int_F \eta \circ T |\det \nabla T| \, d\mathcal{L}^n = \int_F \rho \, d\mathcal{L}^n,$$

ma poiché $F \subset B$ applicando il Lemma 6 si ha che $T(F)$ è misurabile e quindi ponendo $\varphi = \mathbb{1}_{T(F)}$ nella (3.9) si ottiene la tesi. \square

3.2.2 Regolarità BV delle mappe ottimali

Se il differenziale di una mappa permette di controllare come tale mappa altera i volumi anche con ipotesi più deboli di regolarità, non è possibile estendere a questi casi risultati importanti come la formula di integrazione per parti, così come non è sempre vero che le derivate di approssimanti convergono al differenziale. Questo fenomeno si riscontra già in dimensione 1, come mostra l'Esempio seguente.

Esempio 8. Data $f: \mathbb{R} \rightarrow \mathbb{R}$ convessa, si è visto (come conseguenza della Proposizione 3 e del Teorema 7) che esiste quasi ovunque f' ed è una funzione monotona crescente.

Per il Teorema di Aleksandrov la derivata f' è differenziabile \mathcal{L}^1 -quasi ovunque, tuttavia non è detto che valga il teorema fondamentale del calcolo per f' e f'' ; infatti considerando ad esempio

$$u(x) = \begin{cases} 0 & x \leq 0, \\ x & x > 0, \end{cases}$$

abbiamo $u'(x) = \mathbb{1}_{(0, \infty)}$ e quindi $u'' = 0$, quindi u' non è costante pur avendo derivata nulla.

Poiché queste proprietà sono necessarie per rendere rigorosa la (3.2), è necessario considerare la derivata debole, o distribuzionale, delle mappe ottimali. In particolare, usando il fatto che derivano da un potenziale convesso, mostriamo la regolarità BV delle mappe ottimali.

Ricordiamo la definizione e alcune proprietà delle distribuzioni. Per una dimostrazione si rimanda a [F].

Definizione 13 (Distribuzioni). Dato un aperto $\Omega \subset \mathbb{R}^n$, diciamo che $(\phi_k) \subset C_c^\infty(\Omega)$ converge a ϕ in C_c^∞ se esiste un compatto $K \subset \Omega$ tale che $\text{supp } \phi_k \subset K$ e se $\partial^\alpha \phi_k \rightarrow \partial^\alpha \phi$ uniformemente per ogni multiindice α .

Dato $\Omega \subset \mathbb{R}^n$ aperto, una distribuzione su Ω è un funzionale lineare $T: C_c^\infty(\Omega) \rightarrow \mathbb{R}$ continuo rispetto alla convergenza in C_c^∞ , i.e. $T(\phi_k) \rightarrow T(\phi)$ se $\phi_k \rightarrow \phi$ in C_c^∞ .

Indichiamo con $\mathcal{D}'(\Omega)$ lo spazio delle distribuzioni su Ω . Le distribuzioni sono funzioni generalizzate nel senso che se $f \in L_{\text{loc}}^1(\Omega)$ allora la mappa $\phi \mapsto \int_\Omega f \phi \, dx$ è una distribuzione; anche una misura di Radon localmente finita μ induce la distribuzione $\phi \mapsto \int_\Omega \phi \, d\mu$. Identificheremo una funzione f o una misura μ con la distribuzione T_f, T_μ che induce.

Definizione 14 (Derivata distribuzionale). Dato un multiindice α e una distribuzione $T \in \mathcal{D}'(\Omega)$, la derivata distribuzionale (o debole) α -esima di T è la distribuzione $D^\alpha T$ definita da

$$D^\alpha T: \phi \longmapsto (-1)^{|\alpha|} T(\partial^\alpha \phi).$$

La derivata debole è compatibile con la derivata classica, nel senso che se $f \in C^\alpha$ allora $D^\alpha T_f = T_{\partial^\alpha f}$. Inoltre, dato un vettore $v \in \mathbb{R}^n$, si denoterà con $D_v T$ la derivata di T lungo v , definita da $D_v T(\phi) = -T(\partial_v \phi)$. Se e_i, e_j sono vettori della base canonica, indichiamo con $D_{i,j} T = D_{e_i} D_{e_j} T = D_{e_j} D_{e_i} T$.

È utile estendere la convoluzione di funzioni alle distribuzioni, perché permette di dimostrare risultati per approssimazione. Premettiamo alcune notazioni.

Sia $\Omega \subset \mathbb{R}^n$, dato $\delta > 0$ indichiamo con $\Omega_\delta := \{x \in \Omega \mid d(x, \partial\Omega) > \delta\}$, mentre con B_δ si intende la palla di raggio δ centrata nell'origine. Data una funzione $f: \Omega \rightarrow \mathbb{R}$ e un punto $x \in \mathbb{R}^n$, indichiamo con $\tau_x f$ la funzione $y \mapsto f(y - x)$.

Proposizione 7 (Convoluzione di distribuzioni). *Sia $\Omega \subset \mathbb{R}^n$ un aperto e siano $\psi \in C_c^\infty(B_\delta)$ e $T \in \mathcal{D}'(\Omega)$. Allora $\psi * T(x) = T(\tau_x \psi)$ è ben definita per $x \in \Omega_\delta$ e inoltre*

$$i) \quad T * \psi \in C^\infty(\Omega_\delta);$$

$$ii) \quad \text{per ogni multiindice } \alpha \text{ si ha } \partial^\alpha (T * \psi) = T * \partial^\alpha \psi = D^\alpha T * \psi;$$

iii) per ogni $\phi \in C_c^\infty(\Omega_\delta)$ si ha

$$\int_{\Omega_\delta} (T * \psi)\phi \, d\mathcal{L}^n = T(\phi * \psi(-\cdot)).$$

La Proposizione precedente permette quindi di definire la convoluzione tra una distribuzione e una funzione C_c^∞ . Tale operazione è compatibile con la convoluzione di funzioni, nel senso che se $f \in L_{\text{loc}}^1(\Omega)$ allora $\psi * T_f = T_{\psi * f}$. Inoltre, come nel caso classico tale operazione è bilineare.

Infine, è possibile approssimare puntualmente una distribuzione $T \in \mathcal{D}'(\Omega)$ con funzioni regolari, come mostra il Lemma seguente.

Lemma 7. *Siano $T \in \mathcal{D}'(\Omega)$ e $\psi \in C_c^\infty(B_1)$ tali che $\psi \geq 0$ e $\int_{\mathbb{R}^n} \psi \, dx = 1$, e poniamo $\psi_\varepsilon(x) := \varepsilon^{-n} \psi(\varepsilon^{-1}x)$. Allora, data $\phi \in C_c^\infty(\Omega)$, se ε è sufficientemente piccolo è ben definito $\psi_\varepsilon * T(\phi)$ e*

$$\lim_{\varepsilon \rightarrow 0^+} T_{\psi_\varepsilon * T}(\phi) = T(\phi). \quad (3.10)$$

Dimostrazione. Innanzitutto mostriamo che se $\phi \in C_c^0(\mathbb{R}^n)$ allora $\psi_\varepsilon * \phi \rightarrow \phi$ uniformemente per $\varepsilon \rightarrow 0^+$. Infatti, detto ω il modulo di continuità di ϕ e dato $x \in \mathbb{R}^n$ si ha

$$|\phi(x) - \phi * \psi_\varepsilon(x)| \leq \int_{B_1} \psi(y) |\phi(x) - \phi(x - \varepsilon y)| \, dy \leq \omega(\varepsilon),$$

quindi se $\phi \in C_c^\infty(\Omega)$ allora $\psi_\varepsilon * \phi$ converge a ϕ in C_c^∞ . Data $\phi \in C_c^\infty(\Omega)$, poiché si ha $\text{supp } \psi_\varepsilon \subset B_\varepsilon$, abbiamo che per ε abbastanza piccolo è ben definito $\psi_\varepsilon * T(\phi)$.

Infine, per la Proposizione 7 si ha

$$T_{\psi_\varepsilon * T}(\phi) = T(\phi * \psi(-\cdot)),$$

da cui passando al limite per $\varepsilon \rightarrow 0^+$ si conclude. \square

Considerando l'Esempio 8, si ha che la derivata f'' non è la derivata debole seconda di f . Tuttavia, essendo f' crescente allora è la funzione di ripartizione (a meno di costanti) di una misura μ positiva finita sui limitati (nel caso precedente $\mu = \delta_0$), quindi si ha $f'(b) - f'(a) = \mu((a, b])$, e quindi $D(f') = \mu$.

In dimensione maggiore vale un risultato analogo, enunciato nel Teorema 12. Premettiamo due lemmi, per la cui dimostrazione rimandiamo a [EG].

Lemma 8. *Sia $\Omega \subset \mathbb{R}^n$ aperto e sia $T \in \mathcal{D}'(\Omega)$ una distribuzione positiva, ovvero tale che $T(\phi) \geq 0$ per ogni $\phi \in C_c^\infty(\Omega)$ con $\phi \geq 0$. Allora esiste una misura di Radon non negativa μ tale che*

$$T(\phi) = \int_{\Omega} \phi \, d\mu \quad \text{per ogni } \phi \in C_c^\infty(\Omega).$$

Il seguente enunciato è un risultato di differenziazione di misure che generalizza i teoremi fondamentali del calcolo.

Lemma 9 (Lebesgue-Radon-Nikodym). *Siano $X \subset \mathbb{R}^n$ e $\mu \in \mathcal{M}_+(X)$, allora esistono $\rho \in L^1_{\text{loc}}(X)$ e $\lambda \in \mathcal{M}_+(X)$ tali che $\lambda \perp \mathcal{L}^n$ e $\mu = \rho \mathcal{L}^n + \lambda$. Inoltre*

$$\rho(x) = \lim_{r \rightarrow 0^+} \frac{\mu(B(x, r))}{\mathcal{L}^n(B(x, r))} \quad \text{per } \mathcal{L}^n\text{-quasi ogni } x \in X. \quad (3.11)$$

Teorema 12. *Sia $\Omega \subset \mathbb{R}^n$ un aperto convesso, e sia $u: \Omega \rightarrow \mathbb{R}$ una funzione convessa. Allora per ogni $i, j = 1, \dots, n$ esiste una misura di Radon a variazione localmente finita $\mu_{i,j}$ tale che $D_{i,j}u = \mu_{i,j}$, i.e. $\int_{\Omega} u \partial_{i,j} \phi \, d\mathcal{L}^n = \int_{\Omega} \phi \, d\mu_{i,j}$ per ogni $\phi \in C_c^\infty(\Omega)$. Inoltre, μ_{ii} è non negativa.*

Infine, con le stesse notazioni del Teorema 10, si ha che $D^2u = \nabla^2 u \mathcal{L}^n + D_s^2 u$, con $D_s^2 u \perp \mathcal{L}^n$. Inoltre, $(D_s^2 u)_{ii} \in \mathcal{M}_+$.

Dimostrazione. Siano $1 \leq i, j \leq n$ interi. Ponendo $v = e_i + e_j$, $w = e_i - e_j$ abbiamo

$$D_{i,j}u = D_{\frac{v+w}{2}} D_{\frac{v-w}{2}} u = \frac{1}{4} (D_v^2 u - D_w^2 u),$$

pertanto grazie al Lemma 8 per la prima parte è sufficiente mostrare che dato $v \in \mathbb{R}^n$ la distribuzione $D_v^2 u$ è positiva.

Innanzitutto, se $u \in C^\infty$ allora la tesi è vera perché $D_v^2 u = \partial_v^2 u \mathcal{L}^n \geq 0$. Essendo la proprietà $\partial_v^2 u \geq 0$ una proprietà locale è sufficiente che u sia localmente convessa, ovvero che per ogni punto $x \in \Omega$ esista un intorno aperto $U(x)$ tale che $u|_{U(x)}$ sia convessa.

Data u generica, consideriamo dei regolarizzatori ψ_ε presi come nel Lemma 7. Ponendo $f_\varepsilon = u * \rho_\varepsilon$ abbiamo che $f_\varepsilon \in C^\infty(\Omega_\varepsilon)$ per la Proposizione 7 e inoltre sono localmente convesse. Infatti, sia $x \in \Omega_\varepsilon$, allora se $s, t \in B_r(x) \subset \Omega_\varepsilon$ per convessità di f , $f(\lambda s + (1 - \lambda)t - y) \leq \lambda f(s - y) + (1 - \lambda)f(t - y)$ per ogni $y \in B_\varepsilon$, quindi moltiplicando per $\psi_\varepsilon(y)$ e integrando rispetto a $y \in B_\varepsilon$ otteniamo $f_\varepsilon(\lambda s + (1 - \lambda)t) \leq \lambda f_\varepsilon(s) + (1 - \lambda)f_\varepsilon(t)$.

Pertanto data $\phi \in C_c^\infty(\Omega)$ per la (3.10) si ha

$$\int_{\Omega} f \partial_v^2 \phi \, d\mathcal{L}^n = \lim_{\varepsilon \rightarrow 0^+} \int_{\Omega} f_\varepsilon \partial_v^2 \phi \, d\mathcal{L}^n = \lim_{\varepsilon \rightarrow 0^+} \int_{\Omega} \partial_v^2 f_\varepsilon \phi \, d\mathcal{L}^n \geq 0.$$

Per la seconda parte abbiamo che per il Lemma 9 si ha che esistono ρ, λ tali che $\mu_{i,j} = \rho \mathcal{L}^n + \lambda$. Per concludere è sufficiente mostrare che $\rho = (\nabla^2 u)_{ij} \mathcal{L}^n$ -quasi ovunque. Per mostrarlo consideriamo $\phi \in C_c^\infty(B(x, r))$. Allora per il Teorema 10

per \mathcal{L}^n -quasi ogni x si ha

$$\begin{aligned} \int_{B(x,r)} \phi \, d\mu &= - \int_{B(x,r)} T_j(y) \partial_i \phi(y) \, dy \\ &= -T_j(x) \int_B \partial_i \phi \, dy - \int_B \nabla T_j(x) \cdot (y-x) \partial_i \phi \, dy + o(r^n) \\ &= (\nabla T_j)_i \int_B \phi \, dy + o(r^n). \end{aligned}$$

Poiché $\mu(B(x,r)) = \sup_{\phi} \int_B \phi \, d\mu$ e analogamente per \mathcal{L}^n , dove $\phi \in C_c^\infty(B(x,r))$ e $|\phi| \leq 1$, dalla (3.11) segue la tesi. \square

Si noti che in realtà si dimostra il Teorema 10 a partire dal precedente, in quanto presa la densità della derivata seconda rispetto alla misura di Lebesgue si mostra che tale funzione è effettivamente il differenziale del gradiente.

Osservazione 3. Dalla dimostrazione del Teorema 12 emerge come si possa scrivere $D_{ij}u = 1/4(D_v^2u - D_w^2u)$, con $v = e_i + e_j$, $w = e_i - e_j$. Lo stesso vale anche per le parti singolari e assolutamente continue, quindi se $(D_v^2u)_s = 0$ per ogni $v \in \mathbb{R}^n$ allora si ha $(D^2u)_s = 0$.

Corollario 2 (Regolarità BV). *Sia $E \subset \mathbb{R}^n$ un aperto limitato con bordo C^1 , e sia B la palla centrata nell'origine tale che $\mathcal{L}^n(B) = \mathcal{L}^n(E)$. Posto $\mu = \mathcal{L}^n \llcorner E$ e $\nu = \mathcal{L}^n \llcorner B$, allora $T_\mu^\nu \in BV(E)$.*

Dimostrazione. Per il Teorema 8 abbiamo che la mappa T è indotta da un potenziale convesso u , i.e. $T = \nabla u$. Ma poiché $|T| \leq R$, con R il raggio della palla B , per l'Osservazione 2 abbiamo che $u: \mathbb{R}^n \rightarrow \mathbb{R}$, ovvero $\text{dom } u = \mathbb{R}^n$. Pertanto per il Teorema 12 si ha che $DT = D^2u$ è una misura vettoriale con ogni componente localmente finita. Essendo E limitato abbiamo che esistono due misure vettoriali finite μ_E^+, μ_E^- tali che $|D_i T_j|(E) = |\mu_E^+(E) - \mu_E^-(E)| \leq \mu_E^+(E) + \mu_E^-(E) < \infty$ e quindi $T \in BV(E)$. \square

3.3 La disuguaglianza isoperimetrica

Con i risultati ottenuti finora è possibile rendere rigorosa la soluzione del problema (P) esposta precedentemente. Si ricorda che si è enunciato il problema nella classe di aperti con bordo C^1 , di cui di seguito diamo la definizione.

Definizione 15. Sia $E \subset \mathbb{R}^n$ un aperto. Si dice che il bordo è C^1 se per ogni punto $x \in \partial E$ esiste un intorno $U(x)$ di x e delle coordinate opportune tali che esistano un aperto $V \subset \mathbb{R}^{n-1}$ e una funzione $\phi \in C^1(V; \mathbb{R})$ tale che $\text{graph}(\phi) = \partial E \cap U(x)$.

Si noti come il bordo di un aperto con bordo C^1 è localmente un'ipersuperficie regolare. È quindi ben definita la misura $(n - 1)$ -dimensionale del bordo dell'insieme, che se è limitato è anche finita.

Possiamo quindi mostrare l'ottimalità della sfera per il problema **(P)**.

Teorema 13. *Sia $E \subset \mathbb{R}^n$ un aperto con bordo compatto C^1 tale che $\mathcal{L}^n(E) < \infty$, sia $B \subset \mathbb{R}^n$ la palla centrata nell'origine tale che $\mathcal{L}^n(B) = \mathcal{L}^n(E)$. Allora*

$$\sigma_{n-1}(B) \leq \sigma_{n-1}(E). \quad (3.12)$$

Inoltre vale l'uguaglianza se e solo se esiste $x \in \mathbb{R}^n$ tale che $E = x + B$.

Diamo due dimostrazioni di questo risultato. Entrambe si basano sulla soluzione formale esposta in (3.2). Se la prima utilizza meno risultati di teoria geometrica della misura, la seconda permette di caratterizzare i casi di uguaglianza.

Prima dimostrazione. Poniamo

$$\mu := \frac{1}{\mathcal{L}^n(E)} \mathcal{L}^n \llcorner E, \quad \nu := \frac{1}{\mathcal{L}^n(B)} \mathcal{L}^n \llcorner B.$$

Allora μ, ν soddisfano le ipotesi del Teorema 8 e quindi esiste un'unica mappa di trasporto ottimale T_μ^ν con u come potenziale di Kantorovich, i.e. $T = \nabla u$. Poiché la (3.12) è omogenea si può supporre che $B \subset B(0, 1)$. Allora dall'Osservazione 2 abbiamo che $u: \mathbb{R}^n \rightarrow \mathbb{R}$ è 1-lipschitziana. Preso quindi dei regolarizzatori ψ_ε come nel Lemma 7, poniamo $u_\varepsilon := u * \psi_\varepsilon$ e $T_\varepsilon := \nabla u_\varepsilon = T * \psi_\varepsilon$.

Essendo T_ε regolari abbiamo che vale

$$\int_E \operatorname{div} T_\varepsilon \, d\mathcal{L}^n = \int_{\partial E} \nu_E \cdot T_\varepsilon \, d\sigma_{n-1} \leq \sigma_{n-1}(E), \quad (3.13)$$

dove l'ultima disuguaglianza vale perché, essendo u 1-lipschitziana, si ha $|T_\varepsilon| \leq 1$.

Per la Proposizione 7 e per linearità abbiamo che vale

$$\operatorname{div} T_\varepsilon = (\operatorname{tr} D^2 u) * \psi_\varepsilon = \operatorname{tr}(\nabla^2 u) \mathcal{L}^n * \psi_\varepsilon + \operatorname{tr}(D_s^2 u) * \psi_\varepsilon \geq \operatorname{tr}(\nabla^2 u) * \psi_\varepsilon,$$

dove l'ultima disuguaglianza segue dalla seconda parte del Teorema 12. Per il Teorema 10 abbiamo che $\nabla^2 u$ è una matrice simmetrica semidefinita positiva, quindi usando la disuguaglianza tra media aritmetica e media geometrica agli autovalori (reali non negativi) di questa matrice si ottiene $\operatorname{tr}(\nabla^2 u) \geq n \det(\nabla^2 u)^{1/n}$, che per la (3.7) diventa $\operatorname{tr}(\nabla^2 u) \geq n$ in E , da cui si ottiene infine $\operatorname{div} T_\varepsilon \geq n + o(1)$ in E . Pertanto, sostituendo nella (3.13) e passando al limite per $\varepsilon \rightarrow 0$ si ottiene

$$\sigma_{n-1}(E) \geq \int_E \operatorname{div} T_\varepsilon \, d\mathcal{L}^n \geq n \mathcal{L}^n(E) = n \mathcal{L}^n(B) = \sigma_{n-1}(B). \quad \square$$

Premettiamo alla seconda dimostrazione un risultato di teoria geometrica della misura, che garantisce la validità della formula di Gauss-Green sotto ipotesi di minore regolarità. Rimandiamo a [AFP] per la dimostrazione.

Lemma 10 (Traccia e teorema della divergenza). *Sia $\Omega \subset \mathbb{R}^n$ un aperto limitato con bordo C^1 . Allora esiste un operatore $\nu: BV(\Omega) \rightarrow L^1(\partial\Omega, \sigma_{n-1})$ limitato tale che per σ_{n-1} -quasi ogni $x \in \Omega$ si ha*

$$\lim_{r \rightarrow 0^+} \int_{\Omega \cap B(x,r)} |f - f(x)| d\sigma_{n-1} = 0.$$

Inoltre vale

$$\int_{\partial\Omega} \nu(f) \cdot \nu_\Omega d\sigma_{n-1} = \text{tr } Df(E). \quad (3.14)$$

Seconda dimostrazione. Poniamo $\mu = \mathcal{L}^n \llcorner E$ e $\nu = \mathcal{L}^n \llcorner B$. Possiamo supporre che $B = B_1$. Per il Teorema 8 esiste una funzione convessa 1-lipschitziana $u: \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $T = \nabla u$ è la mappa di trasporto ottimale da μ a ν . Poiché u è 1-lipschitziana allora $|T| \leq 1$. Per il Corollario 2 la mappa T è a variazione limitata su E , e per il Lemma 10 è ben definito $\nu(T)$ e si ha $|\nu(T)| \leq 1$.

Inoltre, per il Teorema 10 la mappa T è differenziabile \mathcal{L}^n -quasi ovunque e ∇T è una matrice simmetrica semidefinita positiva. Usando la disuguaglianza tra media aritmetica e media geometrica sugli autovalori (reali non negativi) della matrice ∇T si ottiene che $\text{tr}(\nabla T) \geq n \det \nabla T^{1/n}$ che per la (3.7) diventa $\text{tr}(\nabla T) \geq n$. Infine, per il Teorema 12 si ha che $\text{div}(T) \geq \text{tr}(\nabla T) \mathcal{L}^n$ perché $(D_s^2 u)_{ii} \geq 0$. Pertanto, per la (3.14) si ottiene

$$\sigma_{n-1}(\partial B) = n \mathcal{L}^n(E) \leq \int_E \text{div } T \leq \text{tr } DT(E) = \int_{\partial E} \nu(T) \cdot \nu_E d\sigma_{n-1} \leq \sigma_{n-1}(\partial E).$$

Infine, se vale l'uguaglianza nella (3.12) allora ogni disuguaglianza vista prima è un'uguaglianza, in particolare la disuguaglianza tra media aritmetica e geometrica. Questo implica che gli autovalori di ∇T sono uguali, da cui ricordando la (3.7) $\nabla T = Id$. Inoltre $(D_s^2 u)_{ii} = 0$; poiché il ragionamento è valido per qualunque scelta di coordinate per l'Osservazione 3 $(D_v^2 u)_s = 0$ per ogni v e quindi $(D^2 u)_s = 0$. Esistono allora $b \in \mathbb{R}$ e $a \in \mathbb{R}^n$ tali che $u(x) = |x|^2/2 + a \cdot x + b$, da cui la tesi. \square

Un'immediata conseguenza del Teorema 13 è la disuguaglianza isoperimetrica.

Corollario 3 (Disuguaglianza isoperimetrica). *Sia $E \subset \mathbb{R}^n$ un aperto limitato con bordo C^1 . Allora*

$$\mathcal{L}^n(E) \leq C_n \sigma_{n-1}(\partial E)^{\frac{n}{n-1}},$$

con C_n come nella (3.1). Inoltre vale l'uguaglianza se e solo se E è una palla.

Dimostrazione. Sia B la palla centrata nell'origine tale che $\mathcal{L}^n(B) = \mathcal{L}^n(E)$. Poiché il problema è invariante per riscalamento si può supporre che $B = B_1$. Per il Teorema 13 si ha $\sigma_{n-1}(B) \leq \sigma_{n-1}(E)$, pertanto

$$\mathcal{L}^n(E) = \mathcal{L}^n(B) = |B| \sigma_{n-1}(B)^{n/(1-n)} \sigma_{n-1}(B)^{n/(n-1)} \leq C_n \sigma_{n-1}(E)^{n/(n-1)},$$

dove si è usato che $\sigma_{n-1}(\partial B) = n |B_1|$. Inoltre vale l'uguaglianza se e solo se E è una palla. \square

3.3.1 Cenni alle disuguaglianze isoperimetriche quantitative

È lecito chiedersi se un insieme che sia "quasi ottimale" per il problema isoperimetrico è anche "quasi un cerchio", domanda che è legata alla stabilità del problema isoperimetrico. La risposta a questo problema è data dalle disuguaglianze isoperimetriche quantitative, che permettono di stimare la distanza di un insieme E dall'essere un cerchio con la distanza dall'essere ottimale nel problema (P).

Precisiamo cosa si intenda con essere "quasi ottimale" e "quasi un cerchio".

Definizione 16 (Indice di asimmetria e deficit isoperimetrico). Sia $E \subset \mathbb{R}^n$ un aperto limitato con bordo C^1 . L'indice di asimmetria è definito come

$$A(E) := \inf_{x \in \mathbb{R}^n} \frac{|E \Delta (B+x)|}{|E|},$$

dove B è la palla centrata nell'origine tale che $\mathcal{L}^n(B) = \mathcal{L}^n(E)$. Il deficit isoperimetrico è

$$\delta(E) := \frac{\sigma_{n-1}(E)}{C_n \mathcal{L}^n(E)^{(n-1)/n}} - 1.$$

Siamo interessati a disuguaglianze del tipo $A(E) \leq C\delta(E)^\alpha$, e in particolare al miglior valore possibile dell'esponente per $\delta(E) \rightarrow 0$. Considerando degli ellissi con piccola eccentricità si può vedere che l'esponente non può essere più grande di $1/2$, e la congettura di Hall è che effettivamente il valore $\alpha = 1/2$ sia ottimale. Raffinando la dimostrazione presentata nel paragrafo precedente con ulteriori idee (non banali) è possibile mostrare la congettura di Hall. In particolare vale il seguente risultato (si veda [FMP]).

Teorema 14. *Esiste una costante $C_n > 0$, dipendente soltanto dalla dimensione, tale che per ogni $E \subset \mathbb{R}^n$ aperto limitato con bordo C^1 vale*

$$A(E) \leq C_n \delta(E)^{\frac{1}{2}}.$$

In questa sezione illustriamo senza entrare nei dettagli un modo (presentato sempre in [FMP]) per mostrare un risultato non ottimale, ovvero che esiste una costante $C > 0$ tale che $A(E) \leq C\delta(E)^{1/4}$, nella classe più ristretta degli insiemi convessi e in dimensione 2.

L'idea è di stimare dapprima la distanza L^1 della derivata ∇T dall'identità con il deficit isoperimetrico, ovvero

$$\int_E |\nabla T - Id| \leq C\delta(E)^{1/2}.$$

Questo è possibile in quanto, ripercorrendo la seconda dimostrazione del Teorema 13, si ottiene

$$\int_E \left(\frac{\operatorname{div} T}{n} - \det \nabla T^{1/n} \right) d\mathcal{L}^n \leq C_1\delta(E);$$

ricordando che se $\delta(E) = 0$ questo implica che $\nabla T = Id$ si vede che è ragionevole dimostrare da questo la stima precedente.

Il passo successivo consiste nello stimare in modo indiretto la quantità $A(E)$ tramite la distanza L^p di T dall'identità per un opportuno $p \geq 1$, grazie alla disuguaglianza di Sobolev-Poincarè, per la quale esiste una costante $C(E)$ tale che

$$\left(\int_E |T - \langle T \rangle|^p d\mathcal{L}^n \right)^{1/p} \leq C(E) \int_E |\nabla T| d\mathcal{L}^n,$$

dove $\langle T \rangle$ denota la media di T su E . Si noti che nel nostro caso corrisponde ad una traslazione dell'insieme E e che quindi si può supporre nulla. Tuttavia è necessario usare una costante non dipendente dall'insieme; nel nostro caso, in cui consideriamo insiemi convessi, è possibile mostrare (vedere per esempio [AD]) che la costante C dipende soltanto dal diametro dell'insieme nel caso in cui $p = 1$, in particolare vale

$$\int_E |T - \langle T \rangle| d\mathcal{L}^n \leq \frac{1}{2} \operatorname{diam}(E) \int_E |\nabla T| d\mathcal{L}^n.$$

Poiché $\operatorname{diam}(E) \leq \sigma_1(E)$ si può supporre che il diametro degli insiemi che consideriamo sia uniformemente limitato e quindi che la costante nella disuguaglianza di Sobolev non dipenda dall'insieme. Si ottiene quindi

$$\int_E |T - Id| d\mathcal{L}^n \leq C_2\delta(E)^{1/2}.$$

Come ultimo passaggio serve stimare direttamente l'asimmetria di E , ovvero stimare $A(E)$. L'idea è che essendo T vicino all'identità allora la misura dei punti

di E che vengono spostati più di ε è piccola. Poiché E e B hanno la stessa misura allora $|E \setminus B| = |B \setminus E|$. Fissato $\varepsilon > 0$ si ha

$$|E \setminus B| \leq |E \setminus (1 + \varepsilon)B| + |(1 + \varepsilon)B \setminus B|,$$

dove dato $\lambda > 0$ si pone $\lambda B = \{\lambda x \mid x \in B\}$. Allora un semplice conto mostra che $|(1 + \varepsilon)B \setminus B| = O(\varepsilon)$ per $\varepsilon \rightarrow 0$; per stimare l'altro termine abbiamo che se $x \in E \setminus (1 + \varepsilon)B$ allora $d(x, B) \geq \varepsilon$ e pertanto $|E \setminus (1 + \varepsilon)B| \leq 1/\varepsilon \int_{E \setminus (1 + \varepsilon)B} d(x, B)$. Poiché $T(x) \in B$ allora si ha $|T(x) - x| \geq d(x, B)$ e quindi si ottiene

$$|E \setminus B| \leq C_3 \left(\frac{1}{\varepsilon} \|T - Id\|_1 + \varepsilon \right) \leq C_4 \left(\frac{\delta(E)^{1/2}}{\varepsilon} + \varepsilon \right).$$

Infine ponendo $\varepsilon = \delta(E)^{1/4}$ si ottiene

$$A(E) \leq C\delta(E)^{1/4},$$

che è il risultato voluto.

Bibliografia

- [AD] G. ACOSTA AND R. G. DURÀN, An optimal Poincaré inequality in L^1 for convex domains, Proc. Amer Math. Soc., 132 (2003), pp. 195-202.
- [AFP] L. AMBROSIO, N. FUSCO, AND D. PALLARA, Functions of bounded variation and free discontinuity problems, Oxford Mathematical Monographs, Clarendon Press, Oxford, 2000.
- [B] H. BREZIS, Functional analysis, Sobolev spaces and partial differential equations, Universitext, Springer, New York, 2011.
- [G] E. DE GIORGI, Sulla proprietà isoperimetrica dell'ipersfera, nella classe degli insiemi aventi frontiera orientata di misura finita, Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Nat. Sez. I, (8) 5 (1958), 33-44.20-4792 (L. C. Young).
- [P] G. DE PHILIPPIS, Regularity of Optimal Transport Maps and Applications, Edizioni della Normale, 2013.
- [ET] I. EKELAND AND R. TÉMAM, Convex analysis and variational problems, vol. 28 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english ed., 1999. Translated from the French.
- [EG] L. C. EVANS AND R. GARIEPY, Measure Theory and Fine Properties of Functions, CRC Press - Studies in Advanced Mathematics (1992).
- [FMP] A. FIGALLI, F. MAGGI, AND A. PRATELLI, A mass transportation approach to quantitative isoperimetric inequalities, Invent. Math., 182, 2010, pp. 167-211.
- [F] G. B. FOLLAND, Real analysis, Wiley, 1984.
- [MS] V. D. MILMAN AND G. SCHECHTMAN, Asymptotic theory of finite-dimensional normed spaces. With an appendix by M. Gromov. Lecture Notes in Mathematics, 1200. Springer-Verlag, Berlin, 1986. viii+156 pp.

- [O] J. C. OXTOBY, Homeomorphic measures in metric spaces, Proc. Amer. Math. Soc., 24 (1970), pp. 419-423.
- [P] A. PRATELLI, On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation, Ann. Inst. H. Poincaré Probab. Statist., 43 (2007), pp. 1-13.
- [Ro] R. T. ROCKAFELLAR, Convex analysis, Princeton University Press, 1970.
- [Ru] W. RUDIN, Real and complex analysis, McGraw-Hill Book Co., New York, third ed., 1987.
- [T] T. TAO, An introduction to measure theory, American Mathematical Society, 2011.