Indice

Introduzione

Il problema del trasporto ottimale fu proposto per la prima volta nel 1781 da Monge [18], che si chiedeva quale fosse il modo di spostare un mucchio di terra da una posizione iniziale nota ad una finale stabilita facendo il minimo lavoro possibile: detti f^+ ed f^- i grafici delle posizioni iniziali e finali della terra e $\Omega \subseteq \mathbb{R}^n$ l'ambiente, il lavoro fatto con uno spostamento s risulta naturalmente

$$\int_\Omega |s(x) - x| f^+(x)\ dx, \tag{1}$$

e il fatto che s trasporti correttamente la terra equivale alla condizione

$$\int_B f^-(x)\ dx = \int_{s^{-1}(B)} f^+(x)\ dx \quad \forall B \in \mathfrak{B}(\Omega),$$

dove $\mathfrak{B}$ indica la classe dei boreliani. Un modo equivalente per esprimere questa condizione è dato dall'equazione

$$\int_\Omega h(s(x)) f^+(x)\ dx = \int_\Omega h(y) f^-(y)\ dy \quad \forall h \in C(\Omega). \tag{1.1”}$$

Per quanto la formalizzazione del problema, come si vede, sia piuttosto semplice, si vide subito che la sua risoluzione era ardua: la difficoltà è dovuta prevalentemente all'estrema nonlinearità di (1) e (1.1”). Per questo Monge non riuscì a risolvere la questione, ed ebbe soltanto la possibilità di capire alcune caratteristiche del problema: in particolare intuì il fatto fondamentale che, nel caso in cui vi sia un trasporto ottimale, esiste una funzione potenziale il cui gradiente determina le direzioni lungo le quali viene spostata la massa; una prima dimostrazione di questa proprietà, fatta utilizzando i moltiplicatori di Lagrange, venne proposta solo negli anni '20 da Appell [2], che considerava anche l'ipotesi che il trasporto ottimale fosse regolare, dopo più di un secolo trascorso senza sviluppi significativi del problema.

Dopo il contributo di Appell l'interesse riguardo alla questione si riaccese gradualmente finché negli anni '40 vi fu l'intervento decisivo di Kantorovich, che mutò completamente l'approccio al problema: la sua idea fu di considerare singolarmente ogni punto x in cui inizialmente vi sia massa e di ammettere la possibilità che questa massa non venga tutta trasportata in un unico punto $s(x)$, ma che possa essere distribuita in un modo qualsiasi nello spazio di arrivo; il modo naturale per formalizzare questa idea è considerare come "trasporto generalizzato" (mentre diremo "trasporto classico" quello pensato da Monge) una misura $\gamma \in \mathcal{M}^+(\Omega\times\Omega)$, in modo che la restrizione di γ a $\{x\}\times\Omega$ –o, più precisamente, la misura ottenuta tramite disintegrazione– rappresenti il modo in cui la massa contenuta in x viene distribuita lungo Ω. Con questo approccio il fatto che γ sia un trasporto ammissibile equivale alla condizione

$$\gamma(B \times \Omega) = f^+(B),\ \gamma(\Omega \times B) = f^-(B) \quad \forall B \in \mathfrak{B}(\Omega), \tag{2}$$

mentre il lavoro relativo al trasporto γ diventa

$$\int_{\Omega\times\Omega} |y-x|\ d\gamma(x,y).$$

Si può capire facilmente la portata del cambiamento indotto da questa nuova prospettiva: non si tratta solo del fatto che il problema risulta generalizzato, ma soprattutto dell'enorme facilitazione che ne consegue nella sua trattazione; il lavoro, infatti, a questo punto dipende in maniera lineare dal trasporto, e il vincolo di ammissibilità individua una regione limitata e convessa nello spazio delle misure. L'importanza dell'apporto dovuto a questo nuovo punto di vista è stata riconosciuta a tal punto che ormai ci si riferisce al problema di trasporto con il nome di Monge–Kantorovich.

Dall'intervento di Kantorovich in poi, l'impostazione del problema è stata generalizzata sempre di più; in particolare ci si è interessati alla situazione in cui f^+ ed f^- sono misure non negative, o in cui il "lavoro" ha la forma più generale

$$\int_{\Omega\times\Omega} c(x,y)\ d\gamma(x,y)$$

con un costo $c : \Omega \times \Omega \longrightarrow \mathbb{R}$ opportuno, in alcuni casi addirittura non necessariamente positivo; lo studio fatto in ambiente più generale ha poi permesso in alcuni casi di tornare indietro ottenendo risultati di esistenza per trasporti fatti in senso classico: il primo risultato di questo tipo è stato trovato nel '76 da Sudakov [27], come conseguenza dei suoi studi riguardanti la decomposizione di misure, mentre altri sono stati pubblicati più recentemente.

Negli ultimi anni l'importanza di questo problema è cresciuta a dismisura, e si sono trovati collegamenti con problemi fisici riguardanti i campi più disparati, dalla meteorologia all'astrofisica, fino alla chimica quantistica: un elenco impressionante, a questo riguardo, si può trovare nell'introduzione a [3], che tra l'altro si occupa soltanto del problema con il costo quadratico $c(x,y) = |x-y|^2$.

Uno studio molto esauriente dell'argomento è fatto nel recente libro [21], mentre in [14] viene studiata la situazione generale in cui vengano richieste le minime proprietà possibili sul costo c; estremamente interessante, infine, è l'introduzione a [12], che riassume la storia dei principali passi avanti compiuti nello studio del problema partendo da Monge fino ad arrivare a Sudakov.

L'intento di questo lavoro è quello di offrire inizialmente una panoramica generale del problema e delle sue caratteristiche più importanti, per poi considerarne in particolare tre formulazioni; dallo studio del rapporto tra queste formulazioni emergerà il ruolo di una misura, che risulta importante in contesti diversi; la parte originale di questa tesi, infine, consisterà nello studio delle proprietà di regolarità di questa misura.

Il primo capitolo è dedicato allo studio generale del problema di Monge–Kantorovich. Nella prima sezione viene formalizzata l'impostazione del problema con i trasporti classici come immaginati da Monge, ma in un contesto più generale possibile, in cui f^+ ed f^- sono misure non negative e il costo c è generico. Nella seconda sezione si introduce l'estensione proposta da Kantorovich con le sue conseguenze principali, prima fra tutte l'esistenza di un trasporto generalizzato ottimale (mentre per i trasporti classici non è nemmeno detto che ve ne siano di ammissibili). La terza sezione si occupa del concetto di monotonia ciclica; la definizione è molto semplice: un insieme $S \subseteq \Omega \times \Omega$ viene detto $c-$*ciclicamente monotono* se vale la

$$\sum_{i=1}^{k} c(x_i, y_i) \leq \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}) \tag{1.2}$$

per una qualsiasi scelta di finite coppie $(x_1, y_1), \ldots, (x_k, y_k) \in S$ e di una permutazione $\sigma \in \mathfrak{S}_k$; una prima proprietà che vedremo è che il supporto di una misura ottimale nel senso di Kantorovich è $c-$ciclicamente monotono. L'utilità della monotonia ciclica è dovuta principalmente al rapporto con la concavità; per questo nella quarta sezione estenderemo il concetto di concavità e di superdifferenziale in modo da tener conto del costo c e poi studieremo i collegamenti con la monotonia ciclica. La quinta sezione è dedicata allo studio di alcune importanti proprietà dei trasporti, fino alla dimostrazione dell'esistenza di "potenziali" le cui caratteristiche li rendono indispensabili in più contesti; ad esempio questi potenziali hanno la proprietà predetta da Monge di determinare le direzioni di spostamento nel caso di trasporti ottimali classici. La sesta sezione, infine, riporta gli enunciati di tre teoremi di esistenza di trasporti ottimali classici.

Il secondo capitolo è più tecnico, e si occupa di alcuni strumenti riguardanti le misure. Nella prima sezione si studiano due possibili idee per definire la dimensione di una misura, date da

$$\mathcal{H}\text{-}\dim(\mu) := \inf\left\{k : \exists B \subseteq \mathbb{R}^n, \mu(B) > 0,\ \mathcal{H}^k(B) < +\infty\right\} \tag{2.1}$$

$$\mathcal{W}\text{-}\dim(\mu) := (\mu - \text{ess}) \inf_{x\in\Omega} \Lambda_\mu(x), \tag{2.3}$$

indicando come di consueto con $\mathcal{H}^k$ la misura di Hausdorff $k-$dimensionale e definendo

$$\Lambda_\mu(x) := \sup\left\{k : \exists\ C \in \mathbb{R},\ \forall\rho > 0\ \mu(B(x,\rho)) \leq C\rho^k\right\}, \tag{2.2}$$

e si dimostra che coincidono; nel seguito risulterà comodo poter utilizzare alternativamente una definizione o l'altra sapendo che danno lo stesso risultato. Nella seconda sezione si introduce la generalizzazione del concetto di derivata debole proposta in [5]: data la misura $\mu \in \mathcal{M}^+(\Omega)$ si dicono "campi tangenti" alla misura gli elementi dell'insieme

$$X_\mu^p := \left\{\varphi \in L_\mu^p(\Omega, \mathbb{R}^n) : \operatorname{div}(\varphi\mu) \in L_\mu^p(\Omega)\right\} \tag{2.5}$$

e si definisce lo "spazio tangente" alla misura come

$$T_\mu^p := (\mu - \text{ess}) \bigcup \left\{\varphi(\cdot) : \varphi \in X_\mu^{p'}\right\}. \tag{2.6}$$

A questo punto si chiama "derivata tangenziale" di una funzione $v \in \mathcal{D}(\Omega)$ la

$$D_\mu v(x) := P_\mu(x)(Dv(x)), \tag{2.8}$$

dove $P_\mu(x)$ indica la proiezione sullo spazio tangente; con un opportuno completamento di $\mathcal{D}(\Omega)$, infine, è possibile definire $\mathrm{W}_\mu^{1,p}$, una generalizzazione degli spazi di Sobolev con interessanti proprietà. Nella terza sezione, poi, ci occuperemo della dimensione di questo spazio tangente –che in quasi ogni punto è effettivamente un sottospazio vettoriale– confrontandola con le due nozioni di dimensione introdotte nella prima sezione.

Nel terzo capitolo concentriamo la nostra attenzione su tre possibili formulazioni del problema di Monge–Kantorovich, aggiungendo a quella di Kantorovich le due proposte da Bouchitté, Buttazzo e Seppecher e da Brenier: ci troveremo in pratica a dover confrontare i tre estremali

$$\mathrm{L}_{mk} := \inf\left\{\int |x-y|\ d\gamma(x,y) :\ \gamma \in \Gamma(f^+, f^-)\right\} \tag{3.11}$$

$$\mathrm{L}_{bb} := \sup\left\{\int u\ d(f^+ - f^-) :\ u \in \mathrm{Lip}_1(\Omega)\right\} \tag{3.8}$$

$$\mathrm{L}_{br} := \inf\left\{\int_{t=0}^1 \|E_t\|\ dt :\ \{(\mu_t, E_t)\} \in \Pi(f^+, f^-)\right\}; \tag{3.10}$$

con $\Gamma(f^+, f^-)$ e $\Pi(f^+, f^-)$ indicheremo rispettivamente le misure ammissibili secondo Kantorovich –ovvero le misure di probabilità su $\Omega \times \Omega$ che verificano la (2)– e le configurazioni ammissibili secondo Brenier. Una configurazione secondo Brenier è una famiglia di misure scalari non negative μ_t e di misure vettoriali E_t con $t \in [0, 1]$, e viene detta ammissibile se $\mu_0 = f^+$ e $\mu_1 = f^-$, vale la relazione di compatibilità

$$\frac{d}{dt}\mu_t + \nabla \cdot E_t = 0 \tag{3.9}$$

in senso distribuzionale e $\bigcup_{t\in[0,1]} supp(\mu_t) \cup supp(E_t) =: \Omega_0 \subset\subset \Omega$. Le prime due sezioni saranno dedicate allo studio, rispettivamente, del secondo e del terzo problema, mentre il primo viene già ampiamente considerato nel primo capitolo; nella terza studieremo l'equivalenza tra i tre problemi: mostreremo l'uguaglianza tra i tre estremali, noteremo che ciascuno di loro ammette delle soluzioni e cercheremo di trovare collegamenti tra le soluzioni dei diversi problemi. Nell'ultima sezione si introdurrà la misura ottimale μ, facendo emergere il suo ruolo nell'ambito dei tre problemi considerati.

L'ultimo capitolo, che contiene i risultati originali di questa tesi, è dedicato allo studio della regolarità della misura ottimale μ in dipendenza da quella dei dati f^+ ed f^-. Una prima sezione introduttiva serve per dare un'idea di come si comporta la misura μ e per anticipare i risultati che verranno dimostrati nel seguito. Nella seconda, grazie ad una stima di tipo geometrico conseguente alla monotonia ciclica, si dimostra che la dimensione di μ è maggiore del massimo fra le dimensioni di f^+ ed f^-, dove la dimensione è intesa nel senso discusso nel secondo capitolo; in particolare, come sarà dimostrato successivamente, è sufficiente che una fra f^+ ed f^- sia assolutamente continua rispetto alla misura di Lebesgue per dedurre che anche μ lo è. La terza sezione contiene solamente un risultato tecnico, una stima geometrica accurata simile a quella usata nella seconda sezione. Nella quarta sezione, infine, si studia la sommabilità di μ: in particolare si dimostra che se f^+ ed f^- sono in L^p ed hanno i supporti separati allora anche $\mu \in L^p$ e che, se una tra f^+ ed f^- sta in L^p con $p > 1$, senza bisogno di ipotesi sull'altra si può dedurre che $\mu \in L^q$ con un opportuno $q > 1$. L'ipotesi di separazione dei supporti è forse solo apparentemente "tecnica" dal momento che, anche se il suo effettivo ruolo non è sempre ben chiaro, risulta molto comoda in varie situazioni diverse; una breve discussione a questo riguardo è fatta alla fine dell'ultima sezione del primo capitolo.

Capitolo 1

Il problema di Monge–Kantorovich

Questo capitolo intende offrire una panoramica generale del problema di trasporto di Monge–Kantorovich; dopo aver visto, nelle prime due sezioni, una formulazione generale della questione posta da Monge e il diverso approccio ideato da Kantorovich, studieremo alcune delle principali proprietà conosciute.

1.1 La formulazione classica di Monge

Come già discusso nell'introduzione, il problema del trasporto ottimale è stato proposto per la prima volta da Monge [18] nel 1781, in una forma molto intuitiva: avendo un cumulo di terra in una data posizione ed una buca di eguale volume da riempire esattamente con essa, si voleva trovare il modo di farlo con il minimo lavoro.

Formalizzare questo problema, in maniera anche più generale, è estremamente facile: si vuole semplicemente trovare

$$J(f^+, f^-) \ := \ \inf_{s \in \mathcal{A}} C(s)$$

dove f^+ ed f^- indicano la forma del cumulo di terra e della buca, s è un generico trasporto, $\mathcal{A}$ l'insieme di quelli ammissibili e $C(s)$ il costo totale relativo al trasporto s. Cerchiamo ora di esplicitare tutte queste condizioni; indicheremo d'ora in poi con Ω un sottoinsieme aperto, limitato e convesso di $\mathbb{R}^n$ e con R il suo diametro: si considerino quindi date, su Ω, f^+ ed f^-, che rappresentano l'altezza della terra e la profondità della buca; seguendo l'idea originale dovrebbero essere funzioni non negative, continue e limitate, ma risulta molto più interessante estendere il nostro campo di interesse in modo da contemplare una certa classe di misure: chiediamo solo che siano Borel regolari, a supporto compatto e non negative, visto che altrimenti si perde l'analogia con la questione originale. Per richiedere l'eguaglianza dei volumi tra il cumulo e la buca, sarà sufficiente supporre

$$\int_\Omega df^+ \;=\; \int_\Omega df^- \;=\; 1$$

considerando per convenzione unitario questo volume. Un *trasporto* della massa non sarà altro che una funzione boreliana $s : supp(f^+) \longrightarrow supp(f^-)$: la massa che si trovava in x verrà spostata in $s(x)$; visto che per ragioni di simmetria vogliamo che il problema sia identico scambiando il ruolo di f^+ ed f^-, chiediamo inoltre che s sia biunivoca, cosa ovviamente compatibile con il caso particolare offerto dal problema originario. Perché infine il trasporto sia ammissibile bisogna colmare esattamente la buca, il che vuol dire

$$f^-(B) = f^+(s^{-1}(B)) \quad \forall\, B \in \mathfrak{B}(\Omega), \tag{1.1}$$

indicando con $\mathfrak{B}$ la classe dei boreliani. Per poter esprimere comodamente tutto questo introduciamo alcune notazioni: con $\mathcal{M}(\Omega)$, $\mathcal{M}^+(\Omega)$ e $\mathcal{P}(\Omega)$ indichiamo rispettivamente lo spazio delle misure Borel regolari, il suo sottoinsieme convesso dato dalle misure non negative e l'ulteriore sottoinsieme convesso dato dalle misure di probabilità, ovvero quelle non negative di massa unitaria. Dati poi $X \subseteq \mathbb{R}^m$, $Y \subseteq \mathbb{R}^n$ e $s : X \longrightarrow Y$, si definisce il *push–forward* come $s_\# : \mathcal{M}(X) \longrightarrow \mathcal{M}(Y)$, essendo

$$s_\#\mu(B) := \mu(s^{-1}(B)) \quad \forall B \in \mathfrak{B}(Y);$$

la terminologia indica il fatto che $s_\#$ porta μ in $s_\#\mu$.

Vediamo subito un enunciato riguardante il push–forward, nell'ipotesi che la misura μ sia positiva:

Proposizione 1.1 *Per ogni $\varphi : Y \longrightarrow \mathbb{R}$ misurabile si ha che φ è $s_\#\mu-$integrabile se e solo se $\varphi \circ s$ è $\mu-$integrabile; in questo caso, inoltre, $\int_Y \varphi \; ds_\#\mu = \int_X \varphi \circ s \; d\mu$.*

Dim : Questa identità è immediata sulle funzioni caratteristiche, e quindi segue su tutte le funzioni integrabili per definizione di integrale.

■

La classe dei *trasporti ammissibili* sarà dunque data da

$$\mathcal{A} := \left\{ s : supp(f^+) \longrightarrow supp(f^-) \text{ borel. e biuniv.}, \; s_\# f^+ = f^- \right\}.$$

A questo punto si deve semplicemente formalizzare la richiesta di costo totale minimo: il modo più naturale di definire il *costo totale* C relativo ad un trasporto s è

$$C(s) := \int_\Omega c(x, s(x)) \; df^+(x)$$

essendo $c : supp(f^+) \times supp(f^-) \longrightarrow \mathbb{R}$ ciò che verrà chiamata semplicemente funzione *costo*, la cui interpretazione è evidente: $c(x,y)$ è il costo per il trasporto di una massa unitaria dal punto x al punto y; nello spirito di Monge, che si interessava al lavoro fatto, il costo sarebbe semplicemente quello lineare dato da $c(x,y) = |x-y|$. In effetti si vede che questo caso è particolarmente stimolante perché –a dispetto della sua visualizzazione così elementare– risulta estremamente complesso, ben più, ad esempio, del caso quadratico $c(x,y) = |x-y|^2$; tuttavia è interessante anche considerare altri costi, sotto opportune ipotesi: alcune tra le più frequenti e plausibili sono che $c(x,y)$ sia positiva, o continua, o simmetrica, o funzione solo di $|x-y|$, o ancora che c definisca una metrica.

Notiamo che, nella situazione particolare in cui $f^+, f^- \in L^1$ e $s \in C^1$, dalla 1.1 si ottiene

$$\int_{s^{-1}(B)} f^+(x)\ dx = \int_B f^-(y)\ dy = \int_{s^{-1}(B)} f^-(s(x))|\det\ Ds(x)|\ dx,$$

indicando per semplicità, pur con un abuso di notazione, con f^+ ed f^- anche le densità delle omonime misure. Ma allora, grazie alla genericità di B, si può concludere che con queste ipotesi aggiuntive la (1.1) diventa

$$f^+(x) = f^-(s(x))|\det\ Ds(x)|; \tag{1.1'}$$

la (1.1') è la condizione di ammissibilità usata da Appell in [2]; un'altra possibile condizione, citata in [12], è la

$$\int_\Omega h(s(x))f^+(x)\ dx = \int_\Omega h(y)f^-(y)\ dy \quad \forall h \in C(\Omega), \tag{1.1''}$$

che grazie alla proposizione 1.1 coincide con (1.1) nel caso in cui $f^+,\ f^- \in L^1$.

Risolta facilmente la questione di formalizzare il problema, non si riesce però a compiere molti passi avanti nella sua risoluzione, principalmente a causa della più totale mancanza di linearità, concetto che non ha nemmeno senso, dal momento che non vi è un modo ovvio di definire una somma su $\mathcal{A}$; il fatto poi che il costo totale sembri non dipendere affatto dalle derivate della funzione di trasporto toglie molte delle speranze di poter ragionare in termini di compattezza su qualche spazio di Sobolev, utilizzando i metodi più comuni del Calcolo delle Variazioni; oltretutto non è nemmeno sempre vero che la classe $\mathcal{A}$ sia non vuota: se ad esempio f^+ è una delta di Dirac centrata in un punto e f^- sono due delta centrate in punti diversi di peso $1/2$, risulta chiaro che nessun trasporto può essere fatto in maniera ammissibile. Per questi motivi il problema di Monge è sempre stato difficile da trattare, e prima di avere risultati abbastanza generali si è dovuto attendere fino alla prima metà di questo secolo.

1.2 Il contributo di Kantorovich

Negli anni '40 ci fu una svolta decisiva, quando Kantorovich propose in [15], [16] un nuovo punto di vista che mutava sensibilmente la situazione: innanzitutto chiamiamo

$$\Gamma(f^+, f^-) := \left\{\gamma \in \mathcal{P}(\Omega \times \Omega) \ : \ \gamma_1 = f^+, \ \gamma_2 = f^-\right\}$$

indicando con γ_1 e γ_2 i marginali di γ, ovvero $\pi_{1\#}\gamma$ e $\pi_{2\#}\gamma$. Si può a questo punto introdurre una nozione di "costo totale" anche per le misure in $\Gamma(f^+, f^-)$, che è semplicemente dato da

$$C(\gamma) := \int_{\Omega\times\Omega} c(x,y)\ d\gamma(x,y).$$

La proposta di Kantorovich è stata quella di cercare di minimizzare il costo totale nella classe di misure Γ, anziché in quella $\mathcal{A}$ dei trasporti ammissibili. La prima cosa da fare è notare che il secondo problema è una generalizzazione del primo, e questo discende dal seguente

Lemma 1.2 *Per ogni trasporto $s \in \mathcal{A}$ esiste una misura $\gamma \in \Gamma(f^+, f^-)$ tale che $C(\gamma) = C(s)$.*

Dim : Dato s, si definisca $\gamma = (1\times s)_\# f^+$, ricordando che $(1\times s)_\# : \mathcal{P}(\Omega) \longrightarrow \mathcal{P}(\Omega\times\Omega)$; a questo punto per avere la tesi basta verificare che γ sia una misura ammissibile e che ad essa sia associato lo stesso costo totale di s. Si ha, in effetti,

i) $\gamma_1 = f^+$.

$$\gamma_1(B) = \gamma(B\times\Omega) = (1\times s)_\# f^+(B\times\Omega) = f^+\left((1\times s)^{-1}(B\times\Omega)\right) = f^+(B)$$

ii) $\gamma_2 = f^-$.

$$\begin{aligned}\gamma_2(B) &= \gamma(\Omega\times B) = (1\times s)_\# f^+(\Omega\times B) = f^+\left((1\times s)^{-1}(\Omega\times B)\right) = f^+(s^{-1}(B)) \\ &= s_\# f^+(B) = f^-(B)\end{aligned}$$

iii) $C(\gamma) = C(s)$.

$$\begin{aligned}C(\gamma) &= \int_{\Omega\times\Omega} c(x,y)\ d\gamma(x,y) = \int_{\Omega\times\Omega} c(x,y)\ d(1\times s)_\# f^+(x,y) \\ &= \int_\Omega c\circ(1\times s)(x)\ df^+(x) = \int_\Omega c(x,s(x))\ df^+(x) = C(s)\end{aligned}$$

■

Una prima conseguenza di questo risultato è che $\inf_{\gamma\in\Gamma} C(\gamma) \leq \inf_{s\in\mathcal{A}} C(s)$, ma c'è in effetti molto di più: la dimostrazione assicura che le misure contenute in Γ tengono conto dei trasporti ammissibili. Inoltre danno una visione più completa del problema, e per rendersene pienamente conto è utile introdurre il concetto di *trasporto generalizzato*: chiameremo così le misure $\gamma \in \Gamma$, seguendo l'idea che la massa contenuta in un punto x non debba essere tutta spostata in un unico punto $s(x)$, ma possa essere distribuita su $supp(f^-)$ tramite la misura $\gamma(x, \cdot)$, o più precisamente tramite le misure ottenute da γ tramite disintegrazione; allora una parte della massa contenuta nel punto $x \in supp(f^+)$ verrà portata in ciascuno dei punti $y \in supp(f^-)$ tali che $(x, y) \in supp(\gamma)$. Risulterà molto utile visualizzare la situazione chiamando *raggi di trasporto*, o più rapidamente *raggi*, tutti i segmenti xy tali che $(x, y) \in supp(\gamma)$, che quindi collegano $supp(f^+)$ con $supp(f^-)$, e *insieme di trasporto* l'unione dei raggi: pensando che attraverso tali segmenti scorra la massa da $supp(f^+)$ a $supp(f^-)$ si avrà un'idea abbastanza chiara della situazione, tra l'altro molto vicina all'interpretazione originale di Monge, per la quale lungo tali segmenti veniva fisicamente trasportata la terra. Seguendo questo punto di vista si può pensare ad una inclusione $\mathcal{A} \subseteq \Gamma$, e del resto è facile notare che nella dimostrazione del lemma 1.2 la γ che si è costruita avente lo stesso costo totale di s era proprio l'immagine di s tramite questa inclusione.

Si può subito capire come il problema posto da Kantorovich sia estremamente più facile da affrontare, visto che il costo totale è lineare in γ e il vincolo di ammissibilità è convesso in $\mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$. Inoltre non esiste più il rischio che non vi siano trasporti ammissibili, perché in ogni caso c'è quello che distribuisce tutta la massa di f^+ in maniera omogenea secondo f^-, ovvero $\gamma = f^+ \otimes f^-$. Quello che non è chiaro è quale sia esattamente il collegamento tra la formulazione classica del problema data da Monge e l'estensione proposta da Kantorovich: in particolare sarebbe importante poter sapere in quali situazioni i minimi in un senso o nell'altro si corrispondono tramite l'inclusione appena discussa, o almeno quando si ha l'uguaglianza $\inf_{\gamma\in\Gamma} C(\gamma) = \inf_{s\in\mathcal{A}} C(s)$, che come vedremo nella sezione 1.5 assicura importanti proprietà agli eventuali trasporti ottimali.

Abbandoniamo per ora la discussione del problema secondo la formulazione di Monge per studiare le caratteristiche principali derivanti da quella di Kantorovich; grazie a quanto mostreremo in questo caso più generale, potremo poi risalire ad alcuni risultati riguardo ai trasporti secondo Monge nelle sezioni 1.5 e 1.6. Cominciamo per ora con il primo risultato fondamentale, mostrato ad esempio da Kellerer [17]:

Teorema 1.3 *Se $c \geq 0$ è s.c.i., esiste almeno un minimo di C in Γ.*

<u>Dim</u> : E' immediato notare che $\Gamma(f^+, f^-) \subseteq \mathcal{M}\big(supp(f^+) \times supp(f^-)\big) \subseteq \mathcal{M}(\Omega \times \Omega)$; essendo poi $supp(f^+) \times supp(f^-)$ chiuso, su $\mathcal{M}\big(supp(f^+) \times supp(f^-)\big)$ possiamo considerare la topologia debole$*$ indotta dalla dualità con $C\big(supp(f^+) \times supp(f^-)\big)$. Osserviamo ora che $\Gamma(f^+, f^-)$ è un insieme sequenzialmente chiuso: in effetti se $\gamma_k \in \Gamma(f^+, f^-)$

e $\gamma_k \xrightarrow{*} \gamma$, passando al limite si vede che γ è un funzionale lineare positivo su $C\big(supp(f^+) \times supp(f^-)\big)$, e quindi grazie al teorema di rappresentazione di Riesz è una misura positiva. Inoltre, visto che la funzione unitaria è continua, $\int d\gamma_k \longrightarrow \int d\gamma$, per cui $\int d\gamma = 1$ e allora $\gamma \in \mathcal{P}\big(supp(f^+) \times supp(f^-)\big) \subseteq \mathcal{P}(\Omega \times \Omega)$. Bisogna ora solo verificare che γ abbia i giusti marginali: intanto notiamo che se $\varphi(x)$ è una funzione continua, lo è anche $\tilde{\varphi}(x,y) = \varphi(x) = \varphi \circ \pi_1(x,y)$; allora

$$\int \varphi(x) d\pi_{1\#}\gamma_k = \int \varphi \circ \pi_1(x,y) d\gamma_k = \int \tilde{\varphi}(x,y) d\gamma_k \longrightarrow \int \tilde{\varphi}(x,y) d\gamma = \cdots = \int \varphi(x) d\pi_{1\#}\gamma$$

e quindi $\pi_{1\#}\gamma_k \xrightarrow{*} \pi_{1\#}\gamma$; ma $\pi_{1\#}\gamma_k = f^+ \ \forall \ k$, e una successione costante non può convergere debolmente che al suo valore, da cui segue $\pi_{1\#}\gamma = f^+$; chiaramente allo stesso modo si dimostra che $\pi_{2\#}\gamma = f^-$, per cui $\gamma \in \Gamma(f^+, f^-)$ e allora risulta provata la chiusura sequenziale di $\Gamma(f^+, f^-)$. In quanto contenuto in $\mathcal{P}\big(supp(f^+) \times supp(f^-)\big)$, poi, $\Gamma(f^+, f^-)$ è anche limitato e dunque relativamente sequenzialmente compatto. Indicato ora con I_f il funzionale sulle misure che manda μ in $\int f d\mu$, è sufficiente mostrare che I_c è sequenzialmente s.c.i. per avere la tesi, ragionando nel modo standard con la semicontinuità e la compattezza sequenziali. Per costruzione è chiaro che se f è continua lo è anche I_f; scrivendo dunque c come limite puntuale crescente di funzioni continue c_i si ottiene quanto voluto, perché risulta

$$\int c \ d\gamma = \int \lim_i \ c_i \ d\gamma = \lim_i \int c_i \ d\gamma = \lim_i \ \lim_k \int c_i \ d\gamma_k \leq \lim_i \ \underline{\lim}_k \int c \ d\gamma_k = \underline{\lim}_k \int c \ d\gamma_k.$$

Il fatto poi che una funzione s.c.i. positiva sia limite puntuale crescente di funzioni continue è molto semplice da dimostrare: qualunque sia c, basta definire

$$c_i(x) := \inf_y \ c(y) + i \ d(x,y)$$

per ottenere una funzione continua, lipschitziana con costante i, minore di c e crescente in i; è poi immediato verificare che, se $c \geq 0$ è s.c.i., le c_i convergono puntualmente a c.

■

Allora il problema di minimizzazione di Kantorovich ha soluzione; quello che non è chiaro è sotto quali ipotesi il minimo costo totale di questo problema si possa ottenere da un trasporto ammissibile, o almeno coincida con l'inf del costo totale fatto sui trasporti ammissibili, sempre che ve ne siano...

1.3 Monotonia ciclica per le misure γ ottimali

In questa sezione vedremo una proprietà estremamente importante delle misure γ ottimali, la cui esistenza è stata dimostrata con il teorema 1.3, la $c-$ciclica monotonia, che generalizza l'idea di monotonia ciclica data da Rockafellar [22]: una sua descrizione molto dettagliata si può trovare in un articolo di Gangbo e McCann [14]. Intanto vediamo la definizione: un insieme $S \subseteq X \times Y$ viene detto *c–ciclicamente monotono* se prese comunque finite coppie $(x_1, y_1), \dots, (x_k, y_k) \in S$ ed una permutazione $\sigma \in \mathfrak{S}_k$ si ha

$$\sum_{i=1}^{k} c(x_i, y_i) \leq \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}). \tag{1.2}$$

Mostriamo innanzitutto che i supporti delle misure ottimali secondo Kantorovich sono $c-$ciclicamente monotoni, e poi vedremo che da questo discenderanno importanti proprietà:

Teorema 1.4 *Se $c \geq 0$ è continua e il costo totale non è identicamente infinito, ogni γ ottimale nel senso di Kantorovich ha supporto $c-$ciclicamente monotono in $\Omega \times \Omega$.*

Dim : Se fosse falsa la tesi, esisterebbero finite coppie di punti $(x_1, y_1), \dots, (x_k, y_k)$ nel supporto di γ ed una permutazione $\sigma \in \mathfrak{S}_k$ tali che

$$\rho_\sigma(x_1, \dots, x_k, y_1, \dots, y_k) := \sum_{i=1}^{k} c(x_i, y_i) - c(x_i, y_{\sigma(i)}) > 0.$$

Ma $\rho_\sigma : \Omega \times \Omega \longrightarrow \mathbb{R}$ è una funzione continua, in quanto somma di finite funzioni continue; si possono allora trovare degli opportuni intorni U_i degli x_i e V_i degli y_i in modo che $\rho_{\sigma|U_1 \times U_2 \times \cdots \times V_k} > 0$. Per definizione, essendo $(x_i, y_i) \in supp(\gamma)$, deve essere $\lambda_i = \gamma(U_i \times V_i) > 0$; chiamate allora $\gamma_i = \gamma_{|U_i \times V_i}/\lambda_i$ e $\lambda = \min_{i=1,\dots,k} \lambda_i$, si può notare che $\gamma - \lambda\gamma_i$ è una misura positiva, e allo stesso modo è positiva anche $\gamma - \sum_{i=1}^{k} \lambda\gamma_i/k$, usando $\lambda\gamma_i/k$ al posto di $\lambda\gamma_i$ perché gli insiemi $U_i \times V_i$ potrebbero non essere disgiunti. Visto che le γ_i appartengono per costruzione a $\mathcal{P}(\Omega \times \Omega)$, si può trovare uno spazio probabilizzato (Δ, η) e delle funzioni $(u_i, v_i) : \Delta \longrightarrow \Omega \times \Omega$ tali che $\gamma_i = (u_i, v_i)_\# \eta$: questa è una nota proprietà delle misure di probabilità ed è anche molto semplice da mostrare: in questo caso è sufficiente prendere $\Delta = (\Omega \times \Omega)^k$, $\eta = \gamma_1 \otimes \cdots \otimes \gamma_k$ e come (u_i, v_i) le proiezioni. Dunque

$$\gamma - \frac{\lambda}{k} \left(\sum_{i=1}^{k} (u_i, v_i) \right)_\# \eta = \gamma - \frac{\lambda}{k} \sum_{i=1}^{k} \gamma_i$$

è una misura positiva, e quindi lo è a maggior ragione

$$\tilde{\gamma} = \gamma - \frac{\lambda}{k} \left(\sum_{i=1}^{k} (u_i, v_i) - (u_i, v_{\sigma(i)}) \right)_\# \eta.$$

Si vede però immediatamente che, grazie a come è stata costruita, $\tilde{\gamma}$ ha gli stessi marginali di γ, cioè f^+ ed f^-, visto che il resto si semplifica nella sommatoria, e quindi $\tilde{\gamma} \in \Gamma(f^+, f^-)$, usando il fatto che sappiamo già trattarsi di una misura positiva. Ora si può facilmente calcolare

$$\begin{aligned} C(\tilde{\gamma}) - C(\gamma) &= -\frac{\lambda}{k} \int_{\Omega\times\Omega} c(x,y) d \Big(\sum (u_i, v_i) - (u_i, v_{\sigma(i)}) \Big)_{\#} \eta \\ &= -\frac{\lambda}{k} \int_{\Delta} \sum c\big(u_i(x), v_i(x)\big) - c\big(u_i(x), v_{\sigma(i)}(x)\big) d\eta \\ &= -\frac{\lambda}{k} \int_{\Delta} \rho_\sigma(u_1(x), u_2(x), \dots, v_k(x)) d\eta. \end{aligned}$$

Ma per costruzione $(u_1(x), u_2(x), \dots, v_k(x)) \in U_1 \times U_2 \times \cdots \times V_k$ $\eta-$q.o.; visto che $\rho_{\sigma|U_1\times U_2\times\cdots\times V_k} > 0$ e η è una misura positiva, allora, si trova $C(\tilde{\gamma}) < C(\gamma)$, violando la minimalità di γ. L'assurdo mostra la $c-$ciclica monotonia e dunque la tesi.

■

Nel caso in cui non vi sia unicità del minimo, l'enunciato precedente sembrerebbe limitato dall'arbitrarietà nella scelta di γ; in effetti però si può arrivare ad un risultato più generale, il

Teorema 1.5 *Nelle stesse ipotesi del teorema 1.4, è $c-$ciclicamente monotona l'unione dei supporti delle misure ottimali.*

Dim : Visto che il costo è una funzione lineare e sono lineari i push–forward, è immediato notare che $\Gamma(f^+, f^-)$ è convesso ed è convesso il suo sottoinsieme Γ_0 dato dalle misure γ ottimali. Se allora sono dati $(x_1, y_1), \dots, (x_n, y_n) \in \bigcup_{\gamma\in\Gamma_0} supp(\gamma)$, consideriamo delle misure ottimali $\gamma_1, \dots, \gamma_k$ tali che $(x_i, y_i) \in supp(\gamma_i)$; sia ora $\bar{\gamma} = (\gamma_1 + \cdots + \gamma_k)/k$: è anch'essa una misura ottimale, per convessità, e nel suo supporto sono contenute tutte le coppie (x_i, y_i); scelta ora una permutazione $\sigma \in \mathfrak{S}_k$, per avere la disuguaglianza desiderata basta applicare il teorema 1.4 a $\bar{\gamma}$.

■

La disuguaglianza (1.2) può essere più espressiva –anche se evidentemente meno generale– nel caso $k = 2$, ovvero quando vi siano soltanto due coppie di punti: il suo significato intuitivo è che dati due punti x_1, x_2 nel supporto di f^+ e y_1, y_2 nel supporto di f^-, il modo più "conveniente" per collegarli con due raggi è quello che minimizza la somma dei costi. In particolare, se il costo è lineare, questo porta alla

Proposizione 1.6 *Se $c(x, y) = |x - y|$, se vi è un unico punto di intersezione fra due diversi raggi dev'essere un estremo di entrambi.*

Dim : L'equazione (1.2), nel caso del costo lineare e per $k = 2$, assicura che se (x_1, y_1)

e (x_2, y_2) sono nel supporto di γ, si ha

$$|x_1 - y_1| + |x_2 - y_2| \leq |x_1 - y_2| + |x_2 - y_1|. \tag{1.3}$$

Grazie alla disuguaglianza triangolare, da questo si ottiene immediatamente la tesi.

■

Si noti che nell'enunciato si è dovuto esplicitamente richiedere che vi fosse un unico punto di intersezione: se due raggi sono parzialmente sovrapposti, la (1.3) non estrae da questo alcuna informazione.

Nel caso di costo lineare la proposizione 1.6 è molto interessante, perché in pratica assicura che due raggi non si possono tagliare; tuttavia la sua utilità è ridotta quasi esclusivamente alla situazione bidimensionale, in cui questa è una condizione forte: se $n \geq 3$, infatti, non si tratta di una costrizione significativa; quello che invece ha grandi conseguenze è la disuguaglianza (1.3), che per questo sarà lo strumento fondamentale per il capitolo 4.

1.4 Concavità e superdifferenziale rispetto a c

In questa sezione interrompiamo la discussione riguardo al problema di Monge–Kantorovich per introdurre e studiare brevemente le nozioni di concavità e di superdifferenziale, in una forma abbastanza generale che permetta di tenere conto del costo c; nella prossima sezione, tornando al problema principale, vedremo come questi strumenti permettano di dedurre alcuni fatti interessanti. Le ipotesi di limitatezza e di convessità di Ω, fatte all'inizio di questo lavoro, non giocano alcun ruolo in questa sezione, i cui risultati sono validi qualunque sia l'aperto $\Omega \subseteq \mathbb{R}^n$, possibilmente $\mathbb{R}^n$ stesso.

Per cominciare, data $\psi : \Omega \longrightarrow \overline{\mathbb{R}}$ definiamo il *superdifferenziale* $\partial^{\cdot}\psi \in \Omega \times \mathbb{R}^n$ come l'insieme

$$\partial^{\cdot}\psi := \big\{(x,y) : \psi(x) \in \mathbb{R},\ \forall\, v \in S^{n-1}\ \psi(x+\epsilon v) \leq \psi(x) + \langle \epsilon v, y\rangle + o(\epsilon)\big\}\,, \qquad (1.4)$$

dove la richiesta $\psi(x) \in \mathbb{R}$ serve ad evitare i valori $\pm\infty$; diremo poi che y è un *supergradiente* di ψ in x se $(x,y) \in \partial^{\cdot}\psi$, e che ψ è *superdifferenziabile* in x se esiste almeno un supergradiente di ψ in x; gli analoghi termini *subdifferenziale*, *subgradiente* e *subdifferenziabile* si ottengono ponendo "$\geq$" al posto di "$\leq$"nella (1.4). Questo concetto estende in modo abbastanza intuitivo la nozione classica di derivata, come si nota grazie ad alcuni semplici esempi: se $\psi \in C^1$ si ha $\partial^{\cdot}\psi = \{(x, D\psi(x))\}$; se $\psi(x) = |x|$, su $\mathbb{R}$, i subgradienti in 0 sono tutti i numeri compresi tra -1 e 1, mentre non vi sono supergradienti; se $\psi = \delta_0$, qualunque vettore è un supergradiente in 0 mentre non vi sono subgradienti. Una prima immediata proprietà che mette in luce il collegamento tra il superdifferenziale e le funzioni concave è il

Lemma 1.7 *Se ψ è concava, $\psi(v) - \langle v, y\rangle \leq \psi(x) - \langle x, y\rangle$ per ogni $(x,y) \in \partial^{\cdot}\psi$ e $v \in \Omega$.*

<u>Dim</u> : Presa una coppia $(x,y) \in \partial^{\cdot}\psi$ e $v \in \Omega \setminus \{x\}$ (visto che per $v = x$ la tesi è vera a vuoto) si definisca $\alpha = \psi(v) - \psi(x) - \langle v - x, y\rangle$, in modo da dover mostrare $\alpha \leq 0$; per concavità si ha

$$\begin{aligned} \psi\big(x + \epsilon(v-x)\big) &= \psi\big((1-\epsilon)x + \epsilon v\big) \geq (1-\epsilon)\psi(x) + \epsilon\psi(v) \\ &\geq (1-\epsilon)\psi(x) + \epsilon\psi(x) + \epsilon\langle v-x, y\rangle + \epsilon\alpha \geq \psi(x) + \epsilon\langle v-x, y\rangle + \epsilon\alpha; \end{aligned}$$

questo, applicando la definizione (1.4) ponendo attenzione al fatto che in questo caso non è necessariamente $v \in S^{n-1}$, assicura che

$$o(\epsilon|v-x|) \geq \epsilon\alpha \Longrightarrow \varlimsup_{\epsilon\to 0} \frac{\epsilon\alpha}{\epsilon|v-x|} \leq 0 \Longrightarrow \frac{\alpha}{|v-x|} \leq 0 \Longrightarrow \alpha \leq 0$$

e quindi la tesi.

■

Data $\psi : \Omega \longrightarrow \overline{\mathbb{R}}$, si può parlare di iperpiani superiori a $\mathcal{G}r(\psi)$ nel punto $(x, \psi(x))$: si tratta degli iperpiani $n-$dimensionali in $\Omega \times \mathbb{R}$ parametrizzabili tramite le prime n coordinate –e quindi del tipo $\{(x, \langle v, x\rangle + c)\}$– che stanno sopra al grafico di ψ incontrandolo nel punto $(x, \psi(x))$; il lemma appena mostrato permette allora un'utile caratterizzazione del supergradiente per le funzioni concave:

Lemma 1.8 *Se ψ è concava, i supergradienti in x parametrizzano gli iperpiani superiori a $\mathcal{G}r(\psi)$ in $(x, \psi(x))$.*

Dim : Considerando per comodità $(x, \psi(x)) = (0, 0)$, è chiaro che gli iperpiani che incontrano l'origine sono tutti e soli quelli del tipo $\{(y, \langle v, y\rangle) : y \in \Omega\}$ con $v \in \Omega$, e si tratta di iperpiani superiori a $\mathcal{G}r(\psi)$ nell'origine se e solo se $\psi(y) \leq \langle v, y\rangle\ \forall y$; una parametrizzazione naturale, e quindi la tesi, seguirà allora dalla seguente doppia implicazione

$$(0, v) \in \partial^{\bullet}\psi \iff \psi(y) \leq \langle v, y\rangle\ \forall y,$$

che d'altra parte discende immediatamente dal lemma 1.7.

■

Fatte queste osservazioni, cerchiamo di estendere il concetto di concavità ad un generico costo: dal momento che una funzione concava è l'estremo inferiore di un opportuno insieme di funzioni affini, ad esempio quelle maggiori della funzione stessa, diremo che $\psi : \Omega \longrightarrow \mathbb{R} \cup \{-\infty\}$, diversa dalla costante $-\infty$, è *c–concava* se esiste un insieme $\mathfrak{F} \subseteq \Omega \times \mathbb{R}$ tale che

$$\psi(x) = \inf_{(y,\lambda)\in\mathfrak{F}} c(x, y) + \lambda;$$

nel caso in cui $c(x, y) = h(x - y)$ le funzioni $c-$concave sono dunque tutte quelle ottenibili tramite inviluppo inferiore e traslazioni a partire dalla sola h; ad esempio nel caso del costo quadratico $c(x, y) = |x - y|^2$ si ha la seguente caratterizzazione:

Osservazione 1.9 *Se $c(x, y) = |x - y|^2$, ψ è $c-$concava se e solo se $\psi(x) - |x|^2$ è concava.*

Dim : Dicendo $h(x) = |x|^2$, è chiaro che $h(x - y) = h(x) + h(y) - 2\langle x, y\rangle$; allora

$$\begin{aligned}
\psi \text{ è } c\text{−concava} &\iff \exists \mathfrak{F} : \psi(x) = \inf_{(y,\lambda)\in\mathfrak{F}} h(x - y) + \lambda \\
&\iff \exists \mathfrak{F} : \psi(x) = \inf_{(y,\lambda)\in\mathfrak{F}} |x|^2 + |y|^2 - 2\langle x, y\rangle + \lambda \\
&\iff \exists \widetilde{\mathfrak{F}} : \psi(x) - |x|^2 = \inf_{(y,\lambda)\in\widetilde{\mathfrak{F}}} \langle x, y\rangle + \lambda \iff \psi(x) - |x|^2 \text{ è concava.}
\end{aligned}$$

■

L'osservazione assicura che, con il costo quadratico, le funzioni $c-$concave sono tutte e sole le semiconcave con costante minore di 1, e quindi dalla teoria della semiconvessità seguono alcune buone proprietà.

Possiamo a questo punto estendere anche la nozione di superdifferenziale in modo da tener conto del costo c: data $\psi : \Omega \longrightarrow \mathbb{R} \cup \{-\infty\}$ $c-$concava, si dirà *c–superdifferenziale* l'insieme $\partial^c\psi \in \Omega \times \Omega$ definito da

$$\partial^c\psi := \{(x,y) : \psi(x) \in \mathbb{R},\ \forall\, v \in \Omega \ \ \psi(v) - c(v,y) \leq \psi(x) - c(x,y)\}\,;$$

si noti che le nozioni di $c-$concavità e di $c-$superdifferenziale si riconducono alla concavità e al superdifferenziale consueti nel caso del costo $c(x,y) = \langle x, y\rangle$, grazie al lemma 1.7; una prima proprietà del $c-$superdifferenziale che si può osservare è il

Lemma 1.10 *Se c è continuo e ψ è $c-$concava, $\partial^c\psi \subseteq \Omega \times \Omega$ è chiuso.*

Dim : Dal momento che c è continuo ψ, essendo $c-$concava, è l'estremo inferiore di funzioni continue e dunque è semicontinua superiormente; se allora $(\bar{x}, \bar{y}) \notin \partial^c\psi$, vuol dire che esiste un vettore v tale che $\psi(v) \geq \psi(\bar{x}) + c(v,\bar{y}) - c(\bar{x},\bar{y}) + \epsilon$; usando la continuità di c e la semicontinuità superiore di ψ, esiste un intorno di $(\bar{x}, \bar{y})$ in cui

$$c(v,\bar{y}) > (v,y) - \epsilon/3 \qquad -c(\bar{x},\bar{y}) > -c(x,y) - \epsilon/3 \qquad \psi(\bar{x}) > \psi(x) - \epsilon/3,$$

e quindi

$$\psi(v) > \psi(x) + c(v,y) - c(x,y);$$

ne segue che $(x,y) \notin \partial^c\psi$, per cui il complementare di $\partial^c\psi$ è aperto e dunque si ha la tesi.

■

Siamo a questo punto in grado di mostrare il risultato fondamentale di questa sezione, che mette in collegamento la $c-$ciclica monotonia con il $c-$superdifferenziale:

Teorema 1.11 *$S \subseteq \Omega \times \Omega$ è $c-$ciclicamente monotono se e solo se esiste una funzione $c-$concava $\psi : \Omega \longrightarrow \mathbb{R} \cup \{-\infty\}$ tale che $S \subseteq \partial^c\psi$.*

Dim : Un'implicazione è molto semplice: se ψ è $c-$concava, per definizione si ha che

$$\psi(v) \leq \psi(x) + c(v,y) - c(x,y) \tag{1.5}$$

per ogni $v \in \Omega$ e $(x,y) \in \partial^c\psi$ e che esiste un punto p per cui $\psi(p) > -\infty$; applicando dunque la (1.5) con $v = p$ si trova che per ogni coppia (x,y) nel $c-$superdifferenziale di ψ si ha $\psi(x) > -\infty$; si considerino allora finite coppie $(x_1,y_1), (x_2,y_2), \ldots, (x_k,y_k) \in \partial^c\psi$ ed una permutazione $\sigma \in \mathfrak{S}_k$; la (1.5) con $v = x_{\sigma(i)}$, $x = x_i$ e $y = y_i$ assicura che

$$\psi(x_{\sigma(i)}) \leq \psi(x_i) + c(x_{\sigma(i)}, y_i) - c(x_i, y_i);$$

sommando questa disequazione per $i = 1, 2, \ldots, k$ si ottiene

$$\sum_{i=1}^{k} c(x_i, y_i) \leq \sum_{i=1}^{k} c(x_{\sigma(i)}, y_i),$$

dal momento che i termini del tipo $\psi(x_i)$ si semplificano. Riassumendo, abbiamo verificato che se ψ è $c-$concava allora $\partial^c \psi$ è $c-$ciclicamente monotono; essendo questa una proprietà che passa ai sottoinsiemi, ne risulta immediatamente provata una delle implicazioni.

Per ottenere l'altra, dato un insieme S $c-$ciclicamente monotono si deve esibire un'opportuna funzione ψ; per farlo seguiamo la costruzione fatta da Rüschendorf in [23], mostrata anche in [21]: se $S = \varnothing$ si ha la tesi a vuoto, altrimenti si fissi $(x_0, y_0) \in S$ e si definisca $\Lambda = \big\{\{(x_1, y_1), \ldots, (x_k, y_k)\} \in S \times \cdots \times S,\ k \in \mathbb{N}\big\}$; sia quindi

$$\psi(x) := \inf_{\{(x_i, y_i)\} \in \Lambda} c(x, y_k) - c(x_k, y_k) + c(x_k, y_{k-1}) - \cdots + c(x_1, y_0) - c(x_0, y_0);$$

direttamente dalla definizione si ricava che ψ risulta $c-$concava purché non sia identicamente $-\infty$; ma grazie alla $c-$ciclica monotonia di S si ottiene subito che $\psi(x_0) = 0$, per cui ψ è in effetti $c-$concava. Presa allora $(\bar{x}, \bar{y}) \in S$, visto che $\psi(\bar{x}) < +\infty$ si possono trovare $(x_1, y_1), \ldots, (x_k, y_k) \in S$ tali che

$$\psi(\bar{x}) \geq c(\bar{x}, y_k) - c(x_k, y_k) + c(x_k, y_{k-1}) - \cdots - c(x_0, y_0) - \epsilon;$$

chiaramente $\{(x_1, y_1), \ldots, (x_k, y_k), (\bar{x}, \bar{y})\} \in \Lambda$, da cui segue che, scelto $x \in \Omega$,

$$\psi(x) \leq c(x, \bar{y}) - c(\bar{x}, \bar{y}) + c(\bar{x}, y_k) - \cdots - c(x_0, y_0) \leq c(x, \bar{y}) - c(\bar{x}, \bar{y}) + \psi(\bar{x}) + \epsilon,$$

che per la genericità di x ed $\epsilon > 0$ assicura

$$\psi(x) - c(x, \bar{y}) \leq \psi(\bar{x}) - c(\bar{x}, \bar{y}) \quad \forall x \in \Omega; \tag{1.6}$$

per dedurre da questo che $(\bar{x}, \bar{y}) \in \partial^c \psi$ si deve allora soltanto verificare $\psi(\bar{x}) > -\infty$; ma questo segue immediatamente dalla (1.6) con $x = x_0$, visto che $\psi(x_0) = 0$. La ψ costruita, dunque, è tale che $S \subseteq \partial^c \psi$, e quindi si è ottenuta la tesi.

■

Questo teorema è stato mostrato per la prima volta da Rockafellar in [22] nel caso degli insiemi ciclicamente monotoni e delle funzioni concave, mentre la generalizzazione al caso della $c-$ciclica monotonia e della $c-$concavità è stata notata da Smith e Knott [26].

Concludiamo questa sezione evidenziando alcune proprietà della diagonale $\Delta \subseteq \Omega \times \Omega$, che ci permetteranno tra l'altro di mostrare l'utile teorema 1.14; per fare questo saranno utili alcune notazioni: se μ e ν sono misure su Ω, definiremo

$$\mu \wedge \nu := \mu - (\mu - \nu)_+\,;$$

è immediato notare che si tratta della definizione formale del "minimo" tra le due misure; se invece γ è una misura su $\Omega \times \Omega$, con γ_d indicheremo la sua restrizione alla diagonale, che è quindi una misura su $\Omega \times \Omega$ avente supporto nella diagonale: è chiaro che nei casi più regolari $\gamma_d = 0$, ad esempio se γ è assolutamente continua. Diremo inoltre che una metrica d verifica la diseguaglianza triangolare stretta se $d(x,z) < d(x,y) + d(y,z)$ nel caso in cui $x \neq y$ e $y \neq z$. Possiamo a questo punto mostrare il

Lemma 1.12 *Se $\mu, \nu \in \mathcal{P}(\Omega)$ e $\gamma \in \Gamma(\mu,\nu)$, allora $\gamma_d \leq (1 \times 1)_\#(\mu \wedge \nu)$. Se poi il costo c è una metrica su Ω che verifica la disuguaglianza triangolare stretta e γ ha supporto $c-$ciclicamente monotono, la precedente disuguaglianza è in effetti un'uguaglianza.*

Dim : Per prima cosa notiamo che ha senso confrontare le misure γ_d e $(1 \times 1)_\#(\mu \wedge \nu)$: infatti 1×1 è una funzione da Ω a $\Omega \times \Omega$, per cui $(1 \times 1)_\#$ porta una misura su Ω –come $\mu \wedge \nu$– in una misura su $\Omega \times \Omega$; quelle che vogliamo confrontare sono allora due misure su $\Omega \times \Omega$, ed entrambe hanno supporto contenuto nella diagonale.

Indichiamo allora con γ_0 la restrizione di γ al complementare della diagonale, in modo da avere $\gamma = \gamma_0 + \gamma_d$; un semplice conto assicura

$$\pi_{1\#}\gamma_d(A) = \gamma_d(A \times \Omega) = \gamma_d(A \times A) = \gamma_d(\Omega \times A) = \pi_{2\#}\gamma_d(A);$$

definiamo allora $\beta = \pi_{1\#}\gamma_d = \pi_{2\#}\gamma_d$, e notiamo innanzitutto che

$$\gamma_d = (1 \times 1)_\# \beta, \tag{1.7}$$

visto che

$$\begin{aligned} \gamma_d(A \times B) &= \gamma_d(A \cap B \times A \cap B) = \gamma_d(A \cap B \times \Omega) = \beta(A \cap B) = \\ &= (1 \times 1)_\# \beta(A \cap B \times A \cap B) = (1 \times 1)_\# \beta(A \times B). \end{aligned}$$

Definendo allora $\mu_0 = \pi_{1\#}\gamma_0$ e $\nu_0 = \pi_{2\#}\gamma_0$, si ha che

$$\mu = \pi_{1\#}\gamma = \pi_{1\#}(\gamma_0 + \gamma_d) = \pi_{1\#}\gamma_0 + \pi_{1\#}\gamma_d = \mu_0 + \beta$$

e allo stesso modo $\nu = \nu_0 + \beta$; dal momento però che tutte le misure in gioco sono non negative, da questo segue che $\beta \leq \mu$ e $\beta \leq \nu$, e quindi per definizione $\beta \leq \mu \wedge \nu$; grazie a (1.7) si ottiene allora la prima parte della tesi.

Per la seconda, usando le stesse notazioni si deve mostrare che $\beta = \mu \wedge \nu$; ma avendo scritto $\mu = \mu_0 + \beta$ e $\nu = \nu_0 + \beta$, grazie alla definizione di $\mu \wedge \nu$ è chiaramente sufficiente trovare due insiemi disgiunti di misura piena rispettivamente secondo μ_0 e ν_0. Detto $S = supp(\gamma) \setminus \Delta$, i due insiemi saranno $X = \pi_1(S)$ e $Y = \pi_2(S)$: il fatto che siano di misura piena è immediato, per cui ci siamo ricondotti a mostrare $X \cap Y = \varnothing$. Se esistesse $z \in X \cap Y$, dovrebbero esistere $x, y \in \Omega$ tali che (x, z) e (z, y) sono in S, e questo assicurerebbe che $x \neq z$ e $z \neq y$ perché S non interseca la diagonale; essendo inoltre S $c-$ciclicamente monotono in quanto sottoinsieme di $supp(\gamma)$ che lo è, si avrebbe

$$c(x, z) + c(z, y) \leq c(x, y) + c(z, z); \tag{1.8}$$

essendo però c una metrica, $c(z, z) = 0$ e quindi la (1.8) unita all'osservazione che $x \neq z$ e $z \neq y$ mostra un assurdo in quanto c verifica la disuguaglianza triangolare stretta; si può allora concludere che $X \cap Y = \varnothing$ e quindi ottenere la tesi.

■

Un'altra proprietà utile da notare è il

Lemma 1.13 *La diagonale è contenuta nel $c-$superdifferenziale di ogni fuzione ψ che sia $c-$concava, se c è una metrica; per questo, se S è $c-$ciclicamente monotono, lo è anche $S \cup \Delta$.*

Dim : Notiamo subito che la seconda parte della tesi segue dalla prima: se S è $c-$ciclicamente monotono, grazie al teorema 1.11 esiste un funzione ψ $c-$concava tale che $S \subseteq \partial^c\psi$; grazie alla prima parte, però, $\Delta \subseteq \partial^c\psi$, per cui $S \cup \Delta \subseteq \partial^c\psi$ e quindi, usando ancora il teorema 1.11, $S \cup \Delta$ è $c-$ciclicamente monotono.

Per mostrare la prima parte, sia ψ una funzione $c-$concava; usando il fatto che c è una metrica e la definizione di $c-$concavità si ha che

$$\begin{aligned} \psi(z) - c(z, x) &= -c(z, x) + \inf_{(y,\lambda)\in\mathfrak{F}} c(z, y) + \lambda \leq -c(z, x) + \inf_{(y,\lambda)\in\mathfrak{F}} c(z, x) + c(x, y) + \lambda \\ &= \inf_{(y,\lambda)\in\mathfrak{F}} c(x, y) + \lambda = \psi(x) = \psi(x) - c(x, x), \end{aligned}$$

che grazie alla genericità di z assicura per definizione che $(x, x) \in \partial^c\psi$; la genericità di x, a questo punto, dà la tesi.

■

Una conseguenza molto importante di questo ultimo lemma è il

Teorema 1.14 *Se c è una metrica, ogni funzione $c-$concava è a valori reali ed è $1-$lipschitziana rispetto a c.*

Dim : All'inizio della dimostrazione del teorema 1.11 abbiamo notato che $\psi(x) > -\infty$ se ψ è $c-$superdifferenziabile in x; è chiaro, d'altronde, che grazie al lemma 1.13 se c è una metrica ogni funzione $c-$concava è $c-$superdifferenziabile in ogni punto, per cui è intanto mostrato che una funzione $c-$concava sia a valori reali. Dal conto fatto nella dimostrazione del lemma 1.13, inoltre, si può estrarre l'informazione che $\psi(z) - c(z, x) \leq \psi(x)$, e da questo segue che $\psi(z) - \psi(x) \leq c(z, x)$; dal momento che x e z sono generici, la stessa disuguaglianza vale anche scambiandoli fra di loro, e visto che per simmetria $c(z,x) = c(x,z)$ si deduce che

$$|\psi(x) - \psi(z)| \leq c(x,z),$$

trovando così la $1-$lipschitzianità cercata.

$\blacksquare$

1.5 Proprietà dei trasporti e potenziale ottimale

Riprendiamo ora lo studio del problema di Monge–Kantorovich, che avevamo interrotto nella scorsa sezione per introdurre alcune nozioni e studiarne le caratteristiche salienti; in questa sezione vedremo come da queste si possano dedurre diversi risultati interessanti riguardanti i trasporti ottimali di Monge o le misure ottimali di Kantorovich, e osserveremo l'esistenza di un potenziale dotato di proprietà particolari che lo rendono utile in vari contesti. Un primo fatto degno di nota è che, con opportune ipotesi sul costo c, vale il

Teorema 1.15 *Una misura $\gamma \in \Gamma(f^+, f^-)$ è ottimale per il problema di Kantorovich se e solo se il suo supporto è contenuto nel $c-$superdifferenziale di una funzione $c-$concava.*

Dim : Un'implicazione segue da quanto fatto finora: se γ è ottimale, grazie al teorema 1.4 il suo supporto è $c-$ciclicamente monotono e dunque, grazie al teorema 1.11, è contenuto nel $c-$superdifferenziale di un'opportuna funzione $c-$concava. L'altra implicazione è stata mostrata da Rüschendorf in [24], e si può trovare anche in [21].

■

A questo punto si può controllare cosa succede nel caso in cui le misure f^+ ed f^- non abbiano i supporti disgiunti: in particolare è spontaneo chiedersi se sia conveniente mantenere ferma la massa posta nella stessa posizione; è possibile farsi un'idea approssimativa di come debbano andare le cose considerando il semplice caso in cui $f^+ = \delta_A + \delta_B$ ed $f^- = \delta_B + \delta_C$ con A, B, C allineati: in questo caso è immediato notare che il trasporto ottimale (nel senso di Kantorovich) consiste nel mantenere ferma la massa posta in B e portare in C quella posta in A se il costo $c(x, y) = h(x - y)$ è strettamente concavo, ovvero se lo è h; è invece più convienente portare A in B e B in C se il costo è strettamente convesso, mentre i trasporti sono tutti equivalenti nel caso lineare. Per inciso, il caso concavo è quello di maggior interesse quando si affronti la questione dal punto di vista economico, visto che generalmente un tragitto di lunghezza doppia costa di più, ma comunque meno del doppio...

Si mostra allora il risultato seguente:

Teorema 1.16 *Nel caso in cui c sia una metrica che verifica la disuguaglianza triangolare stretta, se i supporti di f^+ ed f^- si intersecano allora una misura ottimale γ non muove la massa comune.*

Dim : Si noti innanzitutto che le ipotesi sono in accordo con l'idea intuitiva che deriva dall'esempio appena discusso: in effetti nelle situazioni più comuni $c(x, y) = h(|x-y|)$, e in questo caso il fatto che c sia una metrica che verifica la stretta disuguaglianza triangolare è equivalente alla stretta concavità di $h : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$.

Cerchiamo di formalizzare la condizione che la misura γ non muova la massa comune: chiaramente è come dire che se $A \subseteq supp(\mu) \cap supp(\nu)$ è tale che $\mu \llcorner A \leq \nu \llcorner A$, allora $\gamma(\{(x,x) : x \in A\}) = \mu(A)$, mentre se $A \subseteq supp(\mu) \cap supp(\nu)$ è tale che $\mu \llcorner A \geq \nu \llcorner A$, allora $\gamma(\{(x,x) : x \in A\}) = \nu(A)$. Più brevemente, allora, la condizione è

$$A \subseteq supp(\mu) \cap supp(\nu) \Longrightarrow \gamma(\{(x,x) : x \in A\}) = \mu \wedge \nu(A); \tag{1.9}$$

si deve allora solo mostrare che vale la (1.9). Notiamo allora che dal momento che γ è ottimale ha supporto $c-$ciclicamente monotono per il teorema 1.4, e quindi dal lemma 1.12 segue $\gamma_d = (1 \times 1)_\#(\mu \wedge \nu)$; se allora $A \subseteq supp(\mu) \cap supp(\nu)$, detto $\widetilde{A} = \{(x,x) : x \in A\}$ si ha

$$\gamma(\widetilde{A}) = \gamma_d(\widetilde{A}) = (1 \times 1)_\#(\mu \wedge \nu)(\widetilde{A}) = \mu \wedge \nu(A),$$

da cui la (1.9) e dunque la tesi.

■

In alcuni casi può essere utile una caratterizzazione che segue facilmente da questo teorema; per evidenziarla, dato il costo c ed una costante $\lambda > 0$ definiamo intanto

$$\tilde{c}(x,y) := \begin{cases} c(x,y) & \text{se } x \neq y \\ -\lambda & \text{se } x = y \end{cases};$$

si noti che $\tilde{c}$ non è continuo né positivo, per cui ad esso non si può applicare quasi nessuno dei teoremi fin qui dimostrati; tuttavia vale il

Corollario 1.17 *Se c è una metrica che verifica la disuguaglianza triangolare stretta e $\gamma \in \Gamma(f^+, f^-)$, γ è ottimale rispetto al costo c se e solo se è ottimale rispetto a $\tilde{c}$.*

Dim : E' una conseguenza immediata del teorema 1.16.

■

Notiamo adesso l'esistenza di un potenziale U che risulta molto comodo in varie situazioni (ad esempio lo useremo nel corso dei capitoli 2 e 3):

Teorema 1.18 *Se c è una metrica, data una misura ottimale γ esiste un potenziale $U : \Omega \longrightarrow \mathbb{R}$ tale che*

$$\begin{array}{ll} |U(x) - U(y)| \leq c(x,y) & \forall (x,y) \in \Omega \times \Omega \\ U(x) - U(y) = c(x,y) & \forall (x,y) \in supp(\gamma). \end{array} \tag{1.10}$$

Dim : Se γ è ottimale, il suo supporto è $c-$ciclicamente monotono per cui, grazie al teorema 1.11, esiste una funzione $c-$concava ψ tale che $supp(\gamma) \subseteq \partial^c\psi$; vogliamo mostrare la tesi con $U = \psi$. Il fatto che $|U(x)-U(y)| \leq c(x,y)$ $\forall (x,y) \in \Omega\times\Omega$ è la $1-$lipschitzianità rispetto a c, che discende dal teorema 1.14; data poi una coppia $(x,y) \in supp(\gamma)$, per ipotesi $(x,y) \in \partial^c U$, per cui dalla definizione del $c-$superdifferenziale e dalla $1-$lipschitzianità già notata si ha

$$
\begin{aligned}
U(v) - c(v,y) \leq U(x) - c(x,y)\ \forall v \quad &\Longrightarrow \quad U(y) \leq U(x) - c(x,y) \\
\Longrightarrow U(x) - U(y) \geq c(x,y) \quad &\Longrightarrow \quad U(x) - U(y) = c(x,y),
\end{aligned}
$$

da cui la tesi risulta provata.

■

In realtà si può mostrare anche un risultato leggermente più forte, il

Teorema 1.19 *Se c è una metrica, date f^+ ed f^- esiste un potenziale U che verifichi la (1.10) per ogni γ ottimale.*

<u>Dim</u> : In effetti nella dimostrazione del teorema 1.18 si è usato soltanto il fatto che il supporto di una misura ottimale γ è $c-$ciclicamente monotono; grazie al teorema 1.5, tuttavia, è $c-$ciclicamente monotona l'unione dei supporti della misure γ ottimali, per cui la stessa costruzione di 1.18 fatta a partire da questa unione conduce ad un potenziale U che soddisfa l'enunciato del teorema.

■

Torniamo adesso a considerare la formulazione originaria del problema, con i trasporti fatti nel senso classico: come abbiamo visto, non è detto in generale che tali trasporti esistano, né abbiamo ancora mostrato alcun teorema di esistenza di trasporti ottimali (la questione verrà considerata nella prossima sezione). Monge stesso, tuttavia, aveva intuito una proprietà importante degli eventuali trasporti ottimali, cioè il fatto che le direzioni di spostamento sono date dal gradiente di un'opportuno potenziale; esiste cioè una funzione U tale che, detto s il trasporto ottimale, per ogni $x \in supp(f^+)$ si ha

$$
s(x) - x = -DU(x)|s(x) - x|; \tag{1.11}
$$

una prima dimostrazione di questo fatto venne data negli anni '20 da Appell [2] aggiungendo l'ipotesi che il trasporto ottimale s fosse regolare, e solo grazie al contributo di Kantorovich divenne possibile avere una visione più completa della questione. Gli strumenti che abbiamo introdotto fin qui permettono di inquadrare abbastanza facilmente la situazione: se consideriamo il costo lineare e supponiamo che i minimi del problema di Monge e di quello di Kantorovich coincidano –mentre in generale sappiamo solo una disuguaglianza, si può mostrare l'enunciato di Monge, grazie al

Teorema 1.20 *Essendo $c(x,y) = |x-y|$ e supponendo $\inf_{\gamma\in\Gamma} C(\gamma) = \inf_{s\in\mathcal{A}} C(s)$, se s è un trasposto ottimale esiste un potenziale U tale che la (1.11) è verificata da tutti i punti tranne quelli contenuti nell'unione di due insiemi, uno trascurabile secondo Lebesgue e l'altro secondo f^+.*

Dim : Consideriamo un potenziale U che verifichi la (1.10) costruito grazie al teorema 1.18; essendo U lipschitziano, esiste il gradiente $DU(x)$ per tutti gli $x \in \Omega$ eccetto un insieme trascurabile secondo Lebesgue; essendo poi s un trasporto ottimale ed essendo $\inf_{\gamma\in\Gamma} C(\gamma) = \inf_{s\in\mathcal{A}} C(s)$ per ipotesi, ad s è associata la misura ottimale $\gamma = (1 \times s)_\# f^+$. Sia ora x un punto in cui sia definito il gradiente di U e tale che $(x, s(x)) \in supp(\gamma)$; supponiamo inoltre $x \neq s(x)$, visto che altrimenti la (1.11) si riduce all'uguaglianza $0 = 0$: allora grazie alla (1.10), ricordando che il costo è quello lineare, si ha che $U(x) - U(s(x)) = |x - s(x)|$; ma U è $1-$lipschitziana, per cui questa uguaglianza assicura che il gradiente di U in x ha modulo unitario ed ha la stessa direzione del vettore congiungente x ad $s(x)$ ma verso opposto; dunque, in accordo con la (1.11),

$$\frac{s(x) - x}{|s(x) - x|} = -DU(x),$$

Per avere la tesi, allora, è sufficiente mostrare che $(x, s(x))$ appartiene al supporto di γ per f^+-quasi ogni x, cosa abbastanza credibile dal momento che l'intuizione suggerirebbe $supp(\gamma) = \{(x, s(x))\}$; se in effetti definiamo

$$\Theta := \{(x, y) \in \Omega \times \Omega : y = s(x)\},$$

si ottiene subito che $supp(\gamma) \subseteq \Theta$, dal momento che, visto che $\gamma = (1 \times s)_\# f^+$,

$$\gamma(\Omega \times \Omega \setminus \Theta) = f^+\Big(\big\{x : (x, s(x)) \in \Omega \times \Omega \setminus \Theta\big\}\Big) = f^+(\varnothing) = 0.$$

Ma allora si ha quanto serviva per la tesi, dal momento che

$$\begin{aligned} 1 &= \gamma(\Theta) = \gamma\Big(\big\{(x, s(x))\big\}\Big) = \gamma\Big(\big\{(x, s(x)) : (x, s(x)) \in supp(\gamma)\big\}\Big) \\ &= f^+\Big(\big\{x : (x, s(x)) \in supp(\gamma)\big\}\Big). \end{aligned}$$

■

Facciamo ora alcune osservazioni riguardo a quanto mostrato fin qui: intanto è utile notare che, nel caso in cui $s(x) \neq x$, la (1.11) si può scrivere in maniera più espressiva come

$$\frac{s(x) - x}{|s(x) - x|} = -DU(x);$$

in effetti la possibilità che sia $s(x) = x$ è alquanto fastidiosa in molti casi: più in generale, un'ipotesi aggiuntiva molto comoda che spesso viene considerata è che i supporti di f^+ e di f^- siano disgiunti, come discuteremo meglio alla fine della prossima sezione.

Notiamo anche che l'enunciato del teorema appena mostrato è interessante prevalentemente nel caso in cui la misura f^+ sia assolutamente continua rispetto alla misura di

Lebesgue (cosa che accade ad esempio nell'interpretazione originale di Monge), visto che allora la tesi assicura che la (1.11) è verificata quasi ovunque secondo f^+; se invece il supporto di f^+ è trascurabile secondo Lebesgue, l'enunciato non dà alcuna informazione. Un caso in cui questo problema è evidente è ad esempio quello in cui f^+ ed f^- sono due delta di Dirac centrate in P e in Q, per cui il trasporto ottimale è anche l'unico possibile, cioè $s(P) = Q$; un potenziale che verifica (1.10) è $U(x) = |x - P|$, che ammette gradiente in ogni punto tranne P: in questo caso, quindi, la (1.11) non è mai verificata...

Bisogna anche osservare che per mostrare il teorema 1.20, in realtà, l'ipotesi che gli estremi inferiori del problema di Monge e di quello di Kantorovich coincidano non è strettamente necessaria: ad esempio la dimostrazione di Appell non ne fa evidentemente uso, essendo precedente alla formulazione di Kantorovich. L'utilità di questa ipotesi, tuttavia, non è solo quella di permettere di impostare un ragionamento semplice come quello con cui abbiamo dimostrato il teorema 1.20: è possibile ricavare un'altra caratteristica sorprendente riguardo ai trasporti ottimali; per vederlo, notiamo intanto che un potenziale per cui valga la (1.11) fornisce l'informazione relativa alle direzioni in cui si sposta la massa, ma non permette di conoscere il valore dello spostamento: in pratica, dunque, un simile potenziale fissa le direzioni di spostamento ma non è sufficiente a caratterizzare il trasporto. Se $\inf_{\gamma \in \Gamma} C(\gamma) = \inf_{s \in \mathcal{A}} C(s)$, allora, i trasporti ottimali sono un sottoinsieme delle misure ottimali; grazie al teorema 1.19, inoltre, sappiamo che esiste un potenziale che verifichi la (1.10) con ogni misura ottimale, per cui vi è un unico potenziale per cui valga la (1.11) con ogni trasporto ottimale. Ne segue che se una configurazione ammette più trasporti ottimali, le direzioni di spostamento sono comunque sempre le stesse; dato un punto x, in altre parole, l'insieme delle sue immagini $s(x)$ al variare dei trasporti ottimali s è contenuto in una retta passante per x: questo vale, naturalmente, per tutti i punti x tranne quelli contenuti nell'unione di due insiemi, uno trascurabile secondo Lebesgue e l'altro secondo f^+. Un esempio per illustrare la situazione può essere il seguente: siano f^+ ed f^- le misure di Hausdorff 2−dimensionali concentrate su due rettangoli posti nel piano xy, ottenuti l'uno dall'altro tramite una traslazione lungo l'asse x; è facile osservare che un trasporto ottimale è la traslazione che porta un rettangolo nell'altro, oppure la simmetria operata per l'asse di simmetria dei due rettangoli; è in realtà ottimale un trasporto che mandi in un modo qualsiasi ogni segmento parallelo all'asse x nel corrispondente segmento traslato: si vede quindi che vi sono trasporti ottimali molto diversi, ma ciascuno di loro sposta comunque i punti parallelamente all'asse x, e quindi le direzioni di spostamento sono in ogni caso le stesse.

1.6 Esistenza di trasporti ottimali in senso classico

In questa sezione enunceremo i più conosciuti risultati riguardanti l'esistenza di trasporti ottimali in senso classico, mostrati da Sudakov [27] nel '76, Gangbo–McCann [14] nel '96 ed Evans–Gangbo [12] nel '99. Il caso considerato da Sudakov e da Evans–Gangbo è quello in cui il costo è quello lineare e le misure f^+ ed f^- sono assolutamente continue; Gangbo–McCann, invece, si occupano di costi con opportune proprietà di convessità o di concavità, e si accontentano dell'assoluta continuità per una sola delle due misure. Il primo risultato ad essere mostrato è stato il

Teorema 1.21 (Sudakov) *Nel caso del costo lineare $c(x, y) = |x - y|$, se f^+ ed f^- sono assolutamente continue esiste un trasporto ottimale.*

La dimostrazione di questo teorema si trova in [27], e segue da risultati riguardanti la decomposizione di misure; in particolare Sudakov considera una misura ottimale nel senso di Kantorovich ed un potenziale U ottenuto come nel teorema 1.18, e si occupa dei segmenti massimali su cui U ha una crescita lineare con derivata unitaria. Dal momento che $U(x) - U(y) = |x - y| \quad \forall (x, y) \in supp(\gamma)$, tra questi segmenti vi sono tutti i raggi di trasporto a meno di insiemi trascurabili secondo Lebesgue –e dunque secondo f^+ ed f^-, che sono assolutamente continue. Fatte queste osservazioni, la costruzione di Sudakov consiste nel trovare un trasporto ottimale all'interno di ogni raggio di trasporto, risolvendo quindi un problema monodimensionale, per poi incollare le funzioni costruite, grazie al fatto che i raggi non si intersecano se non in insiemi trascurabili.

In [14] si può trovare una discussione molto esauriente che riguarda il caso di costo generico, e due enunciati di esistenza con opportune ipotesi: supponendo $c(x, y) = h(x-y)$ con h continua, si considerano le due seguenti possibili caratteristiche della situazione:

$$\begin{aligned} &-h : \mathbb{R}^n \longrightarrow \mathbb{R}^+ \text{ è strettamente convessa.} \\ &-h(x - y) = l(|x - y|) \text{ e } l : \mathbb{R}^+ \longrightarrow \mathbb{R}^+ \text{ è strettamente concava ed i} \\ &\quad \text{supporti di } f^+ \text{ e di } f^- \text{ sono separati.} \end{aligned} \tag{1.12}$$

Allora si può mostrare il

Teorema 1.22 (Gangbo–McCann) *Se il costo verifica una tra le due condizioni (1.12), se f^+ è assolutamente continua esiste una funzione $c-$concava $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}$ tale che $s : \Omega \longrightarrow \Omega$ definito da $s(x) = x - (Dh)^{-1} \circ D\psi(x)$ è un trasporto ottimale; inoltre $\gamma = (1 \times s)_\# f^+$ è l'unica misura ottimale nel senso di Kantorovich, a meno che il costo totale non sia identicamente infinito.*

Questo risultato segue delle proprietà riguardanti le misure ottimali e le funzioni $c-$concave studiate da Gangbo e McCann, di cui abbiamo in parte già discusso nelle sezioni precedenti; in base a queste proprietà viene anche definito opportunamente $(Dh)^{-1}$ nei punti dove non vi sia la differenziabilità classica; tra l'altro viene mostrato anche che, nelle ipotesi del teorema, la misura $\gamma = (1 \times s)_{\#} f^+$ ottimale non solo è l'unica misura ottimale, ma è anche l'unica in $\Gamma(f^+, f^-)$ avente il supporto $c-$ciclicamente monotono. L'enunciato più generale mostrato, in realtà, è più forte di quello che abbiamo appena citato: infatti l'ipotesi di separazione dei supporti può essere notevolmente indebolita, ed inoltre con opportune ipotesi aggiuntive sul comportamento asintotico della funzione costo si può estendere l'enunciato anche al caso in cui $\Omega = \mathbb{R}^n$, mentre noi consideriamo solo il caso in cui Ω sia limitato.

Con ipotesi simili a quelle di Sudakov vi è poi il risultato di Evans e Gangbo, che però hanno bisogno anche della seguente ipotesi aggiuntiva

$$\begin{cases} \text{i) I supporti di } f^+ \text{ e di } f^- \text{ sono disgiunti.} \\ \text{ii) I supporti di } f^+ \text{ e di } f^- \text{ sono regolari.} \\ \text{iii) } f^+ \text{ ed } f^- \text{ sono strettamente positive nell'interno dei rispettivi supporti.} \end{cases} \tag{1.13}$$

Si può quindi enunciare il

Teorema 1.23 (Evans–Gangbo) *Nel caso del costo lineare $c(x, y) = |x - y|$, se f^+ ed f^- sono assolutamente continue ed è verificata la (1.13), esiste un trasporto ottimale.*

La costruzione del trasporto ottimale in questo secondo caso, che si trova in [12], si basa essenzialmente su proprietà di soluzioni di equazioni differenziali, per cui è abbastanza diversa come metodo rispetto a quella operata da Sudakov; in primo luogo si definisce u_p come l'unica soluzione del $p-$laplaciano

$$\begin{cases} -\mathrm{div}\,(|Du_p|^{p-2} u_p) = f & \text{su } \Omega \\ u_p = 0 & \text{su } \partial\Omega, \end{cases}$$

dicendo $f = f^+ - f^-$. A questo punto, grazie ad opportune stime, si vede che è possibile estrarre una successione p_k in modo che si abbia

$$\begin{cases} u_{p_k} \longrightarrow u & \text{uniformemente} \\ Du_{p_k} \overset{*}{\rightharpoonup} Du & \text{in } L^\infty(\Omega, \mathbb{R}^n). \end{cases}$$

Si può mostrare che al di fuori dei supporti di f^+ e di f^- la u così costruita è una soluzione viscosa del problema

$$-u_i u_j u_{ij} = 0.$$

A questo punto l'idea è di fare in modo che questa u sia un potenziale nel senso della (1.11) per un trasporto ottimale s; come già notato nella scorsa sezione, conoscendo un tale potenziale sono fissate le direzioni di spostamento, per cui bisogna soltanto stabilire il modulo $|s(x) - x|$; una parte importante del lavoro di Evans–Gangbo, in effetti, consiste nello studiare la situazione in modo da poter concludere a questo punto la costruzione.

Vale la pena di discutere brevemente le ipotesi (1.13): mentre le parti ii) e iii) appaiono più che altro di natura tecnica, la i) è più particolare; in effetti in varie situazioni si trova che sarebbe molto comodo supporre che i supporti di f^+ ed f^- siano separati: ci è stato utile per la prima delle osservazioni fatte nella scorsa sezione al termine della dimostrazione del teorema 1.20, ne avremo bisogno nelle dimostrazioni di alcuni dei risultati che raggiungeremo nel capitolo 4 e viene usato da Gangbo–McCann e da Evans–Gangbo come abbiamo appena visto (anche se a Gangbo–McCann basta in realtà un'ipotesi più debole); quello che sembrerebbe plausibile, almeno a livello intuitivo, è che questa ipotesi non sia effettivamente necessaria, ma solo molto comoda per queste dimostrazioni; tuttavia non appare per ora evidente una strada né per farne a meno, né per trovare controesempi che ne assicurino la necessità. Per questo motivo risulta senz'altro preferibile evitare di fare ricorso a questa ipotesi, per quanto possibile, e nel capitolo 4 evidenzieremo quali siano i risultati che saremo in grado di mostrare senza usarla.

Capitolo 2

Dimensione, derivata tangenziale e spazio tangente per una misura

In questo capitolo ci occuperemo inizialmente di alcuni modi in cui si può definire la "dimensione" per una misura avente supporto in un aperto limitato Ω in $\mathbb{R}^n$; studieremo successivamente una generalizzazione della teoria delle derivate deboli, trovando una nozione di derivata rispetto ad una misura e di spazio tangente ad una misura in un punto; confronteremo infine la dimensione di questi spazi tangenti alla misura con la dimensione della misura stessa. Quelle con cui lavoreremo in questo capitolo saranno sempre misure non negative, finite e Borel regolari.

2.1 Dimensione di Hausdorff e dimensione minimale

Un primo modo di definire la dimensione di una misura μ è appoggiarsi alla nozione di dimensione di Hausdorff dei sottoinsiemi di $\mathbb{R}^n$ per controllare quale sia il più piccolo insieme visto dalla misura stessa, ponendo quindi

$$\mathcal{H}-\text{dim}\ (\mu) \ := \ \inf\left\{k : \exists B \subseteq \mathbb{R}^n, \mu(B) > 0,\ \mathcal{H}^k(B) < +\infty\right\}. \tag{2.1}$$

Chiameremo *dimensione di Hausdorff* questa $\mathcal{H}-$dim. A seconda delle situazioni e dei punti di vista, può essere comodo utilizzare un'altra caratterizzazione della $\mathcal{H}-$dim, data dalla seguente

Proposizione 2.1 $\mathcal{H}-\text{dim}\ (\mu) = \sup\ \left\{k : \mu \ll \mathcal{H}^k\right\}$

Dim : Definiti $X = \left\{k : \exists B \subseteq \mathbb{R}^n, \mu(B) > 0,\ \mathcal{H}^k(B) < +\infty\right\}$ e $Y = \left\{k : \mu \ll \mathcal{H}^k\right\}$, dobbiamo dimostrare che $\inf X = \sup Y$. Se $k_1 \in X$, sia B tale che $\mu(B) > 0$ e $\mathcal{H}^{k_1}(B) < +\infty$; per ogni $k_2 > k_1$, allora, si ha $\mathcal{H}^{k_2}(B) = 0$ che, dal momento che $\mu(B) > 0$, assicura $\mu \not\ll \mathcal{H}^{k_2}$ e dunque $k_2 \notin Y$; da questo segue $\inf X \geq \sup Y$. Sia d'altra parte $k \notin X$: allora $\nexists B: \ \mu(B) > 0, \ \mathcal{H}^k(B) < +\infty$, per cui $\mathcal{H}^k(B) < +\infty \Rightarrow \mu(B) = 0$, e allora a maggior ragione $\mathcal{H}^k(B) = 0 \Rightarrow \mu(B) = 0$: per definizione ne segue $k \in Y$, e quindi si ottiene anche $\inf X \leq \sup Y$, da cui la tesi.

■

Può essere utile sapere se, nella definizione di $\mathcal{H}-$dim, vengono raggiunti l'inf o il sup; da quanto visto finora si può intanto dedurre che almeno uno dei due è raggiunto: nella dimostrazione della proposizione 2.1, infatti, si è visto che $k \notin X \Rightarrow k \in Y$.

Veniamo ora alla seconda nozione di dimensione: data la misura μ, per ogni $x \in \Omega$ si può definire la dimensione puntuale come

$$\Lambda_\mu(x) := \sup\left\{k : \exists\ C \in \mathbb{R},\ \forall \rho > 0\ \mu(B(x,\rho)) \leq C\rho^k\right\} \tag{2.2}$$

indicando con $B(x,\rho)$ la palla di centro x e raggio ρ; a questo punto la *dimensione minimale* sarà data da

$$\mathcal{W}-\text{dim}\ (\mu)\ :=\ (\mu - \text{ess}) \inf_{x\in\Omega} \Lambda_\mu(x), \tag{2.3}$$

dove con $(\mu - \text{ess}) \inf f(x)$ si indica $\inf\{r : \mu(\{x : f(x) < r\}) > 0\}$.

Per motivi diversi, sia la $\mathcal{H}-$dim che la $\mathcal{W}-$dim risultano molto utili: la prima per il suo collegamento con la teoria della dimensione di Hausdorff, della quale si possono utilizzare tutte le proprietà note, la seconda perché adatta ad un ragionamento "puntuale" come quello che avremo nel capitolo 4. Guardando le definizioni si nota che entrambe si rifanno alla massa delle palle di raggio ρ confrontata con ρ^k, per cui si può intuire un certo rapporto fra le due "dimensioni"; quello che mostreremo è che in effetti coincidono, per cui potremo usare indifferentemente l'una o l'altra, cosa che ci sarà molto utile nel seguito.

Indicata come di consueto ω_d la costante utile per definire la misura di Hausdorff $\mathcal{H}^d$, definiamo

$$\theta^*_\alpha(\mu, x) := \varlimsup_{\rho\to 0} \frac{\mu(B(x,\rho))}{\omega_\alpha \rho^\alpha}. \tag{2.4}$$

In pratica $\theta^*_\alpha(\mu, x)$ è la densità $\mathcal{H}^\alpha$ dimensionale media intorno ad x: se ad esempio $\mu = \mathcal{H}^\beta \llcorner P$, essendo P un piano $\beta-$dimensionale che contiene x, sarà

$$\theta^*_\alpha(\mu, x) = \begin{cases} +\infty & \text{se } \beta < \alpha \\ 1 & \text{se } \beta = \alpha \\ 0 & \text{se } \beta > \alpha \end{cases}.$$

Si può allora enunciare il seguente

Teorema 2.2 *a) Se* $\theta^*_\alpha(\mu,x) \le t$ su B, allora $\mu \llcorner B \le 2^\alpha\ t\ \mathcal{H}^\alpha \llcorner B$*;*
b) se $\theta^*_\alpha(\mu,x) \ge t$ su B, allora $\mu(\Omega) \ge\ t\ \mathcal{H}^\alpha(B)$.

Si tratta di due casi particolari che si possono ottenere dal Teorema 3.2 in [25].

Da questo discende il seguente

Lemma 2.3 $\theta^*_\alpha(\mu,x) < +\infty\ \ \mathcal{H}^\alpha-$q.o.

Dim : Sia $B = \{x \in \Omega\ :\ \theta^*_\alpha(\mu,x) = +\infty\}$; allora in particolare per ogni $t \in \mathbb{R}$ si ha $\theta^*_\alpha(\mu,x) \ge t$ su B, per cui dal teorema 2.2 si ottiene $\mathcal{H}^\alpha(B) \le \mu(\Omega)/t$; visto che μ è una misura finita e che questo vale per ogni t, ne segue la tesi.

■

Si noti che questo lemma assicura la finitezza di $\theta^*_\alpha(\mu,x)$ quasi ovunque secondo $\mathcal{H}^\alpha$, non certo secondo μ: se ad esempio $\mu = \mathcal{H}^\beta \llcorner P$, essendo P un piano $\beta-$dimensionale (con $\beta < n$) e $\alpha = n$, e quindi $\mathcal{H}^\alpha$ è la misura di Lebesgue,

$$\theta^*_\alpha(\mu,x) = \begin{cases} +\infty & \text{se } x \in P \\ 0 & \text{se } x \notin P \end{cases} ;$$

perciò, visto che la misura di Lebegue non vede i piani $\beta-$dimensionali, θ^*_α è nullo quasi ovunque secondo $\mathcal{H}^\alpha$, in accordo con il lemma; tuttavia è infinito quasi ovunque secondo μ! Finalmente possiamo mostrare il

Teorema 2.4 $\mathcal{H}-\dim = \mathcal{W}-\dim$*, ovvero coincidono la dimensione di Hausdorff e quella minimale.*

Dim : **i)** $\mathcal{W}-\dim \le \mathcal{H}-\dim$.

Sia $d < \mathcal{W}-\dim$: se mostriamo $\mu \ll \mathcal{H}^d$ chiaramente otteniamo anche la disuguaglianza voluta, grazie alla proposizione 2.1; dividiamo allora Ω nell'unione disgiunta degli Ω_i con $i \in \mathbb{N} \cup \infty$, definendo $\Omega_i = \{x : \theta^*_d(\mu,x) \in [i, i+1)\}$ e $\Omega_\infty = \{x : \theta^*_d(\mu,x) = +\infty\}$. Ora dalla definizione di $\mathcal{W}-\dim$ segue che per $\mu-$q.o. x si ha $\Lambda_\mu(x) > d$, e ciascuno di questi x appartiene evidentemente ad un Ω_i, per un certo $i \in \mathbb{N}$, per cui $\mu(\Omega_\infty) = 0$. A questo punto sia B tale che $\mathcal{H}^d(B) = 0$: si deve mostrare $\mu(B) = 0$; chiamato $B_i = \Omega_i \cap B$, si ha che $\mu(B_\infty) \le \mu(\Omega_\infty) = 0$. Se ad Ω_i, poi, si applica il teorema 2.2, risulta $\mu \llcorner \Omega_i \le 2^d\ (i+1)\ \mathcal{H}^d \llcorner \Omega_i$, che assicura $\mu(B_i) \le 2^d\ (i+1)\ \mathcal{H}^d(B_i) = 0$, dal momento che $B_i \subseteq B$ e $\mathcal{H}^d(B) = 0$; ma allora, visto che $\mu(B_i) = 0\ \forall\ i \in \mathbb{N} \cup \infty$, in effetti B è $\mu-$trascurabile.

ii) $\mathcal{W}-\dim \ge \mathcal{H}-\dim$.

Sia $d < \mathcal{H}-\dim$; grazie al lemma 2.3, $\theta^*_d(\mu,x) < +\infty\ \ \mathcal{H}^d-$q.o.; ma allora è anche $\theta^*_d(\mu,x) < +\infty\ \mu-$q.o., perché $\mu \ll \mathcal{H}^d$. Questo assicura che per $\mu-$q.o. x si ha $\Lambda_\mu(x) \ge d$, e quindi che $\mathcal{W}-\dim(\mu) \ge d$, da cui anche questa disuguaglianza è provata.

■

2.2 Derivata tangenziale e spazio tangente ad una misura

In questa sezione vedremo il modo di definire la derivata rispetto ad una misura e lo spazio ad essa tangente introdotto da Bouchitté, Buttazzo e Seppecher in [5]. Data una misura finita $\mu \in \mathcal{M}^+(\Omega)$, poniamo intanto

$$L^p_\mu(\Omega) := \left\{ f : \int_\Omega |f|^p d\mu < +\infty \right\}.$$

A questo punto definiamo l'insieme dei *campi tangenti* a μ come

$$X^p_\mu := \left\{ \varphi \in L^p_\mu(\Omega, \mathbb{R}^n) : \operatorname{div}(\varphi\mu) \in L^p_\mu(\Omega) \right\}, \tag{2.5}$$

intendendo la divergenza nel senso distribuzionale; equivalentemente, si può dire che $\varphi \in L^p_\mu$ appartiene a X^p_μ se e solo se esiste una costante K tale che

$$\int \varphi \cdot Dv \; d\mu \leq K \|v\|_{L^{p'}_\mu} \qquad \forall v \in \mathcal{D}(\Omega).$$

E' opportuno notare che questa definizione è in accordo con l'intuizione nei casi più semplici: se ad esempio μ è la misura di Hausdorff k dimensionale su un $k-$piano, l'insieme X^p_μ è formato soltanto da campi tangenti al piano sufficientemente regolari; questo può essere mostrato facilmente osservando l'esistenza di funzioni regolari di norma $L^{p'}_\mu$ arbitrariamente piccola con derivate normali ad un piano arbitrariamente grandi sul piano stesso, purché rapidamente decrescenti fuori. Un altro esempio confortante è il caso in cui μ sia la misura $\mathcal{H}^1$ su una curva regolare: si mostra in modo simile che X^p_μ è formato solo da campi tangenti puntualmente alla curva.

Sfruttando la definizione di unione essenziale discussa nell'appendice B, allora, si può definire lo *spazio tangente* a μ come

$$T^p_\mu := (\mu - \text{ess}) \bigcup \left\{ \varphi(\cdot) : \varphi \in X^{p'}_\mu \right\}; \tag{2.6}$$

T^p_μ è una multifunzione, per cui per ogni x appartenente al supporto di μ si ha lo spazio tangente in x a μ, $T^p_\mu(x)$; i due esempi considerati sopra sono ancora utili per farsi un'idea del significato di questa definizione: se μ è la misura di Hausdorff su un piano, da quanto osservato sopra segue che $T^p_\mu(x)$ è il piano stesso per ogni x, e allo stesso modo se μ è la misura $\mathcal{H}^1$ su una curva regolare $T^p_\mu(x)$ è la retta tangente in x alla curva. Una conseguenza delle proprietà generali dell'unione $\mu-$essenziale è che

$$\begin{aligned} &a)\ \forall \varphi \in X^{p'}_\mu, \varphi(x) \in T^p_\mu(x)\ \widetilde{\forall} x; \\ &b)\ \text{se per una multifunzione } \Sigma \text{ vale la proprietà a), allora } T^p_\mu(x) \subseteq \Sigma(x)\ \widetilde{\forall} x. \end{aligned} \tag{2.7}$$

Grazie alla definizione, è chiaro che X^p_μ è un sottospazio vettoriale di L^p_μ e da questo si può dedurre che, per $\mu-$q.o. x, $T^p_\mu(x)$ è un sottospazio vettoriale di $\mathbb{R}^n$, che quindi ha senso indicare come spazio tangente a μ in x. Chiameremo allora $P_\mu(x) : \mathbb{R}^n \longrightarrow T^p_\mu(x)$ la proiezione sullo spazio tangente, definita per $\mu-$quasi ogni punto.

Ci proponiamo a questo punto di dare una nuova definizione di derivata rispetto alla misura μ, che chiameremo *derivata tangenziale*; cominciamo a farlo per funzioni regolari: se $v \in \mathcal{D}(\Omega)$, per $\mu-$quasi ogni punto x definiamo

$$D_\mu v(x) := P_\mu(x)(Dv(x)). \tag{2.8}$$

Per poter estendere la definizione, risulta molto comodo il

Lemma 2.5 *Per ogni $v \in \mathcal{D}(\mathbb{R}^n)$ e $\varphi \in X^{p'}_\mu$ si ha*

$$\int D_\mu v \cdot \varphi \; d\mu = -\langle \operatorname{div}(\varphi\mu), v\rangle \tag{2.9}$$

Dim : Per funzioni in $\mathcal{D}(\Omega)$ la derivata tangenziale è la proiezione della derivata classica sullo spazio tangente; tuttavia, per definizione di $X^{p'}_\mu$, $\varphi(x)$ appartiene allo spazio tangente per $\mu-$q.o. x, e quindi il suo prodotto scalare con $Dv(x)$ o con $D_\mu v(x)$ è lo stesso; essendo la divergenza intesa nel senso delle distribuzioni, ne segue

$$\int D_\mu v \cdot \varphi \; d\mu = \int Dv \cdot \varphi \; d\mu = -\langle \operatorname{div}(\varphi\mu), v\rangle$$

e quindi la tesi.

■

Una prima conseguenza della formula (2.9) è il seguente

Lemma 2.6 *Se $\{v_h\} \subseteq \mathcal{D}(\Omega)$ e $w \in L^p_\mu(\Omega, \mathbb{R}^n)$ sono tali che $v_h \xrightarrow[L^p_\mu(\Omega)]{} 0$ e $D_\mu v_h \xrightarrow[L^p_\mu(\Omega,\mathbb{R}^n)]{} w$, allora $w = 0$.*

Dim : Per prima cosa si prenda una generica $\varphi \in X^{p'}_\mu$; ricordando che per definizione φ e $\operatorname{div}(\varphi\mu)$ appartengono rispettivamente a $L^{p'}_\mu(\Omega, \mathbb{R}^n)$ e $L^{p'}_\mu(\Omega)$ e usando la (2.9) si ha

$$\int w \cdot \varphi \; d\mu = \lim \int D_\mu v_h \cdot \varphi \; d\mu = \lim \; -\langle \operatorname{div}(\varphi\mu), v_h\rangle = 0.$$

Da questo, per la genericità di φ, si può dedurre che

$$0 = \sup_{\varphi \in X^{p'}_\mu} \int w \cdot \varphi \; d\mu = \int \sup_{z \in T^p_\mu(x)} w(x) \cdot z \; d\mu; \tag{2.10}$$

per mostrare l'ultimo passaggio si può applicare il corollario A.7 con $j(x,z) = w(x) \cdot z$ e $K = X^{p'}_\mu$: è chiaro che per ogni x $j(x,\cdot)$ sia s.c.s. e concavo, visto che in effetti è lineare e continuo, ed è anche evidente che se $\varphi \in X_\mu^{p'}$ e $\psi \in C^\infty(\overline{\Omega}) \cap L^\infty$ allora $\varphi\psi \in X_\mu^{p'}$. Si deve infine controllare che, per ogni $\varphi \in X_\mu^{p'}$, $\int w(x) \cdot \varphi(x)\, d\mu$ sia finito, cosa ovvia dal momento che $\varphi \in L_\mu^{p'}$ e $w \in L_\mu^p$: allora le ipotesi del teorema sono verificate e l'ultimo passaggio in (2.10) è provato; da questo si ottiene subito che $w(x) \in \left(T_\mu^p(x)\right)^\perp$ per $\mu-$q.o. x. Consideriamo allora lo spazio $X := \left\{ f \in L_\mu^p(\Omega, \mathbb{R}^n) : f(x) \in T_\mu^p(x)\ \widetilde{\forall} x \right\}$, chiaramente un sottospazio lineare e quindi convesso di L_μ^p; notiamo poi che una successione convergente fortemente in L_μ^p ammette una sottosuccessione dove la convergenza è anche puntuale $\mu-$quasi ovunque: questo fatto è ben noto con la misura di Lebesgue, e la stessa dimostrazione funziona anche in questo caso più generale. Ma allora X è un chiuso forte in topologia L_μ^p, e grazie alla convessità è anche un chiuso debole; essendo allora la proiezione puntuale di Dv_h sullo spazio T_μ^p, $D_\mu v_h \in X$, e quindi $w \in X$; ma allora si ha $w(x) \in T_\mu^p(x)\ \widetilde{\forall} x$. Mettendo insieme questo con quanto visto prima si ha che, per $\mu-$quasi ogni x, $w(x) \in T_\mu^p(x) \cap \left(T_\mu^p(x)\right)^\perp = \{0\}$, e quindi $w = 0$, la tesi.

■

A questo punto possiamo definire su $\mathcal{D}(\Omega)$ la norma

$$\|v\|_{\mathrm{W}_\mu^{1,p}} := \|v\|_{L_\mu^p} + \|D_\mu v\|_{L_\mu^p} \tag{2.11}$$

e chiamare $\mathrm{W}_\mu^{1,p}$ il conseguente completamento di $\mathcal{D}(\Omega)$; che questa costruzione generalizzi con qualche utilità il concetto degli spazi di Sobolev è assicurato dalla seguente

Proposizione 2.7 *a) $\mathrm{W}_\mu^{1,p}$ è naturalmente incluso in L_μ^p;*
b) si può estendere in modo unico l'operatore D_μ su tutto $\mathrm{W}_\mu^{1,p}$;
c) la (2.9) e la (2.11) valgono per ogni $v \in \mathrm{W}_\mu^{1,p}$.

Dim : *a)* Sia $v \in \mathrm{W}_\mu^{1,p}$; per definizione è possibile scrivere $v = \lim v_n$ con $v_n \in \mathcal{D}(\Omega)$, per cui v_n è di Cauchy in $\mathrm{W}_\mu^{1,p}$ e dunque a maggior ragione in L_μ^p, e quindi $v_n \xrightarrow[L_\mu^p]{} w$: a $v \in \mathrm{W}_\mu^{1,p}$ si può dunque associare $w \in L_\mu^p$. Perché questo sia ben fatto basta notare che, per come è definita la norma su $\mathrm{W}_\mu^{1,p}$, una diversa successione in $\mathcal{D}(\Omega)$ corrispondente a v porta allo stesso elemento w in L_μ^p. Perché poi questa risulti una inclusione, supponiamo che $u = \lim u_n$ e $v = \lim v_n$, con u_n e v_n in $\mathcal{D}(\Omega)$, siano associati allo stesso $w \in L_\mu^p$; ma allora $u_n - v_n \xrightarrow[L_\mu^p]{} 0$ ed esiste $z \in L_\mu^p$ tale che $D_\mu u_n - D_\mu v_n \xrightarrow[L_\mu^p]{} z$, visto che $D_\mu u_n$ e $D_\mu v_n$ sono due successioni di Cauchy; ma dal lemma 2.6 segue che $z = 0$, e allora $u = v$, il che assicura che quella che abbiamo dato è una buona inclusione.
b) Se $v \in \mathrm{W}_\mu^{1,p}$ si può scrivere $v = \lim v_n$ con $v_n \in \mathcal{D}(\Omega)$, e quindi in particolare $D_\mu v_n$ è di Cauchy ed è possibile chiamare $D_\mu v$ il suo limite; che questo sia consistente segue dal fatto che un'altra successione corrispondente a v avrebbe portato alla stessa definizione.
c) Per la (2.11), grazie ai punti *a)* e *b)*, è sufficiente un passaggio al limite; per quanto

riguarda la (2.9), invece, è possibile scrivere $v = \lim v_n$ come al solito e, ricordando che le convergenze $v_n \longrightarrow v$ e $D_\mu v_n \longrightarrow D_\mu v$ valgono in L^p_μ e che φ e $\mathrm{div}(\varphi\mu)$ appartengono a $L^{p'}_\mu$, agire per passaggio al limite in entrambi i membri dell'uguaglianza.

■

E' opportuno notare che, se $1 < p < +\infty$, L^p_μ risulta riflessivo e quindi è riflessivo anche $\mathrm{W}^{1,p}_\mu$, visto che c'è una naturale inclusione $\mathrm{W}^{1,p}_\mu \subseteq (L^p_\mu)^{n+1}$ la cui immagine è chiusa.

L'utilità di questa nozione di derivata tangenziale rispetto ad una misura è legata in gran parte al fatto che rende possibile rilassare opportunamente alcuni problemi di tipo integrale mantenendo una forma integrale; per essere più precisi, consideriamo $g : \Omega \times \mathbb{R}^n \longrightarrow \mathbb{R}$ per cui valgano le seguenti proprietà con $0 < K_1 \leq K_2$ e $1 < p < +\infty$:

$$\begin{array}{l} \text{a) } g(\cdot, z) \text{ è } \mu - \text{misurabile per ogni } z; \\ \text{b) } g(x, \cdot) \text{ è convessa per } \mu - \text{q.o. } x; \\ \text{c) } K_1|z|^p \leq g(x,z) \leq K_2\,(1 + |z|^p)\,. \end{array} \tag{2.12}$$

Interessiamoci a questo punto al funzionale su $\mathcal{D}(\Omega)$ definito da $\int g(x, Dv)\ d\mu$; è chiaro che può risultare comodo estenderlo ad uno spazio più ampio, e si può sperare che $\mathrm{W}^{1,p}_\mu$ sia quello adatto; definiamo allora, in maniera standard, $F : \mathrm{W}^{1,p}_\mu \longrightarrow [0, +\infty]$ come

$$F(v) := \begin{cases} \int g(x, Dv)\ d\mu & \text{se } v \in \mathcal{D}(\Omega) \\ +\infty & \text{altrimenti} \end{cases}$$

e $\overline{F} = \max\left\{G : \mathrm{W}^{1,p}_\mu \longrightarrow \mathbb{R} : G \text{ debolmente s.c.i.}, G \leq F\right\}$: il fatto che $\overline{F}$ sia effettivamente un massimo è chiaro, visto che l'estremo superiore di funzionali semicontinui inferiormente è anch'esso s.c.i.; il nostro obiettivo, adesso, è quello di riuscire a scrivere $\overline{F}$ in una forma integrale.

Purtroppo si nota che la proprietà utile per arrivare ad una simile scrittura sarebbe la semicontinuità nella topologia debole $\sigma(L^p_\mu, L^{p'}_\mu)$, mentre su $\overline{F}$ si ha per adesso solo quella $\sigma(\mathrm{W}^{1,p}_\mu, \mathrm{W}^{1,p'}_\mu)$; estesa F su tutto L^p_μ come identicamente $+\infty$ su $L^p/\mathrm{W}^{1,p}_\mu$, definiamo allora anche $\widetilde{F} = \max\{G : L^p_\mu \longrightarrow \mathbb{R} : G \text{ debolmente s.c.i.}, G \leq F\}$, in modo che $\widetilde{F}$ sia s.c.i. in topologia $\sigma(L^p_\mu, L^{p'}_\mu)$, e controlliamo il rapporto tra $\overline{F}$ e $\widetilde{F}$ con il

Lemma 2.8 *Si ha la seguente proprietà:*

$$\widetilde{F} = \begin{cases} \overline{F} & \text{su } \mathrm{W}^{1,p}_\mu \\ +\infty & \text{su } L^p_\mu \setminus \mathrm{W}^{1,p}_\mu. \end{cases}$$

Dim : Innanzitutto definiamo

$$\widehat{F} = \begin{cases} \overline{F} & \text{su } \mathrm{W}^{1,p}_\mu \\ +\infty & \text{su } L^p_\mu \setminus \mathrm{W}^{1,p}_\mu \end{cases}$$

in modo da dover mostrare $\widehat{F} = \widetilde{F}$; per farlo, risulta comodo dividere la dimostrazione in più passi:

i) $\widetilde{F}_{|\mathrm{W}^{1,p}_\mu} \leq \overline{F}$.

Su $\mathrm{W}^{1,p}_\mu$ è possibile mettere sia la topologia $\sigma(\mathrm{W}^{1,p}_\mu, \mathrm{W}^{1,p'}_\mu)$ che la restrizione di $\sigma(L^p_\mu, L^{p'}_\mu)$; visto che su $\mathrm{W}^{1,p}_\mu$ la norma $\| \, \|_{\mathrm{W}^{1,p}_\mu}$ è maggiore della $\| \, \|_{L^p_\mu}$, una funzione lineare da $\mathrm{W}^{1,p}_\mu$ a $\mathbb{R}$ continua secondo $\| \, \|_{L^p_\mu}$ lo è a maggior ragione secondo $\| \, \|_{\mathrm{W}^{1,p}_\mu}$; essendo poi le topologie $\sigma(\mathrm{W}^{1,p}_\mu, \mathrm{W}^{1,p'}_\mu)$ e $\sigma(L^p_\mu, L^{p'}_\mu)$ generate dagli aperti $\{x : f(x) \in (a,b)\}$ con funzioni f lineari e continue rispettivamente secondo $\| \, \|_{\mathrm{W}^{1,p}_\mu}$ e $\| \, \|_{L^p_\mu}$, la prima risulta essere più fine. Ma allora $\widetilde{F}_{|\mathrm{W}^{1,p}_\mu}$, s.c.i. in topologia $\sigma(L^p_\mu, L^{p'}_\mu)$, lo è a maggior ragione in topologia $\sigma(\mathrm{W}^{1,p}_\mu, \mathrm{W}^{1,p'}_\mu)$ e quindi, grazie alla definizione, $\overline{F} \geq \widetilde{F}_{|\mathrm{W}^{1,p}_\mu}$.

ii) $\overline{F}$ è convessa.

Intanto notiamo che F è convessa: sul sottospazio vettoriale $\mathcal{D}(\Omega)$, infatti, lo è perché g è convessa sul secondo argomento per $\mu-$q.o. x, mentre fuori da $\mathcal{D}(\Omega)$ vale identicamente $+\infty$. Fissati $t \in (0,1)$ e $x \in \Omega$, chiamiamo allora

$$\tau(y) := \frac{\overline{F}(tx + (1-t)y) - tF(x)}{1-t};$$

visto che τ è ottenuta da $\overline{F}$ con un'omotetia, una traslazione ed un cambio di coordinate, essendo $\overline{F}$ s.c.i. in topologia $\sigma(\mathrm{W}^{1,p}_\mu, \mathrm{W}^{1,p'}_\mu)$ lo è anche τ; inoltre, poiché per definizione $\overline{F} \leq F$, grazie alla convessità di F si ha

$$\begin{aligned} tF(x) + (1-t)F(y) &\geq F(tx + (1-t)y) \geq \overline{F}(tx + (1-t)y) \Longrightarrow \\ F(y) &\geq \frac{\overline{F}(tx + (1-t)y) - tF(x)}{1-t} = \tau(y) \Longrightarrow F \geq \tau. \end{aligned}$$

Allora τ è una funzione s.c.i. minore di F, e visto che $\overline{F}$ è il massimo di tali funzioni si ha $\overline{F} \geq \tau$, che assicura

$$\frac{\overline{F}(tx + (1-t)y) - tF(x)}{1-t} \leq \overline{F}(y) \Longrightarrow \frac{\overline{F}(tx + (1-t)y) - (1-t)\overline{F}(y)}{t} \leq F(x). \quad (2.13)$$

Visto che questa disuguaglianza vale per ogni scelta di t, x ed y, fissiamo ora t ed y e definiamo

$$\psi(x) = \frac{\overline{F}(tx + (1-t)y) - (1-t)\overline{F}(y)}{t}:$$

esattamente come fatto sopra con τ, si nota subito che ψ è s.c.i. e, grazie a (2.13), si ha anche $\psi \leq F$ e quindi $\psi \leq \overline{F}$. Ma questo è come dire che, per ogni t, x ed y,

$$\frac{\overline{F}(tx + (1-t)y) - (1-t)\overline{F}(y)}{t} \leq \overline{F}(x) \Longrightarrow t\overline{F}(x) + (1-t)\overline{F}(y) \geq \overline{F}(tx + (1-t)y),$$

e quindi, come asserito, $\overline{F}$ è convessa.

iii) $\widehat{F}$ è s.c.i. nella topologia forte di L^p_μ.

La topologia di L^p_μ è metrizzabile, per cui la semicontinuità inferiore è equivalente alla semicontinuità inferiore sequenziale; si prenda allora una successione $v_h \underset{L^p_\mu}{\longrightarrow} v$: dobbiamo mostrare $\widehat{F}(v) \le \underline{\lim}\, \widehat{F}(v_h)$. Innanzitutto si estragga una sottosuccessione per la quale $\widehat{F}$ converga al liminf, chiamandola $w_i = v_{h(i)}$; se $\widehat{F}(w_i) \longrightarrow +\infty$ la disuguaglianza è certamente verificata, altrimenti si può supporre –eventualmente eliminando i primi termini della successione– che sia $\widehat{F}(w_i) \le C\ \forall i$. Ma allora la successione è contenuta in $\mathrm{W}^{1,p}_\mu$, visto che al di fuori $\widehat{F}$ vale $+\infty$; si definisca allora $\tau : \mathrm{W}^{1,p}_\mu \longrightarrow \mathbb{R}$ tramite la $\tau(w) = K_1 \|D_\mu w\|^p_{L^p_\mu}$, dove la costante K_1 è quella introdotta in (2.12 c). Il fatto che τ sia s.c.i. forte in norma $\|\ \|_{\mathrm{W}^{1,p}_\mu}$ è immediato, visto che equivale alla semicontinuità inferiore di $\|D_\mu w\|_{L^p_\mu}$ e che questa discende immediatamente dalla definizione della norma su $\mathrm{W}^{1,p}_\mu$; inoltre τ è convessa, grazie alla convessità della norma L^p_μ e dell'elevazione alla potenza $p-$esima: ma allora τ, s.c.i. forte e convessa, è anche s.c.i. debole $\sigma(\mathrm{W}^{1,p}_\mu, \mathrm{W}^{1,p'}_\mu)$. Si ha poi anche $\tau \le F$: fuori da $\mathcal{D}(\Omega)$ non c'è nulla da mostrare, visto che $F = +\infty$, mentre in $\mathcal{D}(\Omega)$, grazie alla (2.12 c), si ha

$$F(w) = \int g(x, Dw)\, d\mu \ge \int K_1 |Dw|^p\, d\mu \ge K_1 \int |D_\mu w|^p\, d\mu = K_1 \|D_\mu w\|^p_{L^p_\mu} = \tau(w),$$

visto che $D_\mu w$ è ottenuto proiettando Dw sullo spazio tangente e quindi $|D_\mu w| \le |Dw|$. Da tutto questo segue, come già visto prima, $\tau \le \overline{F}$ e quindi, su $\mathrm{W}^{1,p}_\mu$, $\tau \le \widehat{F}$; allora τ è limitato sulla successione w_i, visto che $\widehat{F}$ lo è, e quindi $\|D_\mu w_i\|_{L^p_\mu}$ è limitato; per la riflessività di L^p_μ segue che si può estrarre un'ulteriore sottosuccessione in modo che $D_\mu w_i \underset{L^p_\mu}{\rightharpoonup} \psi$; ma per ipotesi $w_i \underset{L^p_\mu}{\longrightarrow} v$, e quindi a maggior ragione $w_i \underset{L^p_\mu}{\rightharpoonup} v$; assumendo $v \in \mathrm{W}^{1,p}_\mu$, grazie al lemma 2.6 si ottiene che $\psi = D_\mu v$, e quindi $w_i \underset{\mathrm{W}^{1,p}_\mu}{\rightharpoonup} v$. Visto ora che $\overline{F}$ è debolmente s.c.i. $\sigma(\mathrm{W}^{1,p}_\mu, \mathrm{W}^{1,p'}_\mu)$, se ne deduce

$$\widehat{F}(v) = \overline{F}(v) \le \underline{\lim}\, \overline{F}(w_i) = \underline{\lim}\, \widehat{F}(w_i) = \lim \widehat{F}(w_i) = \underline{\lim}\, \widehat{F}(v_h);$$

la s.c.i. forte di $\widehat{F}$ è dunque provata pur di mostrare $v \in \mathrm{W}^{1,p}_\mu$, proprietà che si può ottenere con una costruzione simile a quella del lemma di Mazur. Indicando infatti per un momento con $\overline{}$ e $\overline{\overline{}}$ rispettivamente la chiusura debole e la forte L^p_μ e con $co(\ )$ la convessificazione, dal fatto che $D_\mu w_i \rightharpoonup \psi$ si ottiene che $\psi \in \overline{\{D_\mu w_i\}}$, e quindi a maggior ragione $\psi \in \overline{co\,(\{D_\mu w_i\})}$; ma la chiusura di un insieme convesso è banalmente convessa con una topologia metrizzabile, per cui $\overline{\overline{co\,(\{D_\mu w_i\})}}$ è convessa e fortemente chiusa, per cui debolmente chiusa: allora $\overline{co\,(\{D_\mu w_i\})} = \overline{\overline{co\,(\{D_\mu w_i\})}}$ (in generale, in uno spazio di Banach, la chiusura debole e la forte di un insieme convesso coincidono). Ne segue immediatamente, come nella dimostrazione del lemma di Mazur, che esistono due successioni crescenti M_i ed N_i, e opportuni λ_j^i, tali che $\sum_{j=M_i}^{N_i} \lambda_j^i = 1$ e che, chiamando $\widetilde{w}_i = \sum_{j=M_i}^{N_i} \lambda_j^i w_j$, si abbia

$$D_\mu \widetilde{w}_i \underset{L^p_\mu}{\longrightarrow} \psi.$$

Visto poi che la successione w_i tende fortemente a v in norma L^p_μ e che $\widetilde{w}_i$ sono combinazioni convesse crescenti, si ha che anche $\widetilde{w}_i$ tende fortemente a v; visto che $\mathrm{W}^{1,p}_\mu$ è completo, da questo segue che $\psi = D_\mu v$ e che $v \in \mathrm{W}^{1,p}_\mu$.

iv) $\widehat{F} = \widetilde{F}$.

Visto che $\overline{F}$ è convesso, $\widehat{F}$ è convessa su un sottospazio vettoriale del suo dominio, e vale identicamente $+\infty$ sul complementare, per cui è convessa. Essendo poi anche s.c.i. fortemente nella topologia data da $\| \ \|_{L^p_\mu}$, risulta s.c.i. debole; è anche ovvio che $\widehat{F} \leq F$, e da questo segue per definizione che $\widetilde{F} \geq \widehat{F}$; allora si ha chiaramente che, almeno in $L^p_\mu/\mathrm{W}^{1,p}_\mu$, $\widetilde{F} = \widehat{F}$. Su $\mathrm{W}^{1,p}_\mu$, d'altra parte, $\widehat{F} = \overline{F} \geq \widetilde{F}$ grazie al punto i), e quindi si ottiene che l'uguaglianza $\widetilde{F} = \widehat{F}$ è verificata su tutto L^p_μ, trovando così la tesi.

■

Questo lemma permetterà di esprimere, come volevamo, il funzionale $\overline{F}$ in una forma di tipo integrale: definito infatti

$$g_\mu(x,z) := \inf \left\{ g(x, z+\xi) : \xi \in T^p_\mu(x)^\perp \right\},$$

si può concludere questa sezione con il

Teorema 2.9

$$\overline{F}(u) = \int g_\mu(x, D_\mu u) \ d\mu.$$

<u>Dim</u> : Consideriamo lo spazio $\mathrm{W}^{1,p}_\mu$ come sottospazio di L^p_μ, quindi con la topologia data dalla norma $\| \ \|_{L^p_\mu}$ e messo in dualità con $L^{p'}_\mu$; dal lemma 2.8 sappiamo che $\overline{F}$ è s.c.i. debole, ed è il sup delle funzioni s.c.i. deboli minori di F.

i)$F^{**} = \overline{F}$.

Infatti F^{**} è il sup delle funzioni s.c.i. (fortemente) e convesse, che come noto sono anche debolmente s.c.i., per cui F^{**} è minore o uguale al sup delle funzioni debolmente s.c.i. minori di F e quindi $F^{**} \leq \overline{F}$; d'altra parte $\overline{F}$ è convesso, come visto all'interno della dimostrazione del lemma 2.8, e quindi la semicontinuità inferiore debole assicura quella forte, per cui si ha $F^{**} \geq \overline{F}$ e dunque l'uguaglianza asserita.

Definiti ora $X = \mathrm{W}^{1,p}_\mu$, $Y = L^p_\mu(\Omega, \mathbb{R}^n)$, $D(A) = \mathcal{D}(\Omega)$, $A(\varphi) = D\varphi$ e $f(\varphi) = \int g(x,\varphi) \ d\mu$, si ha

ii) $F^*(v) = \inf_{\eta \in L^{p'}_\mu} \{f^*(\eta) : A^*(\eta) = v\}$.

Vorremmo applicare il teorema A.5; si noti che, dal momento che anche su $\mathrm{W}^{1,p}_{\mu}$ usiamo la norma $\| \ \|_{L^p_\mu}$, l'operatore A non è continuo. Perchè si possa applicare il teorema bisogna controllare che f sia convessa e continua in un punto dell'immagine di $\mathcal{D}(\Omega)$: la convessità è immediata dalla (2.12 b), per cui mostriamo la continuità di f in 0; sia dunque data una successione $\varphi_n \underset{L^p_\mu}{\longrightarrow} 0$: grazie a (2.12 b) e (2.12 c), per $\mu-$quasi ogni x $g(x,\cdot)$ è convessa e a valori reali, dunque continua; data allora una sottosuccessione qualsiasi di φ_n, da questa se ne può estrarre un'altra per cui $\varphi_n(x)$ tenda a 0 quasi ovunque: in questo modo per quasi ogni x $g(x,\varphi_n(x)) \to g(x,0)$; si estragga poi un'ulteriore sottosuccessione in modo che $\sum \|\varphi_n\|_{L^p_\mu} \leq +\infty$, per cui $\sum |\varphi_n|$ è una funzione in L^p_μ; ora

$$g(x,\varphi_n(x)) \leq C_2(1+|\varphi_n(x)|^p) \leq C_2(1+\sum |\varphi_n(x)|^p)$$

e quindi, per convergenza dominata, si ha che

$$f(\varphi_n) = \int g(x,\varphi_n)\ d\mu \longrightarrow \int g(x,0)\ d\mu = f(0).$$

Dunque da ogni sottosuccessione di φ_n se ne può estrarre una per cui f tende a $f(0)$ e questo, come noto, assicura che $f(\varphi_n) \to f(0)$ e quindi che f è continua in 0, come volevamo: si può allora applicare il teorema A.5 ottenendo effettivamente l'uguaglianza asserita.

iii) $f^*(\eta) = \int g^*(x,\eta)\ d\mu$.

Visto che $g(x,\cdot) : \Omega \longrightarrow \mathbb{R}$, per definizione $g^*(x,p) = \sup_{y\in\Omega}\{p\cdot y - g(x,y)\}$, e quindi per ogni $\varphi \in L^p_\mu$ si ha

$$\begin{aligned}\int g^*(x,\eta)\ d\mu &= \int \sup_y \{\eta(x)\cdot y - g(x,y)\}\, d\mu \geq \int \eta(x)\cdot\varphi(x) - g(x,\varphi(x))\ d\mu \\ &= \langle \eta,\varphi\rangle - f(\varphi);\end{aligned}$$

prendendo l'estremo superiore sulle possibili φ si ottiene dunque $\int g^*(x,\eta)\ d\mu \geq f^*(\eta)$; per mostrare l'altra disuguaglianza supponiamo che $\int g^*(x,\eta)\ d\mu > -\infty$, altrimenti è vera a vuoto; si costruisca inoltre una funzione misurabile φ in modo che per quasi ogni x si abbia

$$\eta(x)\cdot\varphi(x) - g(x,\varphi(x)) \geq \sup_y \{\eta(x)\cdot y - g(x,y)\} - \epsilon :$$

se mostriamo che la φ così costruita appartiene ad L^p_μ abbiamo chiaramente la disuguaglianza cercata; sia allora $\Omega_n = \{x \in \Omega : |\varphi(x)| \leq n\}$: visto che gli Ω_n invadono Ω si ha

$$\lim_{n\to+\infty} \int_{\Omega_n} \varphi\cdot\eta - g(x,\varphi)\ d\mu = \int_\Omega \varphi\cdot\eta - g(x,\varphi)\ d\mu \geq \int_\Omega g^*(x,\eta)\ d\mu - \epsilon\|\mu\| > -\infty. \quad (2.14)$$

Su ogni Ω_n, tuttavia, φ risulta in L^p_μ, e quindi

$$\int_{\Omega_n} \varphi \cdot \eta - g(x,\varphi) \leq \|\varphi\|_{L^p_\mu(\Omega_n)} \|\eta\|_{L^{p'}_\mu} - K_1 \|\varphi\|^p_{L^p_\mu(\Omega_n)} = \|\varphi\|_{L^p_\mu(\Omega_n)} \left(\|\eta\|_{L^{p'}_\mu} - K_1 \|\varphi\|^{p-1}_{L^p_\mu(\Omega_n)} \right);$$

essendo $p > 1$ questo, grazie a (2.14), assicura che $\|\varphi\|_{L^p_\mu(\Omega_n)} < K$, il che mostra che φ è in L^p_μ e dunque la disuguaglianza che cercavamo.

iv) $\overline{F}(u) = \int g_\mu(x, D_\mu u)\, d\mu$.

Dai punti ii) e iii) si ha che $F^*(v) = \inf_{\eta \in L^{p'}_\mu} \left\{ \int g^*(x,\eta)\, d\mu : A^*(\eta) = v \right\}$; ma per definizione

$$\langle A^*(\eta), \psi \rangle_X = \langle \eta, A(\psi) \rangle_Y = \langle -\mathrm{div}(\eta\mu), \psi \rangle$$

e quindi $A^*(\eta) = -\mathrm{div}(\eta\mu)$; allora, se $v \in L^{p'}_\mu(\Omega)$, $A^*(\eta) = v$ è possibile solo se $\eta \in X^{p'}_\mu$. Si può dunque calcolare

$$\begin{aligned}
\overline{F}(u) &= F^{**}(u) = \sup_{v \in L^{p'}_\mu} \{ \langle v, u \rangle - F^*(v) \} \\
&= \sup_{v \in L^{p'}_\mu} \left\{ \langle v, u \rangle - \inf_{\eta \in L^{p'}_\mu} \left\{ \int g^*(x,\eta)\, d\mu : A^*(\eta) = v \right\} \right\} \\
&= \sup_{v \in L^{p'}_\mu} \sup_{\eta \in X^{p'}_\mu} \left\{ \langle v, u \rangle - \int g^*(x,\eta)\, d\mu : -\mathrm{div}(\eta\mu) = v \right\} \\
&= \sup_{\eta \in X^{p'}_\mu} \left\{ \langle -\mathrm{div}(\eta\mu), u \rangle - \int g^*(x,\eta)\, d\mu \right\} = \sup_{\eta \in X^{p'}_\mu} \left\{ \int D_\mu u \cdot \eta - g^*(x,\eta)\, d\mu \right\}.
\end{aligned}$$

Come nel lemma 2.6, si può portare il sup dentro l'integrale grazie al corollario A.7, essendo stavolta $j(x,z) = D_\mu u(x) \cdot z - g^*(x,z)$. Per vedere che le ipotesi siano verificate, si noti anzitutto che $g^*(x,\cdot)$ è convesso e s.c.i., per cui $j(x,\cdot)$ è concavo e s.c.s.; il fatto che se $\eta \in X^{p'}_\mu$ e $\psi \in C^\infty(\overline{\Omega}) \cap L^\infty$ allora $\eta\psi \in X^{p'}_\mu$ è evidente e comunque è già stato notato nella dimostrazione del lemma 2.6; bisogna infine notare che, per ogni $\eta \in X^{p'}_\mu$, $\int D_\mu u \cdot \eta - g^*(x,\eta)\, d\mu$ sia finito: visto che $\left|\int D_\mu u \cdot \eta\right| \leq \|D_\mu u\|_{L^p_\mu} \|\eta\|_{L^{p'}_\mu}$, è sufficiente far vedere che $\left|\int g^*(x,\eta)\right| < +\infty$; ora però grazie a (2.12 c) si ha che $K_1|z|^p \leq g(x,z) \leq K_2(1 + |z|^p)$, e questo assicura $\widetilde{K}_2(|q|^{p'} - 1) \leq g^*(x,q) \leq \widetilde{K}_1|q|^{p'}$; ma allora, visto che $\eta \in L^{p'}_\mu$, anche $\int g^*(x,\eta)$ è finito. Si può dunque, come asserito, portare il sup dentro l'integrale ottenendo

$$\overline{F}(u) = \int \sup_{p \in T^p_\mu(x)} \{ D_\mu u(x) \cdot p - g^*(x,p) \}\, d\mu;$$

chiamando allora $\tilde{g}(x,z) = \sup_{p \in T^p_\mu(x)} \{ z \cdot p - g^*(x,p) \}$, si ha che $\overline{F}(u) = \int \tilde{g}(x, D_\mu u)$, e quindi ci siamo ricondotti a mostrare che $\tilde{g} = g_\mu$. Ma questa è una conseguenza immediata del teorema A.4, e quindi abbiamo ottenuto la tesi.

$\blacksquare$

2.3 Dimensione dello spazio tangente

Nella prima sezione di questo capitolo abbiamo introdotto due nozioni di dimensione di una misura e le abbiamo confrontate, mentre nella seconda sezione abbiamo studiato i concetti di derivata tangenziale e di spazio tangente ad una misura in un punto. Dal momento che quest'ultimo è un sottospazio vettoriale di $\mathbb{R}^n$, si potrebbe pensare come dimensione puntuale in x la dimensione dello spazio tangente, e definire quindi una terza nozione di dimensione di una misura come l'inf $\mu-$essenziale delle dimensioni puntuali; questa definizione, tuttavia, dipenderebbe dalla scelta di p, dal momento che, almeno a priori, T_μ^p dipende da p; in realtà in tutti gli esempi semplici lo spazio tangente è indipendente da p, e non è nemmeno chiaro se vi siano situazioni in cui può non essere così. Per evitare comunque questo problema, risulta utile introdurre un'estensione del concetto dei campi tangenti e degli spazi tangenti come suggerito in [6]: definiamo infatti lo spazio dei campi tangenti a μ come

$$X_\mu := \left\{\varphi : \Omega \longrightarrow \mathbb{R}^n : \varphi \in L^1_{\mathrm{loc}}(\mu),\ \operatorname{div}(\varphi\mu) \in \mathcal{M}\right\}$$

e di conseguenza lo spazio tangente a μ come

$$T_\mu := (\mu - \mathrm{ess}) \bigcup \left\{\varphi(\cdot) : \varphi \in X_\mu\right\}.$$

Chiaramente queste nuove nozioni sono più generali delle precedenti, e quindi per ogni p si hanno le inclusioni $X_\mu^p \subseteq X_\mu$ e $T_\mu^p(x) \subseteq T_\mu(x)$; negli esempi più semplici, in effetti, si nota che $X_\mu^p \subsetneq X_\mu$ ma $T_\mu^p(x) = T_\mu(x)$. Ripetendo lo studio fatto nella scorsa sezione, poi, si può mostrare che anche in questo caso lo spazio tangente è un sottospazio vettoriale di $\mathbb{R}^n$ per $\mu-$q.o. x; consideriamo allora come dimensione puntuale la dimensione di $T_\mu(x)$ e diamo una terza nozione di dimensione per la misura μ, la *dimensione tangenziale*, che sarà l'inf $\mu-$essenziale delle dimensioni puntuali. Poiché la dimensione tangenziale è evidentemente intera, non si può certo sperare che coincida con le altre due; in questa sezione, in particolare, mostreremo che è minore, dando anche esempi espliciti in cui lo è strettamente, basandoci sui risultati di Fragalà e Mantegazza in [13]. Grazie alle inclusioni notate, inoltre, per ogni p la dimensione di $T_\mu^p(x)$ non supera quella di $T_\mu(x)$, e quindi lo stesso risultato varrà a maggior ragione con T_μ^p al posto di T_μ.

Per prima cosa dato $x \in supp(\mu)$ e $\rho > 0$ consideriamo su B, la palla unitaria di $\mathbb{R}^n$, la misura

$$\mu_{x,\rho} := \frac{1}{\mu(B(x,\rho))}\, \tau_{\#}\mu,$$

dove τ è l'applicazione che manda $B(x,\rho)$ in B tramite una traslazione ed un'omotetia: in pratica le $\mu_{x,\rho}$ sono le restrizioni di μ alle palle in Ω, considerate tutte sulla palla unitaria ed riscalate in modo che abbiano massa unitaria per poterle confrontare; si noti che supporre x appartenente al supporto di μ assicura che nella definizione il denominatore non si annulli. Definiamo poi lo spazio della misure tangenti a μ in x, $\mathrm{Tan}(\mu, x)$, come l'insieme delle misure ν su B tali che esiste una successione $\rho_i \searrow 0$ per cui $\mu_{x,\rho_i} \xrightarrow{*} \nu$, utillizzando la dualità con le funzioni continue; si noti che, visto che le $\mu_{x,\rho}$ sono tutte di massa unitaria, risultano un insieme debolmente* compatto, e quindi $\mathrm{Tan}(\mu, x)$ non è vuoto; in particolare nei punti dotati di maggior regolarità –ad esempio i punti in cui μ risulta una funzione continua– consiste di una sola misura. Questa nozione di spazio delle misure tangenti a μ in un punto è un caso particolare di quella proposta da Preiss in [20].

Dimostriamo intanto il seguente lemma, nel quale per comodità indicheremo con $\mathcal{H}^r_V$ la misura di Hausdorff $r-$dimensionale ristretta al sottospazio vettoriale $V \subseteq \mathbb{R}^n$:

Lemma 2.10 *Per $\mu-$q.o. x, se $\nu \in \mathrm{Tan}(\mu, x)$ si può scrivere $\nu = \left(\mathcal{H}^{k(x)}_{T_\mu(x)} \times \sigma\right) \llcorner B$, dove $k(x)$ è la dimensione di $T_\mu(x)$ e σ è una misura su $T_\mu(x)^\perp$.*

<u>Dim</u> : **i)** Fissata $\varphi \in X_\mu$, per $\mu-$q.o. x si ha $D_{\varphi(x)}\nu = 0$ per ogni $\nu \in \mathrm{Tan}(\mu, x)$.
Per mostrare che $D_{\varphi(x)}\nu = 0$, basta mostrare che per una generica $\psi \in C^\infty_0$ si ha $\langle D_{\varphi(x)}\nu, \psi\rangle = 0$; essendo ovviamente la derivata intesa in senso distribuzionale, questo è come dire $\int \partial\psi(y)/\partial\varphi(x)\ d\nu(y) = 0$, che si può anche riscrivere più espressivamente come

$$\int_B \nabla\psi(y) \cdot \varphi(x)\ d\nu(y) = 0;$$

visto però che ν appartiene a $\mathrm{Tan}(\mu, x)$, esiste una successione ρ_i tendente a 0 tale che $\mu_{x,\rho_i} \xrightarrow{*} \nu$; definendo allora $\psi_i(y) = \psi(\frac{y-x}{\rho_i})$, si ha chiaramente che

$$\nabla\psi_i(y) = \frac{1}{\rho_i}\nabla\psi(\frac{y-x}{\rho_i}), \tag{2.15}$$

e quindi

$$\begin{aligned} -\langle D_{\varphi(x)}\nu, \psi\rangle &= \int_B \nabla\psi(y)\cdot\varphi(x)\ d\nu(y) = \lim_{i\to\infty}\int_B \nabla\psi(y)\cdot\varphi(x)\ d\mu_{x,\rho_i}(y) \\ &= \lim_{i\to\infty}\int_{B(x,\rho_i)} \nabla\psi\left(\frac{y-x}{\rho_i}\right)\cdot\varphi(x)\ \frac{d\mu(y)}{\mu(B(x,\rho_i))} \\ &= \lim_{i\to\infty}\frac{\rho_i}{\mu(B(x,\rho_i))}\int_{B(x,\rho_i)} \nabla\psi_i(y)\cdot\varphi(x)\ d\mu(y). \end{aligned} \tag{2.16}$$

Visto che la variabile di integrazione è la y, il termine $\varphi(x)$ è un vettore costante che potrebbe anche essere estratto dall'integrale e dal limite: sarebbe invece molto più utile se fosse $\varphi(y)$; senza perdita di generalità, però, si può supporre che x sia un punto di Lebesgue per φ, in modo che la successione definita da

$$\lambda_i := \frac{1}{\mu(B(x,\rho_i))}\int_{B(x,\rho_i)} |\varphi(y) - \varphi(x)|\ d\mu$$

tenda a 0. Allora per una generica $f \in L^\infty$ si ha immediatamente che

$$\left|\frac{1}{\mu(B(x,\rho_i))}\int_{B(x,\rho_i)} f(y)\varphi(y)\,d\mu - \frac{1}{\mu(B(x,\rho_i))}\int_{B(x,\rho_i)} f(y)\varphi(x)\,d\mu\right| \le \|f\|_{L^\infty}\lambda_i; \quad (2.17)$$

visto poi che $\psi \in C_0^\infty$, si ha $\|\nabla\psi\|_{L^\infty} < +\infty$; per ogni i, inoltre, grazie a (2.15) vale la stima $\|\rho_i \nabla\psi_i\|_{L^\infty} \le \|\nabla\psi\|_{L^\infty}$, e quindi dalla (2.17) si ottiene

$$\left|\frac{\rho_i}{\mu(B(x,\rho_i))}\int_{B(x,\rho_i)} \nabla\psi_i(y)\varphi(y)\,d\mu - \frac{\rho_i}{\mu(B(x,\rho_i))}\int_{B(x,\rho_i)} \nabla\psi_i(y)\varphi(x)\,d\mu\right| \le \|\nabla\psi\|_{L^\infty}\lambda_i;$$

la (2.16) allora assicura che

$$\begin{aligned} \left|\langle D_{\varphi(x)}\nu, \psi\rangle\right| &= \cdots = \lim_{i\to\infty} \frac{\rho_i}{\mu(B(x,\rho_i))}\left|\int_{B(x,\rho_i)} \nabla\psi_i(y)\varphi(y)\,d\mu\right| \\ &\le \lim_{i\to\infty} \frac{\rho_i}{\mu(B(x,\rho_i))}\langle|\mathrm{div}(\varphi\mu)|, |\psi_i|\rangle \le \lim_{i\to\infty} C\rho_i \frac{\gamma(B(x,\rho_i))}{\mu(B(x,\rho_i))}, \end{aligned}$$

essendo $C = \|\psi\|_{L^\infty} \ge \|\psi_i\|_{L^\infty}$ e $\gamma = |\mathrm{div}(\varphi\mu)|$, che è una misura perché $\varphi \in X_\mu$.

A questo punto, grazie al teorema di Radon–Nikodym, si può decomporre $\gamma = \gamma_s + h\mu$, dove $\gamma_s \perp \mu$ e $h \in L^1_\mu$; a meno di escludere un insieme $\mu-$trascurabile, allora, si può supporre che x non appartenga al supporto di γ_s e che sia un punto di Lebesgue per h; allora $\gamma(B(x,\rho_i))/\mu(B(x,\rho_i)) \to h(x)$ e dunque, visto che $\rho_i \to 0$,

$$\lim_{i\to\infty} C\rho_i \,\frac{\gamma(B(x,\rho_i)}{\mu(B(x,\rho_i)} = 0.$$

Ricapitolando, abbiamo mostrato che, scelta $\varphi \in X_\mu$, per $\mu-$q.o. x si ha la proprietà che $D_{\varphi(x)}\nu = 0$ per ogni $\nu \in \mathrm{Tan}(\mu, x)$.

ii) Per $\mu-$q.o. x, $D_v\nu = 0\ \forall\, v \in T_\mu(x)$, $\forall\, \nu \in \mathrm{Tan}(\mu, x)$.

Grazie alle proprietà dell'unione $\mu-$essenziale, come già notato in (2.7) si ha che se una multifunzione Σ è tale che $\varphi(x) \in \Sigma(x)$ per $\mu-$q.o. x per una qualunque $\varphi \in X_\mu$, allora $T_\mu(x) \subseteq \Sigma(x)$ per $\mu-$q.o. x. Definito allora $\Sigma(x) = \{v : D_v\nu = 0 \forall\, \nu \in \mathrm{Tan}(\mu, x)\}$, da i) si ottiene quanto asserito.

iii) $D_v\nu = 0 \iff \nu = (\mathcal{H}^1_{<v>} \times \sigma) \llcorner B$, essendo σ una misura su $<v>^\perp$.

Consideriamo per comodità un sistema di coordinate in cui il primo asse sia parallelo a v e chiamiamo z e w rispettivamente la prima coordinata e il blocco delle altre. Si nota subito che $\Longleftarrow$ è molto semplice: se $\psi \in C_0^\infty$,

$$\begin{aligned} -\langle D_v\nu, \psi\rangle &= \iint \frac{\partial\psi}{\partial v}(z,w)\,d\nu(z,w) = \iint \frac{\partial\psi}{\partial v}(z,w)\,d\mathcal{H}^1_{<v>} \times \sigma(z,w) \\ &= \int_w d\sigma(w) \int_z \frac{\partial\psi(\cdot,w)}{\partial z}(z)\,dz = 0, \end{aligned}$$

visto che anche $\psi(\cdot, w)$ è una funzione C_0^∞.

Per mostrare $\Longrightarrow$ conviene innanzitutto operare una disintegrazione di ν, come illustrato ad esempio in [1], che permette di scrivere

$$\iint \varphi(z,w)\ d\nu(z,w) = \int_w d\sigma(w) \int_z \varphi(z,w)\ d\nu_w(z),$$

dove $\sigma = \pi_{1\#}\nu$ e per ogni w è $\nu_w(\mathbb{R}) = 1$; chiamiamo allora $z_0 = \sqrt{1-|w|^2}$ e, definita su $< v >^\perp$ la misura $\tilde{\sigma} = \sigma/2z_0$, ci proponiamo di mostrare che $\nu = (\mathcal{H}^1_{<v>} \times \tilde{\sigma}) \llcorner B$. Per poterlo fare, presa $\varphi \in C_0^\infty$ diamo le seguenti definizioni:

$$\tau(w) := \int_{-z_0}^{z_0} \varphi(z,w)\ dz \qquad\qquad \psi(z,w) := \int_{-z_0}^{z} \varphi(s,w)\ ds\ - \frac{z+z_0}{2z_0}\ \tau(w);$$

la ψ così definita non è in C_0^∞, però si annulla sul bordo e quindi, per densità, si ottiene $\langle D_v\nu, \psi\rangle = 0$; inoltre $\partial\psi/\partial z\ (z,w) = \varphi(z,w) - \tau(w)/2z_0$, da cui risulta

$$0 = \iint \frac{\partial\psi}{\partial z}(z,w)\ d\nu(z,w), \qquad \text{e quindi}$$
$$\iint\varphi(z,w)\ d\nu(z,w) = \iint \frac{\tau(w)}{2z_0}\ d\nu(z,w) = \int_w \frac{\tau(w)}{2z_0}\ d\sigma(w) \int_z d\nu_w(z) = \int_w \tau(w)\ d\tilde{\sigma}(w)$$
$$= \int_w d\tilde{\sigma}(w) \int_z \varphi(z,w)\ d\mathcal{H}^1_{<v>} = \iint \varphi(z,w)\ d\left(\mathcal{H}^1_{<v>} \times \tilde{\sigma}\right)(z,w);$$

dunque le misure ν e $(\mathcal{H}^1_{<v>} \times \tilde{\sigma}) \llcorner B$ coincidono se applicate alle funzioni C_0^∞, e dunque sono uguali.

iv) Mostriamo ora la tesi.

Grazie a ii) e iii), per $\mu-$q.o. x è possibile trovare per ν una decomposizione del tipo $(\mathcal{H}^1_{<v>} \times \sigma) \llcorner B$ per ogni vettore $v \in T_\mu(x)$; fissato x per cui questo vale e chiamando $k(x)$ la dimensione di $T_\mu(x)$, a meno di cambiare le coordinate si può supporre che sia $T_\mu(x) = V_{k(x)}$, indicando con V_i il sottospazio vettoriale di $\mathbb{R}^n$ generato dai primi i vettori coordinati: convenendo allora di chiamare per comodità $\mathcal{H}^k_X$ la misura di Hausdorff $k-$dimensionale ristretta non ad X, bensì alla sua intersezione con la palla unitaria B, e definito $\Gamma = \big\{i \in \{1, 2, \ldots, n\} : \exists\sigma \in \mathcal{M}(V_i{}^\perp),\ \nu = \mathcal{H}^i_{V_i} \times \sigma\big\}$, in pratica si deve mostrare che $k(x) \in \Gamma$. Per ottenere la tesi ragioniamo per induzione: $1 \in \Gamma$ per ipotesi, visto che $V_1 =<e_1>$: supponiamo dunque che $k \in \Gamma$ e $k < k(x)$; indichiamo allora con $\pi_1,\ \pi_2$ e π_3 le proiezioni sui sottospazi generati rispettivamente dalle prime k coordinate, dalla $(k+1)-$esima e dalle prime $k+1$, e con $\pi'_1,\ \pi'_2$ e π'_3 le proiezioni sui rispettivi sottospazi ortogonali. Per ipotesi induttiva, per ogni sottoinsieme $A \subseteq \mathbb{R}^n$ si ha

$$\nu(A) = \mathcal{H}^k_{V_k}(\pi_1(A)) \cdot \sigma_1(\pi'_1(A)),$$

mentre grazie a quanto già mostrato è anche

$$\nu(A) = \mathcal{H}^1_{<e_{k+1}>}(\pi_2(A)) \cdot \sigma_2(\pi_2'(A)).$$

In particolare si ha

$$\nu(V_k \times \pi_1'(A)) = \mathcal{H}^k_{V_k}(V_k) \cdot \sigma_1(\pi_1'(A)) = \mathcal{H}^1_{<e_{k+1}>}(\pi_2(A)) \cdot \sigma_2(V_k \times \pi_3'(A)),$$

da cui segue che

$$\nu(A) = \mathcal{H}^k_{V_k}(\pi_1(A)) \cdot \mathcal{H}^1_{<e_{k+1}>}(\pi_2(A)) \cdot \frac{\sigma_2(V_k \times \pi_3'(A))}{\mathcal{H}^k_{V_k}(V_k)} = \mathcal{H}^{k+1}_{V_{k+1}}(\pi_3(A)) \cdot \sigma_3(\pi_3'(A)),$$

che assicura che $\nu = \mathcal{H}^{k+1}_{V_{k+1}} \times \sigma_3$, ovvero $k+1 \in \Gamma$ e quindi si ha il passo induttivo che assicura la tesi.

■

Grazie a questo lemma si può dimostrare un primo risultato parziale sulla dimensione dello spazio tangente ad una misura, indicando come al solito con $[\alpha]$ la parte intera del numero reale α:

Lemma 2.11 *Sia $\alpha \in [0, n]$ ed $E \subseteq \Omega$: se $\forall x \in E$ si ha $\theta^*_\alpha(\mu, x) = +\infty$, allora per $\mu-$q.o. x appartenente ad E $k(x)$, la dimensione di $T_\mu(x)$, è minore di $[\alpha]$.*

Dim : **i)** Se $x \in E$, per ogni $t \in (0, 1)$ è

$$\overline{\lim_{\rho \to 0^+}} \frac{\mu(B(x, t\rho))}{\mu(B(x, \rho))} \geq t^\alpha. \tag{2.18}$$

Se non fosse così, infatti, si avrebbe $\mu(B(x, t\rho)) \leq t^\alpha \mu(B(x, \rho))$ per ogni $\rho \in (0, \bar{\rho}]$, e di conseguenza per induzione $\mu(B(x, t^n\rho)) \leq t^{n\alpha}\mu(B(x, \rho))$; dato allora $\rho \in (0, \bar{\rho}]$, sia n l'intero tale che $\rho \in (t^{n+1}\bar{\rho}, t^n\bar{\rho}]$, per cui tra l'altro $t^n/\rho < 1/(t\bar{\rho})$: allora

$$\frac{\mu(B(x, \rho))}{\rho^\alpha} \leq \frac{\mu(B(x, t^n\bar{\rho}))}{\rho^\alpha} \leq \left(\frac{t^n}{\rho}\right)^\alpha \mu(B(x, \bar{\rho})) < \left(\frac{1}{t\bar{\rho}}\right)^\alpha \mu(B(x, \bar{\rho})) = K,$$

dove K è una costante indipendente da ρ; questo va contro l'ipotesi che $\theta^*_\alpha(\mu, x) = +\infty$, e dunque per assurdo abbiamo l'asserto.

ii) Per $\mu-$q.o. x in E si ha $k(x) \leq [\alpha]$.

Per ogni x appartenente ad E scegliamo una successione $\rho_i \searrow 0$ in modo che venga raggiunto il $\overline{\lim}$ in (2.18), e quindi estraiamo una sottosuccessione (che per comodità continueremo a chiamare ρ_i) in modo che $\mu_{x,\rho_i} \stackrel{*}{\rightharpoonup} \nu$, con $\nu \in \mathrm{Tan}(\mu, x)$; detta allora B_t la palla centrata nell'origine di raggio t, grazie alla scelta dei ρ_i e al punto i) si può stimare

$$\nu(B_t) = \lim_{i \to \infty} \mu_{x,\rho_i}(B_t) = \lim_{i \to \infty} \frac{\mu(B(x, t\rho_i))}{\mu(B(x, \rho_i))} \geq t^\alpha;$$

grazie al lemma 2.10, poi, a meno di escludere un insieme $\mu-$trascurabile di punti x si può scrivere $\nu = \left(\mathcal{H}^{k(x)}_{T_\mu(x)} \times \sigma\right) \llcorner B$; indicato poi con Q_l il cubo centrato nell'origine di semilato l, se l è sufficientemente piccolo $Q_l \subseteq B$, per cui $\nu(Q_l) \leq \nu(B) = 1$; grazie alla scrittura di ν, però, si ha $\nu(Q_l) = (2l)^{k(x)}\sigma(\pi(Q_l))$, indicando con π la proiezione su $T_\mu(x)^\perp$, per cui

$$(2l)^{k(x)}\sigma(\pi(Q_l)) \leq 1; \tag{2.19}$$

d'altra parte $B_t \subseteq Q_t$ per ogni t, per cui

$$(2t)^{k(x)}\sigma(\pi(Q_t)) = \nu(Q_t) \geq \nu(B_t) \geq t^\alpha. \tag{2.20}$$

Mettendo insieme (2.19) e (2.20) si trova che per ogni $t \leq l$ è

$$(2l)^{k(x)} \leq \frac{1}{\sigma(\pi(Q_l))} \leq \frac{1}{\sigma(\pi(Q_t))} \leq \frac{(2t)^{k(x)}}{t^\alpha} = 2^{k(x)}t^{k(x)-\alpha};$$

fissando l e mandando t a 0, se ne deduce che $k(x) - \alpha \leq 0$, e visto che $k(x)$ è un numero intero ne segue la tesi.

■

Corollario 2.12 *Se $E \subseteq \Omega$, $k(x) \leq \mathcal{H}-\mathrm{dim}(E)$ per $\mu-$q.o. $x \in E$.*

Dim : Si noti che, essendo E un insieme, con $\mathcal{H}-\mathrm{dim}$ si indica la consueta dimensione di Hausdorff sui sottoinsiemi di $\mathbb{R}^n$, non la nozione che abbiamo introdotto per le misure; sia allora $\alpha > \mathcal{H}-\mathrm{dim}(E)$: evidentemente $\theta^*_\alpha(\mu, x) = +\infty$ per $\mu-$q.o. $x \in E$, come si ottiene facilmente dal teorema 2.2, per cui la tesi discende direttamente dal lemma 2.11.

■

Da questi risultati discende finalmente il

Teorema 2.13 *La dimensione tangenziale è minore della dimensione minimale e di quella di Hausdorff.*

Dim : Visto che abbiamo mostrato nel teorema 2.4 l'equivalenza tra la dimensione minimale e quella di Hausdorff, è sufficiente il confronto con una delle due; sia allora $\alpha > \mathcal{H}-\mathrm{dim}(\mu)$: per definizione, esiste un insieme B tale che $\mu(B) > 0$ e $\mathcal{H}^\alpha(B) < +\infty$. Questo assicura che $\mathcal{H}-\mathrm{dim}(B) \leq \alpha$ e quindi, dal corollario appena mostrato, che $k(x) \leq \mathcal{H}-\mathrm{dim}(B) \leq \alpha$ per $\mu-$q.o. $x \in B$; visto che $\mu(B) > 0$, allora, vi è un insieme di $\mu-$misura positiva su cui la dimensione è al più α, e quindi dalla definizione segue che la dimensione tangenziale è minore di α; grazie alla genericità di $\alpha > \mathcal{H}-\mathrm{dim}(\mu)$, infine, si ottiene la tesi.

■

Vediamo infine un esempio in cui le diverse nozioni di dimensione sono diverse, grazie alla

Osservazione 2.14 *La minorazione in 2.13 può essere stretta.*

Dim : Già introducendo la nozione di dimensione tangenziale abbiamo notato che questa, non essendo intera, non può coincidere sempre con le altre due, pur facendolo nella maggior parte dei casi semplici; per trovare un controesempio, allora, sarà sufficiente una qualunque misura con dimensione minimale non intera. Soffermiamoci in particolare, nel caso $n = 1$, su una generalizzazione dell'insieme di Cantor classico, gli *insiemi di Cantor $\alpha-$dimensionali:* è così detto un chiuso contenuto nell'intervallo unitario che abbia $\alpha-$dimensione positiva finita e complementare denso. Notiamo intanto che è possibile trovare insiemi di questo tipo per qualunque $\alpha \in [0, 1]$: per $\alpha = 0$ va bene un insieme finito di punti, per α strettamente compreso tra 0 ed 1 si può effettuare la stessa costruzione dell'insieme di Cantor cambiando opportunamente la costante, mentre per $\alpha = 1$ va bene il complementare dei "razionali ingrassati". E' evidente che se μ è la misura $\mathcal{H}^\alpha$ ristretta ad un insieme di Cantor $\alpha-$dimensionale, si ha $\mathcal{H}-\dim(\mu) = \alpha$; la dimensione tangenziale, invece, risulta 0: trattandosi di un numero intero, questo è evidente nel caso $\alpha < 1$; se $\alpha = 1$, invece, basta mostrare che $X_\mu = \{0\}$, in modo che lo spazio tangente a μ sia formato dalla sola origine per ogni x. Sia allora $\varphi \in X_\mu$: per definizione questo assicura che $\mathrm{div}(\varphi\mu) \in \mathcal{M}([0, 1])$, e si nota facilmente che questo può succedere soltanto se $\varphi\mu \ll \mathcal{H}^1$, ed in particolare solo se la densità f di $\varphi\mu$ è assolutamente continua: visto però che certamente $f = 0$ sul complementare di C, che è denso, per continuità ne segue $\varphi\mu = 0$ e quindi $\varphi = 0$ in $L^1_{\mathrm{loc}}(\mu)$, da cui quanto asserito.

Capitolo 3

Formulazioni equivalenti per il problema di Monge–Kantorovich

In questo capitolo discuteremo due formulazioni del problema di Monge–Kantorovich presenti in letteratura piuttosto diverse da quella di Kantorovich discussa nel capitolo 1 e mostreremo che si tratta di tre problemi equivalenti. La prima è stata proposta da Bouchitté, Buttazzo e Seppecher nel 1997 e successivamente perfezionata mentre la seconda, dovuta a Brenier, è molto recente, anche se alcune idee in questa direzione si possono ritrovare già nei suoi precedenti lavori. Nelle prime due sezioni introdurremo queste formulazioni, nella terza discuteremo la loro equivalenza e nella quarta considereremo una particolare misura ottimale che gioca un importante ruolo, per motivi diversi, in ciascuno dei tre problemi.

3.1 Formulazione di Bouchitté, Buttazzo e Seppecher

Il problema di Bouchitté, Buttazzo e Seppecher è stato proposto per la prima volta in [6]; per impostarlo, chiamiamo innanzitutto $f \in \mathcal{M}(\Omega)$ una misura a media nulla e con variazione totale limitata, che corrisponde alla differenza $f^+ - f^-$ del problema di Monge–Kantorovich. Data poi una misura non negativa μ, definiamo l'energia associata ad una funzione $u \in \mathcal{D}(\Omega)$ come

$$E(\mu, u) := \frac{1}{2} \int |Du|^2 \, d\mu - \langle f, u \rangle;$$

consideriamo quindi l'estremo inferiore dell'energia, avendo fissato μ, che sarà

$$\mathcal{E}(\mu) := \inf_{u \in \mathcal{D}(\Omega)} E(\mu, u).$$

La questione che ci si pone, allora, è di trovare la misura μ per cui $\mathcal{E}(\mu)$ è massimo, considerando dunque il problema

$$(\star) := \sup\left\{\mathcal{E}(\mu) : \mu \in \mathcal{P}(\Omega)\right\}. \tag{3.1}$$

Si può anche dare un'interpretazione fisica della situazione, considerando f come la densità di una sorgente di calore fissata, μ come la densità di conducibilità termica e u come una distribuzione di temperatura: in questo caso $E(\mu, u)$ è l'energia associata alla distribuzione di temperatura u in base alla conducibilità μ, mentre $\mathcal{E}(\mu)$ è l'energia effettiva corrispondente alla conducibilità μ; si desidera, in conclusione, massimizzare questa energia.

Si può notare facilmente che esistono misure μ per le quali $\mathcal{E}(\mu) = -\infty$: per costruire un esempio basta supporre che il supporto di f abbia dimensione più piccola di quello di μ, in modo che esistano funzioni per cui $\langle f, u\rangle$ è arbitrariamente alto mentre $\int \|Du\|^2 \, d\mu$ resta limitato.

Per studiare opportunamente il problema $(\star)$ vedendolo in un'altra forma, si può intanto notare che vale la

Proposizione 3.1 *Per ogni misura $\mu \in \mathcal{P}(\Omega)$ si ha*

$$\mathcal{E}(\mu) = \sup\left\{-\frac{1}{2}\int |\sigma|^2 \, d\mu \ : \sigma \in L^2_\mu(\Omega, \mathbb{R}^n),\ -\mathrm{div}(\sigma\mu) = f\right\}.$$

Se poi $\mathcal{E}(\mu) > -\infty$, allora il sup è in effetti un massimo.

<u>Dim</u> : L'idea è di applicare il teorema A.3, essendo $X = C(\Omega)$, $Y = L^2_\mu(\Omega, \mathbb{R}^n)$, $g(u) = -\langle f, u\rangle$, $h(w) = 1/2 \int |w|^2 \, d\mu$, $D(A) = \mathcal{D}(\Omega)$ e $A(u) = Du$; ovviamente con $\langle\, , \rangle$ indichiamo la dualità tra $C(\Omega)$ e $\mathcal{M}(\Omega)$. Che le ipotesi del teorema siano verificate è immediato da notare, per cui si ha

$$\inf_{x \in D(A)} \left\{h(A(x)) + g(x)\right\} = \sup_{\tilde{y} \in D(A^*)} \left\{-g^*(-A^*(\tilde{y})) - h^*(\tilde{y})\right\}, \tag{3.2}$$

e il primo membro si può riscrivere come

$$\inf_{u \in \mathcal{D}(\Omega)} \left\{\frac{1}{2}\int |Du|^2 \, d\mu - \langle f, u\rangle\right\} = \inf_{u \in \mathcal{D}(\Omega)} E(\mu, u) = \mathcal{E}(\mu).$$

Per quanto riguarda il secondo, che con l'ipotesi $\mathcal{E}(\mu) > -\infty$ è in effetti un massimo, si può notare facilmente che $h^*(\sigma) = 1/2 \int |\sigma|^2 \, d\mu$: infatti

$$\begin{aligned} h^*(\sigma) &= \sup_{w \in L^2_\mu} \left\{\langle \sigma, w\rangle - \frac{1}{2}\int |w|^2 \, d\mu\right\} \leq \sup_{w \in L^2_\mu} \left\{\|\sigma\| \, \|w\| - \frac{1}{2}\|w\|^2\right\} \\ &= \sup_{c \in \mathbb{R}} \left\{\|\sigma\| c - \frac{1}{2}c^2\right\} = \frac{1}{2}\|\sigma\|^2 = \frac{1}{2}\int |\sigma|^2 \, d\mu, \end{aligned}$$

e d'altra parte

$$h^*(\sigma) \geq \langle \sigma, \sigma \rangle - \frac{1}{2}\int |\sigma|^2 = \frac{1}{2}\int |\sigma|^2 \, d\mu;$$

in effetti l'uguaglianza $\left(\int \varphi\right)^* = \int \varphi^*$ vale con ipotesi piuttosto generali, come si può trovare discusso in [11]. Inoltre se $\sigma \in D(A^*)$ e $u \in \mathcal{D}(\Omega)$, per definizione si ha

$$\langle A^*(\sigma), u \rangle = \langle \sigma, A(u) \rangle = \int \sigma Du \, d\mu = \langle -\mathrm{div}(\sigma\mu), u \rangle,$$

per cui $A^*(\sigma) = -\mathrm{div}(\sigma\mu)$. Allora il secondo membro in (3.2) diventa

$$\begin{aligned}
&\sup_{\sigma \in D(A^*)} \left\{ -g^*(\mathrm{div}(\sigma\mu)) - \frac{1}{2}\int |\sigma|^2 \, d\mu \right\} \\
= &\sup_{\sigma \in D(A^*)} \left\{ -\sup_{u \in C(\Omega)} \left\{ \langle \mathrm{div}(\sigma\mu), u \rangle - g(u) \right\} - \frac{1}{2}\int |\sigma|^2 \, d\mu \right\} \\
= &\sup_{\sigma \in D(A^*)} \left\{ -\sup_{u \in C(\Omega)} \langle \mathrm{div}(\sigma\mu) + f, u \rangle - \frac{1}{2}\int |\sigma|^2 \, d\mu \right\} \\
= &\sup \left\{ -\frac{1}{2}\int |\sigma|^2 \, d\mu : \sigma \in D(A^*), \ -\mathrm{div}(\sigma\mu) = f \right\}.
\end{aligned}$$

Per ottenere la tesi bisogna solo mostrare che l'estremo superiore appena trovato rimane invariato se calcolato in $Y^* = L^2_\mu(\Omega, \mathbb{R}^n)$ piuttosto che in $D(A^*)$; ma per definizione $\sigma \in D(A^*)$ se e solo se esiste una misura $\nu \in \mathcal{M}(\Omega)$ tale che

$$\langle \sigma, Du \rangle = \langle \nu, u \rangle \ \forall u \in \mathcal{D}(\Omega),$$

ovvero se e solo se esiste $\nu \in \mathcal{M}(\Omega)$ tale che $-\mathrm{div}(\sigma\mu) = \nu$ o, in altre parole, se e solo se $-\mathrm{div}(\sigma\mu) \in \mathcal{M}(\Omega)$. Da ciò segue che se $\sigma \in L^2_\mu(\Omega, \mathbb{R}^n)$ e $-\mathrm{div}(\sigma\mu) = f$ allora in effetti $\sigma \in D(A^*)$, e quindi in definitiva

$$\mathcal{E}(\mu) = \sup \left\{ -\frac{1}{2}\int |\sigma|^2 \, d\mu : \sigma \in L^2_\mu(\Omega, \mathbb{R}^n), \ -\mathrm{div}(\sigma\mu) = f \right\}.$$

■

A questo punto si definisce

$$I(f) := \sup \left\{ \langle f, \varphi \rangle : \varphi \in \mathrm{Lip}_1(\Omega) \right\}; \tag{3.3}$$

notiamo innanzitutto il

Lemma 3.2 *L'estremo superiore in (3.3) è in effetti un massimo.*

Dim : Visto che f è a media nulla, $\langle f, \varphi \rangle = \langle f, \varphi + k \rangle$ per ogni costante k; ne segue che ci si può limitare a considerare $\langle f, \varphi \rangle$ per le funzioni φ che siano $1-$lipschitziane e a media nulla; il modulo di queste funzioni si può stimare con il diametro di Ω, che supponiamo limitato, e quindi, essendo $f \in \mathcal{M}(\Omega)$, si trova innanzitutto che $I(f) < +\infty$; sia ora φ_n una successione tale che $\langle f, \varphi_n \rangle \to I(f)$: trattandosi di funzioni $1-$lipschitziane a media nulla su un dominio limitato, sono equilimitate ed equicontinue in $C(\overline{\Omega})$, e dunque per il teorema di Ascoli–Arzelà si ha che $\varphi_n \longrightarrow \varphi$ in $C(\Omega)$ a meno di passare ad una sottosuccessione. Ne segue dunque che $\langle f, \varphi \rangle = I(f)$ e da questo si ottiene la tesi, notando che φ è $1-$lipschitziana in quanto limite puntuale di funzioni che lo sono.

■

Quello che ci proponiamo di fare è mostrare che il calcolo di $(\star)$ si riconduce al calcolo del massimo in (3.3) e che anche il primo è in effetti un massimo. Una prima stima si ottiene grazie al seguente

Lemma 3.3 $(\star) \leq -I(f)^2/2$.

Dim : Scelta $\mu \in \mathcal{P}(\Omega)$, siano $\sigma \in L^2_\mu(\Omega, \mathbb{R}^n)$ tale che $-\mathrm{div}(\sigma\mu) = f$ e $\varphi \in \mathcal{D}(\Omega)$ tale che $|D\varphi| \leq 1$; allora

$$\begin{aligned} \langle f, \varphi \rangle^2 &= \langle -\mathrm{div}(\sigma\mu), \varphi \rangle^2 = \left(\int \sigma D\varphi \, d\mu \right)^2 \leq \int |\sigma|^2 \, d\mu \int |D\varphi|^2 \, d\mu \\ &\leq \int |\sigma|^2 \, d\mu \int d\mu = \int |\sigma|^2 \, d\mu, \end{aligned}$$

dal momento che $\mu \in \mathcal{P}(\Omega)$. Quindi

$$-\frac{1}{2} \int |\sigma|^2 \, d\mu \leq -\frac{\langle f, \varphi \rangle^2}{2},$$

che assicura, grazie alla proposizione 3.1 ed alla genericità di σ con $-\mathrm{div}(\sigma\mu) = f$, che

$$\mathcal{E}(\mu) \leq -\frac{\langle f, \varphi \rangle^2}{2};$$

visto poi che questo vale per ogni $\varphi \in \mathcal{D}(\Omega)$ con $|D\varphi| \leq 1$, per densità vale per tutte le funzioni $1-$lipschitziane, visto che $f \in \mathcal{M}(\Omega)$, per cui passando all'estremo inferiore si ottiene

$$\mathcal{E}(\mu) \leq \inf_{u \in \mathrm{Lip}_1} \left\{ -\frac{\langle f, u \rangle^2}{2} \right\} = -\sup_{u \in \mathrm{Lip}_1} \left\{ \frac{\langle f, u \rangle^2}{2} \right\} = -\frac{\left(\sup_{u \in \mathrm{Lip}_1} \langle f, u \rangle \right)^2}{2} = -\frac{I(f)^2}{2};$$

grazie alla genericità di $\mu \in \mathcal{P}(\Omega)$, a questo punto, passando all'estremo superiore nell'ultima equazione si trova la tesi.

■

Un'altra comoda caratterizzazione si ottiene con la seguente

Proposizione 3.4

$$I(f) = \min \left\{ \int d|\nu| : \nu \in \mathcal{M}(\Omega, \mathbb{R}^n), -\mathrm{div}(\nu) = f \right\}.$$

<u>Dim</u> : Anche per questa dimostrazione utilizzeremo il teorema A.3, con $X = C(\Omega)$, $Y = C(\Omega, \mathbb{R}^n)$, $g(u) = -\langle f, u\rangle$, $D(A) = \mathcal{D}(\Omega)$, $A(u) = Du$ e

$$h(w) = \begin{cases} 0 & \text{se } |w(x)| \leq 1 \ \forall x \\ +\infty & \text{altrimenti.} \end{cases}$$

Anche in questo caso si controlla facilmente che le ipotesi del teorema sono verificate e si ha quindi ancora

$$\inf_{x \in D(A)} \{h(A(x)) + g(x)\} = \sup_{\tilde{y} \in D(A^*)} \{-g^*(-A^*(\tilde{y})) - h^*(\tilde{y})\}\,; \tag{3.4}$$

il primo membro, inoltre, non è altro che

$$\inf_{u \in \mathcal{D}(\Omega)} \{h(Du) - \langle f, u\rangle\} = \inf\{-\langle f, u\rangle : u \in \mathcal{D}(\Omega),\ |Du| \leq 1\} = -I(f),$$

usando ancora la densità di $\{u \in \mathcal{D}(\Omega),\ |Du| \leq 1\}$ in $\mathrm{Lip}_1(\Omega)$ e il fatto che $f \in \mathcal{M}(\Omega)$. Dal momento che, come abbiamo già notato, $I(f)$ è finito, l'estremo superiore del secondo membro è in effetti un massimo; se allora $\nu \in D(A^*)$ e $u \in \mathcal{D}(\Omega)$, per definizione si ha

$$\langle A^*(\nu), u\rangle = \langle \nu, A(u)\rangle = \int Du \cdot d\nu = \langle -\mathrm{div}(\nu), u\rangle,$$

per cui $A^*(\nu) = -\mathrm{div}(\nu)$. Allora se $\nu \in D(A^*)$ si ha

$$g^*(-A^*(\nu)) = \sup_{u \in X} \{\langle -A^*(\nu), u\rangle - g(u)\} = \sup_{u \in X} \langle \mathrm{div}(\nu) + f, u\rangle,$$

e quindi $g^*(-A^*(\nu))$ è nullo se $-\mathrm{div}(\nu) = f$, mentre vale $+\infty$ altrimenti. Infine si può notare che h^* calcola la variazione totale, dal momento che

$$h^*(\nu) = \sup_{w \in Y} \{\langle \nu, w\rangle - h(w)\} = \sup \left\{ \int w \ d\nu : |w| \leq 1 \right\} = \int d|\nu|.$$

Da tutto questo segue che il secondo membro in (3.4) diventa

$$\max \left\{ -\int d|\nu| : \nu \in D(A^*),\ -\mathrm{div}(\nu) = f \right\};$$

anche stavolta per ottenere la tesi bisogna mostrare che il massimo appena trovato rimanga invariato se calcolato in $Y^* = \mathcal{M}(\Omega, \mathbb{R}^n)$ piuttosto che in $D(A^*)$; è però immediato notare che $\nu \in D(A^*)$ se e solo se $\mathrm{div}(\nu) \in \mathcal{M}(\Omega)$, e quindi in effetti

$$-I(f) = \max \left\{ -\int d|\nu| : \nu \in \mathcal{M}(\Omega, \mathbb{R}^n),\ -\mathrm{div}(\nu) = f \right\},$$

da cui, cambiando segno, si ottiene la tesi.

■

Prima di mostrare quanto ci eravamo prefissi, è utile dare una scrittura più espressiva per $\mathcal{E}(\mu)$: come discusso nella sezione 2.2, risulterebbe molto comodo poter estendere la definizione di $E(\mu, u)$ alle funzioni $u \in \mathrm{W}^{1,2}_\mu$, invece di averla solo per le funzioni regolari. Per fortuna questo è possibile, e la scrittura che si ottiene generalizza nel modo migliore $E(\mu, u)$, grazie alla

Proposizione 3.5 *Se $\mathcal{E}(\mu) > -\infty$, allora*

$$\mathcal{E}(\mu) = \inf_{u \in \mathrm{W}^{1,2}_\mu} \left\{ \frac{1}{2} \int |D_\mu u|^2 \; d\mu - \langle f, u \rangle \right\}.$$

Dim : Innanzitutto definiamo $g(x, z) = |z|^2/2$, che verifica le proprietà (2.12) con $p = 2$, e chiamiamo $F : \mathrm{W}^{1,2}_\mu \longrightarrow \mathbb{R}$ la

$$F(u) = \begin{cases} \frac{1}{2} \int |Du|^2 \; d\mu & \text{se } u \in \mathcal{D}(\Omega) \\ +\infty & \text{altrimenti} \end{cases}$$

e $\overline{F} = \max \big\{ H : \mathrm{W}^{1,2}_\mu \longrightarrow \mathbb{R} : H \text{ debolmente s.c.i.}, \; H \leq F \big\}$.
i) $\overline{F}(u) = \frac{1}{2} \int |D_\mu u|^2 \; d\mu$.
Grazie al teorema 2.9, si ha

$$\overline{F}(u) = \int g_\mu(x, D_\mu u) \; d\mu;$$

ma

$$g_\mu(x, z) = \inf \left\{ g(x, z + \xi) : \xi \in T^2_\mu(x)^\perp \right\} = \inf \left\{ \frac{1}{2} |z + \xi|^2 : \xi \in T^2_\mu(x)^\perp \right\} = \frac{1}{2} |P_\mu(x)(z)|^2,$$

indicando come nel capitolo 2 con $P_\mu(x)$ la proiezione su $T^2_\mu(x)$; d'altra parte grazie alla definizione è chiaro che per ogni $u \in \mathrm{W}^{1,2}_\mu$ si ha $D_\mu(x) \in T^2_\mu$, ragionando per densità dal momento che è vero per le funzioni regolari. Dunque

$$\overline{F}(u) = \frac{1}{2} \int |D_\mu u|^2 \; d\mu,$$

e quindi si ha quanto asserito.

Definiamo ora $G : \mathrm{W}^{1,2}_\mu \longrightarrow \mathbb{R}$ tramite

$$G(u) = \begin{cases} \frac{1}{2} \int |Du|^2 \; d\mu - \langle f, u \rangle = E(\mu, u) & \text{se } u \in \mathcal{D}(\Omega) \\ +\infty & \text{altrimenti} \end{cases}$$

e $\overline{G} = \max\{H : \mathrm{W}^{1,2}_{\mu} \longrightarrow \mathbb{R} : H \text{ debolmente s.c.i., } H \leq G\}$.
ii) $\mathcal{E}(\mu) = \inf \overline{G}$.
E' chiaro dalla definizione che $\mathcal{E}(\mu) = \inf G$; a questo punto si può notare che, essendo $\overline{G} \leq G$, è ovviamente $\inf \overline{G} \leq \inf G$, per cui si ha già una disuguaglianza; d'altra parte la funzione $H \equiv \inf G$, essendo costante, è debolmente s.c.i., ed è anche ovviamente $H \leq G$: ma allora $\overline{G} \geq H$, e quindi $\inf \overline{G} \geq \inf H = \inf G$, per cui si ha anche l'altra disuguaglianza.
iii) $f \in \mathrm{W}^{-1,2}_{\mu}$.
Se fosse $f \notin \mathrm{W}^{-1,2}_{\mu}$, per ogni $K \in \mathbb{R}$ esisterebbe $u \in \mathcal{D}(\Omega)$ tale che $\|u\|_{\mathrm{W}^{1,2}_{\mu}} \leq 1$ e $\langle f, u\rangle \geq K$; si avrebbe allora $G(u) \leq 1/2 - K$, e dunque $\mathcal{E}(\mu) = \inf G = -\infty$, contro le ipotesi.
iv) $\overline{G} = \overline{F} - \langle f, \cdot\rangle$.
Notiamo anzitutto che $u \longrightarrow \langle -f, u\rangle$ è una funzione debolmente s.c.i. da $\mathrm{W}^{1,2}_{\mu}$ a $\mathbb{R}$, dal momento che $f \in \mathrm{W}^{-1,2}_{\mu}$; indicato quindi con $\langle\ ,\ \rangle_{\mathcal{D}} : \mathcal{M}(\Omega) \times C(\Omega) \longrightarrow \mathbb{R}$ l'operatore definito da

$$\langle \tau, u\rangle_{\mathcal{D}} = \begin{cases} \langle \tau, u\rangle & \text{se } u \in \mathcal{D}(\Omega) \\ +\infty & \text{altrimenti,} \end{cases}$$

è chiaro che $\langle -f, \cdot\rangle \leq \langle -f, \cdot\rangle_{\mathcal{D}}$; visto allora che $\overline{F} \leq F$ e $G = F + \langle -f, \cdot\rangle_{\mathcal{D}}$, si ha che

$$\overline{F} + \langle -f, \cdot\rangle \leq F + \langle -f, \cdot\rangle_{\mathcal{D}} = G;$$

essendo inoltre $\overline{F} + \langle -f, \cdot\rangle$ debolmente s.c.i. in quanto somma di due funzioni che lo sono, si ha

$$\overline{G} \geq \overline{F} - \langle f, \cdot\rangle;$$

d'altra parte è anche vero che $\langle f, \cdot\rangle \leq \langle f, \cdot\rangle_{\mathcal{D}}$, e dunque

$$\overline{G} \leq G \Longrightarrow \overline{G} + \langle f, \cdot\rangle \leq G + \langle f, \cdot\rangle_{\mathcal{D}} = F;$$

essendo anche $\overline{G} + \langle f, \cdot\rangle$ debolmente s.c.i., da questo si deduce $\overline{G} + \langle f, \cdot\rangle \leq \overline{F}$; raccogliendo le due disuguaglianze trovate si ottiene quanto asserito.
v) La tesi segue a questo punto da ii), i) e iv).

■

Possiamo a questo punto mostrare il risultato fondamentale di questa sezione, il

Teorema 3.6 *Il problema* $(\star)$ *ammette una soluzione. Inoltre* $(\star) = -I(f)^2/2$.

Dim : Grazie al lemma 3.2 ed alla proposizione 3.4, esistono $u \in \mathrm{Lip}_1(\Omega)$ e $\nu \in \mathcal{M}(\Omega, \mathbb{R}^n)$ con $-\mathrm{div}(\nu) = f$ tali che

$$I(f) = \langle f, u\rangle = \int d|\nu| : \tag{3.5}$$

definiamo allora $\mu = |\nu|$, e scriviamo $\nu = \theta\mu$, con $\theta : \Omega \longrightarrow \mathbb{R}^n$ e $|\theta| = 1$ $\mu-$q.o.; dal momento che $-\mathrm{div}(\nu) = f$, usando la formula di integrazione per parti (2.9) si ottiene

$$\int d\mu = \int d|\nu| = \langle f, u\rangle = \langle -\mathrm{div}(\nu), u\rangle = \langle -\mathrm{div}(\theta\mu), u\rangle = \int D_\mu u \cdot \theta \; d\mu;$$

il fatto che si possa usare la (2.9) in questo contesto segue da una trattazione delle derivate tangenziali leggermente diversa da quella originale, che è discussa in [13] e che abbiamo già usato nella sezione 2.3. Visto poi che $u \in \mathrm{Lip}_1(\Omega)$, dalla definizione segue facilmente che $|D_\mu u| \leq 1$ $\mu-$q.o.; essendo $|\theta| = 1$ $\mu-$q.o., allora, dall'ultima equazione si ottiene che

$$|D_\mu u| = 1 \; \mu - \text{q.o.} \qquad\qquad \theta = D_\mu u.$$

A questo punto definiamo

$$\bar\mu := \frac{1}{I(f)}\mu \qquad\qquad \bar u := uI(f); \tag{3.6}$$

essendo $\mu = |\nu|$, μ è una misura non negativa, e dunque lo è $\bar\mu$; inoltre per costruzione $\int d\mu = I(f)$, per cui $\bar\mu \in \mathcal{P}(\Omega)$; infine si ha

$$f = -\mathrm{div}(\nu) = -\mathrm{div}(\theta\mu) = -\mathrm{div}(\mu D_\mu u) = -\mathrm{div}(\bar\mu D_{\bar\mu}\bar u);$$

grazie a tutto questo e alla proposizione 3.1 si ha

$$\mathcal{E}(\bar\mu) \geq -\frac{1}{2}\int |D_\mu \bar u|^2 \; d\mu = \frac{1}{2}\langle \mathrm{div}(\bar\mu D_{\bar\mu}\bar u), \bar u\rangle = -\frac{1}{2}\langle f, \bar u\rangle = -\frac{1}{2}I(f)^2.$$

Grazie al lemma 3.3, allora, $\bar\mu$ risulta una misura ottimale ed inoltre $(\star) = -I(f)^2/2$.

■

E' importante osservare come la u e la μ incontrate in questa dimostrazione sono una soluzione di

$$-\mathrm{div}(\mu D_\mu u) = f \qquad |D_\mu u| = 1 \; \mu - \text{q.o.} \qquad u \in \mathrm{Lip}_1(\Omega), \;\; \mu \in \mathcal{M}^+(\Omega); \tag{3.7}$$

questo problema, discusso in [6], è strettamente legato a quello di Monge–Kantorovich, e si può incontrare in situazioni diverse: ad esempio nella costruzione di un trasporto ottimale di Evans–Gangbo in [12], di cui abbiamo già parlato nella sezione 1.6, si trova una soluzione di questo problema essendo u il potenziale e μ una misura assolutamente continua che contiene l'informazione necessaria per costruire il trasporto.

Ma la u e la μ incontrate nella dimostrazione erano ottenute semplicemente prendendo una coppia (u, ν) di estremali per $I(f)$ nel senso dato da (3.5) e definendo $\mu = |\nu|$; prendendo dunque una generica coppia (u, ν) ottimale per $I(f)$ nel senso dato da (3.5), $(u, \mu) = (u, |\nu|)$ è una soluzione di (3.7); un fatto sorprendente è che vale anche il viceversa, grazie alla

Proposizione 3.7 *Se* (u, μ) *è una soluzione del problema (3.7),* $(u, \mu D_\mu u)$ *sono ottimali per* $I(f)$ *nel senso di (3.5) e* $\mu / \int d\mu$ *risolve* $(\star)$.

Dim : Definiamo intanto $\sigma = D_\mu u$ e $\nu = \mu D_\mu u$: visto che (u, μ) risolvono (3.7), si ha che $|\sigma| = 1$, dove l'uguaglianza è intesa quasi ovunque secondo μ, $\nu \in \mathcal{M}(\Omega, \mathbb{R}^n)$ con $-\mathrm{div}(\nu) = f$, e infine $u \in \mathrm{Lip}_1(\Omega)$. Grazie alla definizione (3.3) e alla proposizione 3.4, allora, si ha

$$I(f) \leq \int d|\nu| = \int d\mu = \int |\sigma|^2 \, d\mu = \int D_\mu u \cdot D_\mu u \, d\mu = \langle -\mathrm{div}(\mu D_\mu u), u \rangle = \langle f, u \rangle \leq I(f),$$

per cui le due disuguaglianze sono in realtà uguaglianze e quindi si ottiene la prima parte della tesi. A questo punto definiamo, in analogia con (3.6),

$$\tilde{\mu} := \frac{\mu}{\int d\mu} \qquad\qquad \tilde{u} := u \int d\mu,$$

in modo che $\tilde{\mu} \in \mathcal{P}(\Omega)$; sia inoltre $\tau = D_\mu \tilde{u}$: dalle ipotesi discende che $-\mathrm{div}(\tau \tilde{\mu}) = f$, per cui, grazie alla proposizione 3.1, si ha

$$\begin{aligned} (\star) \quad &\geq \quad \mathcal{E}(\tilde{\mu}) \geq -\frac{1}{2} \int |\tau|^2 \, d\tilde{\mu} = -\frac{1}{2} \int \left| \sigma \int d\mu \right|^2 \, d\left(\frac{\mu}{\int d\mu} \right) = -\frac{1}{2} \int d\mu \int |\sigma|^2 \, d\mu \\ &= \quad -\frac{1}{2} I(f)^2 = (\star), \end{aligned}$$

e dunque $\tilde{\mu}$ risolve il problema $(\star)$.

■

Dopo quest'ultima caratterizzazione notiamo che, grazie al teorema 3.6, il calcolo numerico di $(\star)$ si può ricondurre a quello di $I(f)$; nelle condizioni generali del problema di Monge–Kantorovich, allora, definiamo

$$\mathrm{L}_{bb} := \sup \left\{ \int u \; d(f^+ - f^-) : \; u \in \mathrm{Lip}_1(\Omega) \right\} = I(f); \tag{3.8}$$

questa formulazione del problema verrà poi confrontata con le altre nella sezione 3.3.

Recentemente Bouchitté e Buttazzo in [4] hanno esteso lo studio di questo problema anche al caso vettoriale: per fare questo si considera $j : \mathbb{R}^{n^2} \longrightarrow \mathbb{R}$ che verifichi le proprietà

a) j è convessa;

b) j è positivamente $p-$omogenea con $p > 1$;

c) j dipende solo dalla parte simmetrica del suo argomento;

si definisce poi, data una misura μ ed $u \in \mathcal{D}(\Omega, \mathbb{R}^n)$, l'energia associata come

$$E(\mu, u) := \int j(Du) \; d\mu - \langle f, u \rangle :$$

chiaramente stavolta il dato iniziale sarà una $f \in \mathcal{M}(\Omega, \mathbb{R}^n)$. Un'ulteriore generalizzazione si ottiene considerando il minimo dell'energia non fra tutte le possibili funzioni $u \in \mathcal{D}(\Omega, \mathbb{R}^n)$, bensì solo fra quelle che appartengono ad un cono convesso $\mathcal{U} \subseteq \mathcal{D}(\Omega, \mathbb{R}^n)$, e si definisce quindi

$$\mathcal{E}(\mu) := \inf_{u \in \mathcal{U}} E(\mu, u);$$

anche stavolta l'obiettivo è quello di massimizzare $\mathcal{E}(\mu)$ fra le misure di probabilità. Così come il problema, nel caso scalare, si prestava ad un'interpretazione fisica nel modello della termostatica, nel caso vettoriale si può dare un'interpretazione nell'ambito di un modello dell'elasticità.

Anche questo problema più generale può essere affrontato, grazie ai risultati di dualità che abbiamo già usato nel caso scalare, trovando scritture equivalenti che permettono di trovare un massimo.

3.2 Formulazione di Brenier

Un modello per il problema di trasporto ben diverso da quello che abbiamo visto finora è dovuto a Brenier, che affronta la questione in modo "dinamico": non si chiede, infatti, solo quale sia la funzione ottimale che manda la distribuzione iniziale in quella finale (in altre parole, *dove* portare la massa), ma anche, in un certo senso, quali siano le conseguenti distribuzioni intermedie (ovvero, *come* portarla). Questo punto di vista si può già riconoscere in [8], in cui viene utilizzato per lo studio di un modello della fluidodinamica, ma viene per la prima volta discusso in modo organico in [3]; in [9], poi, permette di costruire un sistema di equazioni alle derivate parziali che coincide con le equazioni di Maxwell. Mentre la trattazione in [3] si occupa del caso di costo quadratico, qui verrà considerato, come in tutto il resto del capitolo, il costo lineare. La formulazione che deriva da questa estensione del problema è abbastanza comoda da usare, e si confronta facilmente sia con quella di Kantorovich che con quella di Bouchitté, Buttazzo e Seppecher, risultando in qualche modo una "via di mezzo".

La formalizzazione di questa idea è la seguente: si considera una famiglia di misure scalari non negative μ_t e di misure vettoriali E_t con $t \in [0,1]$, in modo che valga la relazione di compatibilità

$$\frac{d}{dt}\mu_t + \nabla \cdot E_t = 0 \tag{3.9}$$

in senso distribuzionale; per ogni $\varphi \in \mathcal{D}(\Omega)$, in altre parole, chiamando

$$F(t) := \langle \mu_t, \varphi \rangle = \int_\Omega \varphi(x)\ d\mu_t(x) \qquad\qquad G(t) := \sum_{i=1}^n \int_\Omega \frac{\partial \varphi}{\partial x_i}(x)\ d(E_t)_i$$

si vuole che $F' = G$ in senso debole; come di consueto, d'ora in poi scriveremo per comodità $G(t)$ anche come $\int_\Omega \nabla\varphi \cdot dE_t$.

E' chiaro che se $\|E_t\| \in L^1([0,1])$ allora anche $G \in L^1([0,1])$, visto che si può facilmente stimare $\|G\|_{L^1} \leq \mathrm{Lip}(\varphi)\|E_t\|_{L^1}$; perciò, essendo $F' = G$, F risulta continua e dunque ben definita ad ogni istante t: da questo si può dedurre che anche le misure μ_t possono essere ben definite ad ogni istante, grazie al

Lemma 3.8 *Se $\|E_t\| \in L^1([0,1])$, dato $\bar t \in [0,1]$ esiste una misura α tale che $\mu_t \xrightarrow[t\to\bar t]{*} \alpha$; tale misura, inoltre, è evidentemente unica.*

<u>Dim</u> : Come già notato l'ipotesi assicura che, assegnata φ, è continua la F risultante, e quindi è ben definito $\beta(\varphi) = F(\bar t)$; per definizione, allora, $\langle \mu_t, \varphi \rangle \xrightarrow[t\to\bar t]{} \beta(\varphi)$, e quindi se esiste una misura α come nell'enunciato deve essere $\beta(\varphi) = \langle \alpha, \varphi \rangle$ per ogni $\varphi \in \mathcal{D}(\Omega)$; ci siamo ricondotti allora a mostrare che β sia una misura, e visto che chiaramente β è lineare è sufficiente trovare una costante K tale che $\beta(\varphi) \leq K\|\varphi\|_{C(\Omega)}$. Ma in effetti la

continuità della funzione F relativa ad una qualunque φ assicura la limitatezza puntuale delle misure μ_t e quindi, grazie al teorema di Banach–Steinhaus, la limitatezza globale; ne segue che $K = \sup_{t\in[0,1]} \|\mu_t\| < +\infty$ è una costante adeguata, e dunque si ha la tesi.

■

Nella dimostrazione di questo lemma abbiamo usato il fatto che $\|\mu_t\|$ è limitato; chiamando

$$\Omega_0 := \bigcup_{t\in[0,1]} supp(\mu_t) \cup supp(E_t),$$

se $\Omega_0 \subset\subset \Omega$ dalla relazione di compatibilità segue in effetti molto di più:

Lemma 3.9 *Se $\{(\mu_t, E_t)\}$ verifica la (3.9) e $\Omega_0 \subset\subset \Omega$, allora $\|\mu_t\|$ è costante.*

Dim : Si consideri una funzione $\varphi(x) \in \mathcal{D}(\Omega)$ tale che $\varphi \equiv 1$ su Ω_0: è chiaro che $\langle \nabla \cdot E_t, \varphi\rangle = 0 \; \forall\, t$, e quindi $\mu_t(\varphi) = \|\mu_t\|$ è costante.

■

Visto che la misura μ_t è ben definita ad ogni istante purché sia $\|E_t\| \in L^1([0,1])$, si può considerare l'insieme delle configurazioni ammissibili come

$$\Pi(f^+, f^-) := \Big\{ \big\{(\mu_t, E_t) \in \mathcal{M}^+(\Omega) \times \mathcal{M}(\Omega, \mathbb{R}^n),\ t \in [0,1]\big\} : \\ \Omega_0 \subset\subset \Omega,\ \dot{\mu}_t + \nabla \cdot E_t = 0,\ \|E_t\| \in L^1([0,1]),\ \mu_0 = f^+,\ \mu_1 = f^- \Big\};$$

in sostanza una configurazione ammissibile tiene conto di tutte le misure intermedie μ_t con cui si passa da f^+ ad f^- e della "velocità" $\|E_t\|$ con cui avviene questo passaggio. E' opportuno notare che la condizione di compattezza $\Omega_0 \subset\subset \Omega$ non è un'ipotesi di comodo assunta solo per poter applicare il lemma 3.9, bensì una richiesta naturale visto il problema che stiamo considerando: senza questa ipotesi, infatti, sarebbero permesse configurazioni che facciano "sparire" massa portandola oltre il bordo di Ω e la facciano "ricomparire" da fuori, magari lontano da dove era sparita; chiaramente non si può permettere che questo avvenga, a meno di snaturare completamente l'idea che stiamo formalizzando…

In analogia con il problema di Kantorovich, chiameremo d'ora in poi *costo totale* di una configurazione la

$$C\big(\{(\mu_t, E_t)\}\big) := \int_{t=0}^{1} \|E_t\|\, dt$$

e il problema posto da Brenier si può esprimere semplicemente come

$$\mathrm{L}_{br} := \inf \big\{ C\big(\{(\mu_t, E_t)\}\big) :\ \{(\mu_t, E_t)\} \in \Pi(f^+, f^-)\big\}\,; \tag{3.10}$$

le buone proprietà di questo problema, come ad esempio l'esistenza di una configurazione ottimale, verranno studiate nella prossima sezione, insieme al confronto con gli altri problemi incontrati.

3.3 Equivalenza tra le diverse formulazioni

In questa sezione mostreremo l'equivalenza tra la formulazione di Monge–Kantorovich del problema e le due formulazioni che abbiamo introdotto nelle sezioni 3.1 e 3.2; ovviamente, in analogia con L_{bb} ed L_{br}, definiremo

$$\mathrm{L}_{mk} := \inf \left\{ \int |x-y| \, d\gamma(x,y) : \ \gamma \in \Gamma(f^+, f^-) \right\}, \tag{3.11}$$

che è la formulazione di Kantorovich già discussa nel corso del primo capitolo. L'obiettivo principale di questa sezione sarà la dimostrazione che $\mathrm{L}_{mk} = \mathrm{L}_{bb} = \mathrm{L}_{br}$; visto poi che per ciascuno di questi problemi l'estremo è effettivamente raggiunto, cercheremo anche dei collegamenti tra le soluzioni, notando in quali casi da una soluzione per un problema si riesce facilmente a costruire una soluzione per un altro.

Iniziamo subito il confronto tra L_{mk}, L_{br} e L_{bb} con il primo

Lemma 3.10 *Per ogni $\gamma \in \Gamma(f^+, f^-)$ esiste una configurazione $\{(\mu_t, E_t)\} \in \Pi(f^+, f^-)$ tale che $C\big(\{(\mu_t, E_t)\}\big) \le C(\gamma)$.*

Dim : La costruzione che faremo è molto intuitiva: un elemento di $\Pi(f^+, f^-)$ è in pratica un "tragitto" che porta la distribuzione di massa iniziale $\mu_0 = f^+$ in quella finale $\mu_1 = f^-$; avendo a disposizione un trasporto generalizzato γ, un modo molto semplice per spostare nel tempo la massa è farlo seguendo i raggi di trasporto a velocità costante, e tali spostamenti di massa inducono delle corrispondenti distribuzioni. E' facile formalizzare questa idea: definita, per $t \in [0,1]$, $f_t : \Omega \times \Omega \longrightarrow \Omega$ come $f_t(x,y) = x + t(y-x)$ –cosa che ha senso perché Ω è convesso– le distribuzioni μ_t che abbiamo appena descritte sono date semplicemente da

$$\mu_t := f_{t\#}\gamma;$$

meno immediato è capire quali debbano essere le corrispondenti misure vettoriali E_t: in effetti negli esempi è spesso molto più facile intuire il comportamento delle misure μ_t piuttosto che quello delle E_t, anche se poi il costo totale della configurazione dipende in maniera esplicita da queste ultime. Definiamo comunque

$$E_t := f_{t\#}(y-x)\gamma,$$

notando che $(y-x)\gamma$ è una misura vettoriale su $\Omega \times \Omega$ e quindi le E_t così definite sono effettivamente misure vettoriali su Ω.

Innanzitutto si deve controllare che $\{(\mu_t, E_t)\} \in \Pi(f^+, f^-)$; ma presa una funzione test $\varphi \in \mathcal{D}(\Omega)$ si ha

$$\begin{aligned}
\frac{d}{dt}\langle \mu_t, \varphi\rangle &= \frac{d}{dt}\int_\Omega \varphi(x)\ d\mu_t(x) = \frac{d}{dt}\int_\Omega \varphi(x)\ df_{t\#}\gamma(x) = \frac{d}{dt}\int_{\Omega\times\Omega} \varphi(f_t(x,y))\ d\gamma(x,y) \\
&= \int_{\Omega\times\Omega} \frac{d}{dt}\ \varphi(x+t(y-x))\ d\gamma(x,y) = \int_{\Omega\times\Omega} \nabla\varphi(f_t(x,y))\cdot(y-x)\ d\gamma \\
&= \int_{\Omega\times\Omega} \nabla\varphi(f_t(x,y))\cdot d(y-x)\gamma = \int_\Omega \nabla\varphi\cdot d\big(f_{t\#}(y-x)\gamma\big) \\
&= \int_\Omega \nabla\varphi\cdot dE_t = -\langle \nabla E_t, \varphi\rangle,
\end{aligned}$$

che assicura che vale l'equazione di compatibilità (3.9). Inoltre $f_0(x,y) = x$ e $f_1(x,y) = y$, per cui $\mu_0 = f^+$ e $\mu_1 = f^-$; grazie a tutto questo $\{(\mu_t, E_t)\} \in \Pi(f^+, f^-)$. Resta da stimare il costo totale di $\{(\mu_t, E_t)\}$; ma per definizione

$$\begin{aligned}
C\big(\{(\mu_t, E_t)\}\big) &= \int_{t=0}^1 \|E_t\|\ dt = \int_{t=0}^1 |E_t|(\Omega)\ dt = \int_{t=0}^1 |f_{t\#}(y-x)\gamma|(\Omega)\ dt \\
&\le \int_{t=0}^1 \big(f_{t\#}|y-x|\gamma\big)(\Omega)\ dt = \int_{t=0}^1 \big(|y-x|\gamma\big)\big({f_t}^{-1}(\Omega)\big)\ dt \\
&= \int_{t=0}^1 \big(|y-x|\gamma\big)(\Omega\times\Omega)\ dt = \int_{t=0}^1 \iint_{\Omega\times\Omega} |y-x|\ d\gamma\ dt \\
&= \iint_{\Omega\times\Omega} |y-x|\ d\gamma = C(\gamma),
\end{aligned}$$

per cui la configurazione costruita verifica tutte le richieste e dunque mostra la tesi.

■

Corollario 3.11 $\mathrm{L}_{br} \le \mathrm{L}_{mk}$.

Dim : E' una conseguenza immediata del lemma appena mostrato: basta prendere una misura γ ottimale per il problema di Kantorovich e applicare il lemma ottenendo una configurazione $\{(\mu_t, E_t)\}$; a questo punto

$$\mathrm{L}_{br} \le C\big(\{(\mu_t, E_t)\}\big) \le C(\gamma) = \mathrm{L}_{mk}.$$

■

Mostriamo ora che l'enunciato del lemma 3.10 resta valido anche cambiando il ruolo dei due problemi:

Lemma 3.12 *Per ogni $\{(\mu_t, E_t)\} \in \Pi(f^+, f^-)$ esiste una misura $\gamma \in \Gamma(f^+, f^-)$ tale che $C(\gamma) \le C\big(\{(\mu_t, E_t)\}\big)$.*

Dim : Sia $\{(\mu_t, E_t)\}$ una configurazione ammissibile.

i) Estendiamo intanto la configurazione in un intorno sinistro di 0 semplicemente definendo $E_t \equiv 0$ e lasciando costante $\mu_t = \mu_0 = f^+$, e allo stesso modo in un intorno destro di 1 con $E_t \equiv 0$ e $\mu_t = \mu_1 = f^-$. Si operi quindi una convoluzione (sia nelle variabili spaziali che in quella temporale), passando in questo modo a $\{(\mu_t^\epsilon, E_t^\epsilon)\}$: allora E_t^ϵ e μ_t^ϵ sono C^∞ e, grazie alle buone proprietà della convoluzione, anche $\{(\mu_t^\epsilon, E_t^\epsilon)\}$ verifica l'equazione di compatibilità (3.9), ed in senso classico. Pur di utilizzare dei mollificatori gaussiani, inoltre, $\mu_t^\epsilon(x) > 0$ per ogni ϵ, t ed x visto che, in accordo con il lemma 3.9, $\|\mu_t\| = 1 \; \forall\, t$. E' allora possibile introdurre il campo di vettori $g_t^\epsilon : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ definito da

$$g_t^\epsilon(x) := -\frac{E_t^\epsilon(x)}{\mu_t^\epsilon(x)}$$

e considerare il flusso $\psi^\epsilon : [0,1] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ individuato da

$$\begin{cases} \psi^\epsilon(0,x) = x \\ \dfrac{d}{dt}\psi^\epsilon(t,x) = g_t^\epsilon(\psi^\epsilon(t,x)). \end{cases}$$

ii) A questo punto definiamo una misura γ_t^ϵ su $\mathbb{R}^n \times \mathbb{R}^n$ tramite

$$\langle \gamma_t^\epsilon, \varphi \rangle := \int_{\mathbb{R}^n} \varphi(x, \psi^\epsilon(t,x)) \; d\mu_0^\epsilon \, ; \tag{3.12}$$

il costo totale di questa misura risulta

$$C(\gamma_t^\epsilon) = \iint_{\mathbb{R}^n \times \mathbb{R}^n} |x-y| \; d\gamma_t^\epsilon = \langle \gamma_t^\epsilon, |x-y| \rangle = \int_{\mathbb{R}^n} |\psi^\epsilon(t,x) - x| \; d\mu_0^\epsilon,$$

da cui si deduce immediatamente che $C(\gamma_0^\epsilon) = 0$; la derivata temporale di questo costo totale, inoltre, è

$$\begin{aligned} \frac{d}{dt} C(\gamma_t^\epsilon) &= \frac{d}{dt} \int_{\mathbb{R}^n} |\psi^\epsilon(t,x) - x| \; d\mu_0^\epsilon \leq \int_{\mathbb{R}^n} |g^\epsilon(\psi^\epsilon(t,x))| \; d\mu_0^\epsilon \\ &= \int_{\mathbb{R}^n} |g^\epsilon(x)| \; d\big(\psi^\epsilon(t,\cdot)_\# \mu_0^\epsilon\big). \end{aligned} \tag{3.13}$$

iii) $\psi^\epsilon(t,\cdot)_\# \mu_0^\epsilon = \mu_t^\epsilon$.
Chiamando $\nu_1(t,x) = \mu_t^\epsilon(x)$ e $\nu_2(t,x) = \psi^\epsilon(t,\cdot)_\# \mu_0^\epsilon$, si ha

$$\nu_1(0,x) = \mu_0^\epsilon(x) \qquad\qquad \nu_2(0,x) = \psi^\epsilon(0,\cdot)_\# \mu_0^\epsilon(x) = \mu_0^\epsilon(x);$$

inoltre

$$\dot{\nu}_1(t,x) = \dot{\mu}_t^\epsilon(x) = -\nabla \cdot E_t^\epsilon = \operatorname{div}\big(g_t^\epsilon(x)\mu_t^\epsilon(x)\big) = \operatorname{div}\big(g_t^\epsilon(x)\nu_1(t,x)\big)$$

e analogamente

$$\begin{aligned} \langle \dot{\nu}_2(t,x), \varphi \rangle &= \frac{d}{dt} \int_{\mathbb{R}^n} \varphi(\psi^\epsilon(t,x)) \; d\mu_0^\epsilon = \int_{\mathbb{R}^n} \nabla\varphi(\psi^\epsilon(t,x)) \cdot \dot{\psi}^\epsilon(t,x) \; d\mu_0^\epsilon \\ &= \int_{\mathbb{R}^n} \nabla\varphi \cdot g_t^\epsilon \; d\big(\psi^\epsilon(t,\cdot)_\# \mu_0^\epsilon\big) = \langle \operatorname{div}\big(g_t^\epsilon(x)\nu_2(t,x)\big), \varphi \rangle; \end{aligned}$$

ma allora ν_1 e ν_2 verificano la stessa equazione differenziale ed hanno lo stesso dato iniziale; essendo in ambito C^∞ si possono applicare i risultati standard di unicità per trovare $\nu_1 = \nu_2$, e dunque quanto asserito; questa particolare equazione differenziale è studiata in [19], dove si mostra l'unicità anche in ipotesi molto più deboli.
iv) Grazie a quanto mostrato in iii), si può proseguire il conto lasciato in sospeso in (3.13):

$$\frac{d}{dt} C(\gamma_t^\epsilon) \leq \cdots \leq \int_{\mathbb{R}^n} |g^\epsilon(x)| \; d\mu_t^\epsilon = \int_{\mathbb{R}^n} |E_t^\epsilon| \; dx.$$

Visto poi che $C(\gamma_0^\epsilon) = 0$, si ha

$$\begin{aligned} C(\gamma_1^\epsilon) &= \int_{t=0}^1 \frac{dC(\gamma_t^\epsilon)}{dt} \; dt \leq \int_{t=0}^1 \int_{\mathbb{R}^n} |E_t^\epsilon| \; dx \; dt \leq \int_{t=0}^1 \int_{\mathbb{R}^n} |E_t| \; dx \; dt \\ &= \int_{t=0}^1 \|E_t\| \; dt = C\big(\{(\mu_t, E_t)\}\big); \end{aligned} \tag{3.14}$$

inoltre, grazie alla (3.12), si ottiene che $\|\gamma_1^\epsilon\| \leq \|\mu_0^\epsilon\| \underset{\epsilon \to 0}{\longrightarrow} \|\mu_0\|$: le restrizioni $\tilde{\gamma}_1^\epsilon$ delle misure γ_1^ϵ ad Ω, allora, sono limitate per $\epsilon \to 0$, e si può dunque definire $\gamma \in \mathcal{M}(\Omega)$ tale che $\tilde{\gamma}_1^{\epsilon_k} \xrightharpoonup[k \to +\infty]{*} \gamma$ per un'opportuna successione $\epsilon_k \to 0$.
v) $\gamma \in \Gamma(f^+, f^-)$ e $C(\gamma) \leq C\big(\{(\mu_t, E_t)\}\big)$.
La seconda affermazione segue dalla (3.14) notando che $C(\gamma) = \lim C(\tilde{\gamma}_1^{\epsilon_k}) \leq \underline{\lim}\, C(\gamma_1^{\epsilon_k})$, per cui basta mostrare che $\gamma_1 = f^+$ e $\gamma_2 = f^-$; ma questo, ricordando che $\tilde{\gamma}_1^{\epsilon_k} \xrightharpoonup[k \to +\infty]{*} \gamma$ e facendo uso di quanto mostrato nel punto iii), si ottiene immediatamente dalla (3.12); si noti che questo assicura anche $\gamma \in \mathcal{P}(\Omega)$ mentre prima, avendo ristretto tutte le misure ad Ω, si sarebbe potuto affermare solo che $\gamma \in \mathcal{M}^+(\Omega)$ e che $\|\gamma\| \leq 1$. La misura γ, allora, verifica tutte le richieste dell'enunciato.

■

Corollario 3.13 $\mathrm{L}_{mk} \leq \mathrm{L}_{br}$.
Dim : E' una conseguenza immediata del lemma appena mostrato.

■

Si noti che nella dimostrazione del lemma 3.12 è possibile regolarizzare le μ_t e le E_t tramite convoluzione solo grazie all'ipotesi $\Omega_0 \subset\subset \Omega$ fatta definendo le configurazioni ammissibili; senza questa ipotesi non solo non si potrebbe fare la costruzione descritta, ma sarebbe anche falso l'enunciato del lemma e del conseguente corollario visto che, come discusso nella scorsa sezione, non assumendo la compattezza dei supporti diventa ammissibile far scomparire e riapparire massa attraverso il bordo di Ω. Se allora, ad esempio, $f^+ = \delta_x$ e $f^- = \delta_y$ e i punti x ed y sono lontani tra loro ma vicini al bordo, si può immaginare come sia più conveniente far uscire la massa dal bordo vicino al punto x

per farla ricomparire vicino ad y; in particolare, si verifica facilmente che risulterebbe

$$\mathrm{L}_{br} = d(x, \partial\Omega) + d(y, \partial\Omega) < d(x, y) = \mathrm{L}_{mk}.$$

Occupiamoci adesso del confronto con L_{br}, molto più semplice; intanto si ha il

Lemma 3.14 *Data una soluzione (u, μ) del problema (3.7), esiste una configurazione $\{(\mu_t, E_t)\}$ tale che $C\big(\{(\mu_t, E_t)\}\big) = \int d\mu = \mathrm{L}_{bb}$.*

Dim : Mostrando, con la proposizione 3.7 e l'osservazione precedente, che il problema (3.7) è equivalente al calcolo di $I(f)$, abbiamo anche notato che se (u, μ) risolve (3.7) allora $\int d\mu = I(f) = \mathrm{L}_{bb}$; per avere la tesi, allora, consideriamo una coppia ottimale (u, μ) e definiamo

$$\mu_t := f^+ + t(f^- - f^+) \qquad\qquad E_t := -\mu D_\mu u;$$

si noti l'estrema semplicità delle misure definite: le μ_t variano linearmente da f^+ ad f^-, mentre E_t è in realtà la stessa ad ogni istante. Il fatto che si tratti di una configurazione ammissibile è abbastanza facile: la relazione di compatibilità è verificata in quanto

$$\dot{\mu}_t = f^- - f^+ = -f = \mathrm{div}(\mu D_\mu u) = -\mathrm{div}(E_t),$$

e il fatto che $\mu_0 = f^+$ e $\mu_1 = f^-$ è immediato dalla definizione. Il costo totale di questa configurazione, inoltre, è

$$C\big(\{(\mu_t, E_t)\}\big) = \int_{t=0}^{1} \|E_t\| \; dt = \int_{t=0}^{1} \|\mu D_\mu u\| \; dt = \|\mu\| = \int d\mu$$

ricordando che $|D_\mu u| = 1 \; \mu -$ q.o.; la configurazione esibita, dunque, assicura la tesi.

■

Corollario 3.15 $\mathrm{L}_{br} \leq \mathrm{L}_{bb}$.

Dim : Anche questo corollario è banale, pur di ricordare che effettivamente, come abbiamo discusso nella prima sezione, il problema (3.7) ammette soluzioni.

■

Mostriamo quindi l'altra disuguaglianza:

Lemma 3.16 $\mathrm{L}_{bb} \leq \mathrm{L}_{br}$.

Dim : Sia $\{(\mu_t, E_t)\} \in \Pi(f^+, f^-)$ e $u \in \mathrm{Lip}_1(\Omega)$; allora

$$\begin{aligned} \left| \int_\Omega u \; d(f^+ - f^-) \right| &= \left| \int_\Omega u \; d\mu_0 - \int u \; d\mu_1 \right| = \left| \int_{t=0}^{1} \int_\Omega u \; d\dot{\mu}_t \; dt \right| \\ &= \left| \int_{t=0}^{1} \int_\Omega u \; d\big(\nabla \cdot E_t\big) \; dt \right| = \left| \int_{t=0}^{1} \int_\Omega \nabla u \cdot \; dE_t \; dt \right| \leq \int_{t=0}^{1} \|E_t\| \; dt; \end{aligned}$$

data la genericità di u e $\{(\mu_t, E_t)\}$, allora, ne segue la tesi.

$\blacksquare$

Mentre nelle dimostrazioni dei lemmi 3.10, 3.12 e 3.14 abbiamo effettivamente esibito una misura o una configurazione ottimali, nel lemma 3.16 non c'è stato bisogno di costruire alcuna funzione; per quanto, allora, questi quattro lemmi siano sufficienti per stabilire l'uguaglianza tra gli estremali dei tre problemi considerati, vediamone un altro, il cui scopo è di mostrare come sia possibile costruire una funzione ottimale per L_{bb}:

Lemma 3.17 *Data una misura γ ottimale nel senso di Kantorovich, ogni potenziale $u \in \mathrm{Lip}_1(\Omega)$ per cui valga la (1.10) è ottimale per L_{bb}.*

Dim : Dal momento che stiamo lavorando con il costo lineare, se per u vale la (1.10) si ha che $u \in \mathrm{Lip}_1(\Omega)$ e $u(x) - u(y) = |x - y|$ per ogni $(x, y) \in supp(\gamma)$; che u sia ottimale per L_{bb} discende dal fatto che, essendo $\gamma_1 = f^+$ e $\gamma_2 = f^-$,

$$\begin{aligned}
\int_\Omega u \ d(f^+ - f^-) &= \int_\Omega u(x) \ df^+(x) - \int_\Omega u(y) \ df^-(y) \\
&= \int_{\Omega\times\Omega} u(x) \ d\gamma(x, y) - \int_{\Omega\times\Omega} u(y) \ d\gamma(x, y) \\
&= \int_{\Omega\times\Omega} u(x) - u(y) \ d\gamma(x, y) = \int_{\Omega\times\Omega} |x - y| \ d\gamma(x, y) = C(\gamma) = \mathrm{L}_{bb}.
\end{aligned}$$

Si noti anche che, in accordo con il teorema 1.18, esistono potenziali per cui valga la (1.10).

$\blacksquare$

Riassumendo quanto mostrato finora, si ha il

Teorema 3.18 $\mathrm{L}_{mk} = \mathrm{L}_{bb} = \mathrm{L}_{br}$*; ciascuno dei tre estremali, inoltre, è effettivamente raggiunto.*

Dim : L'uguaglianza segue dai lemmi 3.10, 3.12, 3.14 e 3.16; per quanto riguarda l'esistenza di soluzioni, per il problema di Kantorovich è stato mostrato nel teorema 1.3, per quello di Bouchitté, Buttazzo e Seppecher nel lemma 3.2; a partire da soluzioni degli altri problemi, infine, si possono costruire soluzioni di quello di Brenier usando il lemma 3.10 o il lemma 3.14.

$\blacksquare$

Discutiamo infine la questione del collegamento tra le soluzioni dei diversi problemi raccogliendo quanto emerso finora: per il problema di Brenier si possono facilmente costruire soluzioni esplicite usando soluzioni degli altri due problemi; si noti che questo è stato fatto in due modi radicalmente diversi: partendo da una soluzione del problema di Bouchittè, Buttazzo e Seppecher, infatti, le misure μ_t percorrono semplicemente il "segmento" che collega f^+ ad f^- mentre, partendo da una misura ottimale per

Kantorovich, sono ottenute "trascinando" f^+ lungo i raggi individuati da γ, in un modo forse più vicino a quanto suggerirebbe l'intuizione. Il supporto di μ_t, allora, è l'unione dei supporti di f^+ e di f^- nel primo caso, mentre si sposta lungo Ω partendo dal primo per raggiungere il secondo nell'altro caso.

Per gli altri problemi la situazione è meno buona: la costruzione che abbiamo presentato di una soluzione per Kantorovich a partire da una per Brenier è molto più laboriosa delle altre e non porta alla costruzione esplicita di una misura ottimale, ma solo alla dimostrazione della sua esistenza come limite debole; in generale, inoltre, la misura ottimale che si trova con questo metodo non è univocamente determinata dalla configurazione ottimale di partenza.

Anche la soluzione per Bouchitté, Buttazzo e Seppecher trovata a partire da una misura ottimale per Kantorovich non è univocamente determinata, visto che in generale vi possono essere molti potenziali per γ trovati grazie al teorema 1.18; questi potenziali, tra l'altro, non hanno una chiara dipendenza da γ dal momento che, grazie al teorema 1.19, esiste un particolare potenziale che soddisfa ogni misura γ ottimale. Non abbiamo, infine, esibito alcun metodo diretto per passare da una soluzione per Brenier ad una per Bouchitté, Buttazzo e Seppecher, né da una per Bouchitté, Buttazzo e Seppecher ad una per Kantorovich.

3.4 La misura ottimale μ

In questa sezione discuteremo brevemente il ruolo e le proprietà di una particolare misura che risolve il problema (3.7). Grazie ai forti legami che abbiamo mostrato nella scorsa sezione esservi tra le possibili formulazioni del problema, non ci si deve stupire che questa misura giochi un ruolo particolare in contesti diversi; ad esempio, mentre la dimostrazione che faremo del fatto che sia ottimale utilizzerà il rapporto tra il problema di Kantorovich e quello di Brenier, un'altra dimostrazione, completamente differente, nell'ambito del problema studiato da Bouchitté e Buttazzo si trova in [4]. Il prossimo capitolo, poi, sarà dedicato allo studio delle proprietà di regolarità di questa misura.

Vediamo intanto la definizione: fissato un trasporto ottimale γ, si definirà μ tramite la

$$\langle \mu, \varphi \rangle := \int_{\Omega\times\Omega} \left(\int_{\Omega} \varphi(z)\ d\mathcal{H}^1_{xy}(z) \right) d\gamma(x,y), \tag{3.15}$$

dove con $\mathcal{H}^1_{xy}$ indichiamo la misura di Hausdorff monodimensionale sul segmento xy; indicando con $l(xy \cap A) = \mathcal{H}^1_{xy}(A)$ la lunghezza dell'intersezione tra un segmento xy ed un insieme $A \subseteq \Omega$, allora, la misura di A è data da

$$\mu(A) = \int_{\Omega\times\Omega} l(xy \cap A) d\gamma(x,y). \tag{3.16}$$

Un pregio di questa misura è che il suo comportamento si presta molto bene ad essere visualizzato tenendo conto dell'interpretazione delle misure $\gamma \in \Gamma(f^+, f^-)$ come trasporti generalizzati: immaginando che lungo i raggi di trasporto venga spostata la massa da f^+ ad f^-, $\mu(A)$ rappresenta il lavoro totale compiuto all'interno della regione A. Per mostrare il ruolo di questa misura, risulterà comodo considerare la configurazione $\{(\mu_t, E_t)\}$ ottenuta da γ come mostrato nel lemma 3.10; innanzitutto notiamo che vale il

Lemma 3.19

$$\mu = \int_{t=0}^{1} |E_t|\ dt.$$

Dim : Definiamo $\tilde{\mu} = \int_{t=0}^{1} |E_t|\ dt$, in modo da dover mostrare $\tilde{\mu} = \mu$; grazie all'ottimalità di γ sappiamo che $\mathrm{L}_{mk} = C(\gamma) = \tilde{\mu}(\Omega)$; dato inoltre un qualunque insieme $B \subseteq \Omega$ si ha, ricordando che per definizione $E_t = f_{t\#}(y-x)\gamma$,

$$\begin{aligned}\tilde{\mu}(B) &= \int_{t=0}^{1} |E_t|(B)\ dt = \int_{t=0}^{1} |f_{t\#}(y-x)\gamma|(B)\ dt \le \int_{t=0}^{1} f_{t\#}\big(|y-x|\gamma\big)(B)\ dt \\ &= \int_{t=0}^{1} \int_{f_t{}^{-1}(B)} |y-x|\ d\gamma\ dt = \int_{\Omega\times\Omega} \varphi(x,y)|y-x|\ d\gamma,\end{aligned} \tag{3.17}$$

indicando con $\varphi(x,y)$ la misura secondo dt del sottoinsieme di $[0,1]$ formato dagli istanti t per i quali $f_t(x,y) \in B$; ma allora

$$\varphi(x,y) = \mathcal{H}^1\left\{t : (1-t)x + ty \in B\right\} = \frac{l(xy \cap B)}{|x-y|};$$

sostituendo questa caratterizzazione nella (3.17) si ottiene

$$\tilde{\mu}(B) \leq \int_{\Omega\times\Omega} l(xy \cap B)\ d\gamma = \mu(B);$$

d'altra parte, come abbiamo già fatto notare prima, si ha che $\tilde{\mu}(\Omega) = C(\gamma)$; ma è anche

$$\mu(\Omega) = \int_{\Omega\times\Omega} l(xy \cap \Omega)\ d\gamma = \int_{\Omega\times\Omega} |x-y|\ d\gamma = C(\gamma),$$

e quindi per ogni $B \subseteq \Omega$ usando la disuguaglianza già trovata si ha

$$\tilde{\mu}(\Omega) = \tilde{\mu}(B) + \tilde{\mu}({}^cB) \leq \mu(B) + \mu({}^cB) = \mu(\Omega) = \tilde{\mu}(\Omega),$$

da cui segue anche l'altra disuguaglianza e dunque la tesi.

■

Resta ora da far vedere che questa misura risolva il problema (3.7), proprietà che viene assicurata dal

Teorema 3.20 *Esiste una funzione u tale che (u,μ) risolve (3.7); equivalentemente esiste una misura vettoriale ν ottimale nel senso della proposizione 3.4 tale che $\mu = |\nu|$.*

<u>Dim</u> : Il fatto che i due enunciati di questo teorema siano equivalenti segue dalle caratterizzazioni mostrate nella prima sezione di questo capitolo; mostriamo allora solo il secondo, definendo

$$\nu := -\int_{t=0}^{1} E_t\ dt;$$

è chiaro intanto che

$$|\nu| \leq \int_{t=0}^{1} |E_t|\ dt = \mu.$$

Perché ν possa essere ottimale nel senso della proposizione 3.4 serve anzitutto che $-\mathrm{div}(\nu) = f$; in effetti

$$-\mathrm{div}(\nu) = \mathrm{div}\left(\int_{t=0}^{1} E_t\ dt\right) = \int_{t=0}^{1} \mathrm{div}(E_t)\ dt = -\int_{t=0}^{1} \dot{\mu}_t = \mu_0 - \mu_1 = f^+ - f^- = f;$$

ricordando poi che $\mu(\Omega) = C(\gamma) = \mathrm{L}_{bb}$, come notato nella dimostrazione dell'ultimo lemma, si ha

$$\mathrm{L}_{bb} = I(f) \leq \int d|\nu| \leq \int d\mu = \mathrm{L}_{bb},$$

da cui segue allo stesso tempo che ν è ottimale nel senso della proposizione 3.4 e che $|\nu| = \mu$, per cui la tesi è completa.

■

Capitolo 4

Regolarità della misura ottimale μ

In questo capitolo studieremo la regolarità della misura ottimale μ che viene definita tramite (3.15) e (3.16) una volta che sia fissata una misura $\gamma \in \mathcal{P}(\Omega)$ ottimale per il problema di Kantorovich; ci occuperemo in particolare della dimensione, intesa nel senso dato dalle nozioni di Hausdorff o minimale definite nel capitolo 2, e della sommabilità rispetto alla misura di Lebesgue. Per quanto riguarda la dimensione, troveremo che quella di μ è almeno il massimo fra quelle di f^+ ed f^-; per quanto riguarda la sommabilità, invece, mostreremo che se f^+ ed f^- sono in L^p ed i loro supporti sono separati allora anche $\mu \in L^p$, e che se anche una sola delle due misure appartiene ad un L^q, allora μ è contenuta in un opportuno L^p dove p, che dipende da q e dalla dimensione n dello spazio, è strettamente maggiore di 1 se lo è anche q: in particolare se una fra f^+ ed f^- è assolutamente continua lo è anche μ.

La prima sezione è introduttiva e commenta brevemente i risultati di tutto il capitolo; la seconda affronta la questione della dimensione; la terza contiene semplicemente un risultato tecnico fondamentale per lo studio della sommabilità; la quarta, infine, si occupa della sommabilità di μ.

E' opportuno notare che tutte le proprietà di μ che troveremo in questo capitolo dipendono dalle analoghe proprietà di f^+ ed f^- e non da quelle di γ, che non è un dato del problema. Se la situazione è sufficientemente regolare, in effetti, si trova che μ è univocamente determinata da f^+ ed f^- e non dipende dalla scelta della misura γ ottimale: anche se non è vero in generale che sia così, è tuttavia quello che succede in tutti gli esempi considerati nella sezione 4.1.

4.1 Introduzione

Prima di studiare in dettaglio le proprietà di μ, conviene farsi un'idea di quello che succede tramite alcune considerazioni; innanzitutto notiamo che, grazie alla definizione, il supporto di μ è contenuto nell'insieme di trasporto, o almeno nella sua chiusura, e quindi si può facilmente avere una prima, rozza stima della sua estensione facendo l'inviluppo convesso dei supporti di f^+ ed f^-, che certamente contiene ogni raggio di trasporto. In configurazioni caratterizzate da elevata simmetria, inoltre, si può capire immediatamente quale debba essere la misura γ ottimale, grazie alla proposizione 1.6 che assicura che due diversi raggi non si possono incontrare. Se ad esempio f^+ ed f^- sono le misure di Hausdorff $m-$dimensionali concentrate su due $m-$cubi, ottenuti l'uno dall'altro tramite la traslazione di un vettore parallelo ad uno degli spigoli dei cubi, è chiaro che tutti i raggi di trasporto devono essere paralleli a quel vettore, e quindi direttamente dalla definizione si trova facilmente che il supporto di μ è esattamente l'inviluppo convesso dei due cubi e che μ ha dimensione m. Se invece, essendo $m < n$, i due cubi si corrispondono tramite una traslazione ortogonale a tutte le facce dei cubi, i raggi di trasporto sono ancora tutti paralleli al vettore di traslazione, ma chiaramente la dimensione di μ è stavolta $m+1$. Questi due semplici esempi possono dare un'idea intuitiva del modo in cui μ venga costruita a partire da f^+ ed f^-, e rendono credibile l'idea che la dimensione di μ sia almeno quella di f^+ o di f^-: nella prossima sezione mostreremo che in effetti è almeno il massimo; il primo esempio, inoltre, assicura anche che questa stima è ottimale, visto che la dimensione di μ può non superare questo massimo. Una questione della quale non ci occuperemo, ma che potrebbe sicuramente risultare interessante, è la ricerca di una stima dall'alto della dimensione di μ.

Anche per quanto riguarda la questione della sommabilità alcuni semplici esempi possono risultare illuminanti: innanzitutto si può considerare il caso in cui f^+ sia distribuita uniformemente in una corona sferica e f^- sia una δ di Dirac nel centro della sfera; essendo il supporto di f^- un solo punto, i raggi dovranno essere tutti e soli i segmenti che partono da quel punto e raggiungono la corona, dal che è immediato dedurre che μ ammette come densità una funzione inversamente proporzionale alla distanza dal centro, e quindi $\mu \in L^p$ per ogni $p < n$; in questo caso $f^+ \in L^\infty$, mentre $\mathcal{H}-\dim(f^-) = 0$. Questo esempio può lasciare immaginare che, perché μ sia assolutamente continua, non è necessario che lo siano sia f^+ che f^-: nel teorema 4.9 mostreremo che in effetti è sufficiente che lo sia solo una delle due. Come ultimo esempio, per semplicità fatto in $\mathbb{R}^2$, consideriamo ora due rettangoli uguali, disgiunti, con i lati paralleli ad un sistema di assi cartesiani, ottenuti l'uno dall'altro tramite una traslazione lungo l'asse x. Se (x, y) appartiene al primo rettangolo sia ora $f^+(x,y) = \varphi(y)$, e allo stesso modo sia $f^-(x,y) = \varphi(y)$ se (x,y) appartiene al secondo rettangolo; fuori dai rispettivi rettangoli f^+ ed f^- siano invece nulle. Anche in questo caso è ovvio che i raggi di trasporto siano tutti segmenti orizzontali che congiungono i due rettangoli e che μ abbia supporto nel loro inviluppo; inoltre si trova subito che per ogni y di Lebesgue per φ si ha $\mu(x,y) = \varphi(y)$ nella zona compresa tra i due rettangoli, mentre al loro interno decresce fino a 0 linearmente con l'ascissa. Ne segue che la sommabilità di f^+, f^- e μ è la stessa e coincide con quella di φ. Questo ultimo esempio fornisce una valida idea generale per la questione della sommabilità di μ nel caso in cui f^+ e f^- siano in L^p: si può sperare che anche μ sia in L^p, e in effetti nel teorema 4.13 mostreremo che è così con l'ipotesi che i supporti siano separati. Inoltre il fatto che nell'esempio le sommabilità di f^+, f^- e μ coincidano assicura che anche questa stima è ottimale.

L'ultimo teorema di questo capitolo, come già accennato, asserisce che se anche una sola delle due misure sta in uno spazio L^q con $q > 1$, allora μ sta in uno spazio L^p con un certo $p > 1$; grazie all'ultimo esempio visto è chiaro che non si può estendere questo enunciato con $p > 1$ anche nel caso $q = 1$. L'utilità di questo risultato sarà principalmente dovuta a due fattori: il primo è che non sarà necessaria l'ipotesi aggiuntiva della separazione dei supporti delle misure date, già discussa nel primo capitolo; il secondo è che grazie a questo enunciato diventa possibile usare caratteristiche di compattezza dovute alla riflessività di L^p con $p > 1$. Allora è chiaro che, tutto sommato, per la maggior parte delle applicazioni non è fondamentale sapere quale sia l'esponente p, purché esista; oltretutto non si può sperare che sia molto maggiore di 1, come intuibile vista la mancanza di ipotesi sull'altra misura e come alcuni semplici esempi possono mostrare. Per questo motivo eviteremo di preoccuparci di quale sia l'esponente ottimale, trovandone quindi forse uno più basso.

4.2 Dimensione della μ ottimale

Il risultato principale di questa sezione sarà il teorema 4.6, che assicura che la dimensione di μ è almeno il massimo fra quelle di f^+ ed f^-.

Visto che la definizione di μ trovata in (3.15) non coinvolge esplicitamente i dati del problema non è immediato riuscire a stabilire delle proprietà per μ direttamente a partire da quelle di f^+ ed f^-, tuttavia scopriremo che queste sono estremamente legate; dal momento, poi, che per ottenere le stime di questo capitolo faremo molti ragionamenti di tipo geometrico, sarà comodo adottare notazioni sui nomi dei punti leggermente diverse rispetto a quanto fatto finora, più adatte a questo tipo di studio: quando fisseremo un punto per calcolarne la dimensione puntuale, sarà per comodità l'origine $\mathrm{O} \in \mathbb{R}^n$, mentre nei ragionamenti di tipo globale verrà chiamato x; la generica coppia di punti appartenente al supporto di γ sarà inoltre (A, B) piuttosto che (x, y). Una prima conseguenza immediata della definizione è che, senza alcuna ipotesi su f^+ ed f^-, si ha la

Proposizione 4.1 $\mathcal{W}-\dim(\mu) \geq 1$. *In particolare,* $\mu(B) = 0$ *per ogni insieme* B $\mathcal{H}^1$*−trascurabile.*

Dim : Preso $x \in \Omega$, si ha che evidentemente $l(\mathrm{AB} \cap B(x, \rho)) \leq 2\rho$ per qualsiasi coppia $(\mathrm{A}, \mathrm{B}) \in \Omega \times \Omega$. Ma allora, grazie a (3.16), è

$$\mu(B(x,\rho)) = \int_{\Omega\times\Omega} l(\mathrm{AB} \cap B(x,\rho))d\gamma \leq \int_{\Omega\times\Omega} 2\rho \; d\gamma \leq 2\rho, \tag{4.1}$$

e quindi (2.2) assicura $\Lambda_\mu(x) \geq 1$, per cui da (2.3) si ottiene che $\mathcal{W}-\dim \geq 1$. Grazie a (4.1), in particolare, direttamente dalla definizione di misura di Hausdorff si evince che $\mu(B) \leq \mathcal{H}^1(B)$, per cui $\mu \ll \mathcal{H}^1$, il che mostra anche la seconda parte dell'enunciato.

■

A parte questa semplicissima proprietà, per ottenere qualcosa avremo bisogno di una strategia: lo schema generale sarà di concentrarci su un punto fissato, O, per ottenere stime su $\Lambda_\mu(\mathrm{O})$; da questo risaliremo a informazioni riguardo alla dimensione di μ, usando quindi prevalentemente –ma non solo– la dimensione minimale. Per dare una stima su $l(\mathrm{AB} \cap B(\mathrm{O}, \rho))$ cercheremo di limitare l'insieme delle coppie $(\mathrm{A}, \mathrm{B}) \in supp(\gamma)$ per cui l'intersezione è non vuota: avremo dunque una maggiorazione immaginando la lunghezza pari a 2ρ su questo insieme e 0 fuori. Per fare questa limitazione costruiremo tale insieme modellandolo su una data coppia $(\mathrm{A}, \mathrm{B}) \in supp(\gamma)$ tale che AB contiene O; in questa sezione considereremo il caso in cui O sia interno al segmento AB, quindi distinto da A e B. Supporremo poi che un'altra coppia $(\widetilde{\mathrm{A}}, \widetilde{\mathrm{B}})$ tale che $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ intersechi $B(\mathrm{O}, \rho)$ appartenga al supporto di γ senza rispettare la limitazione enunciata, e cercheremo di trovare un

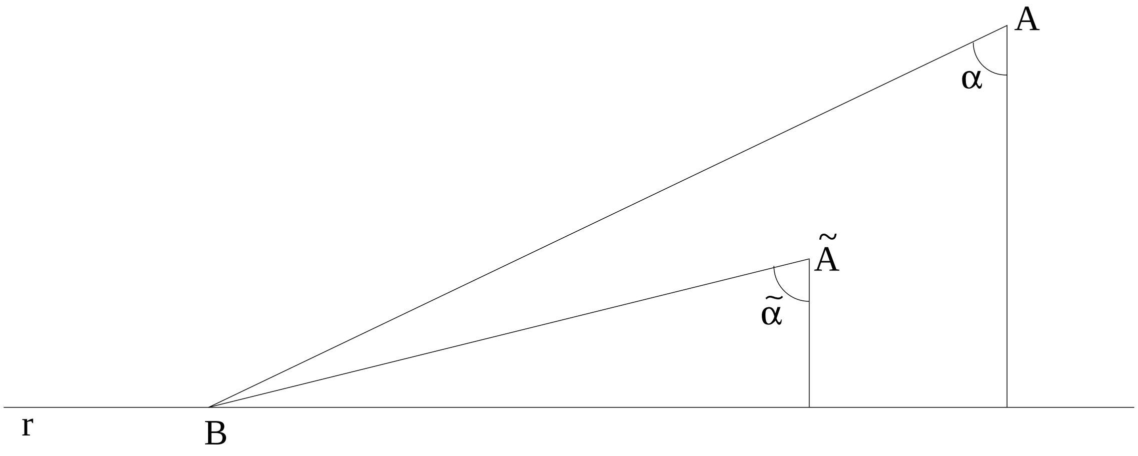

Figura 4.1: Spostamenti ammissibili

assurdo grazie alla (1.3), che in questo capitolo sarà più utile nella forma

$$\overline{\mathrm{AB}} + \overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}} \leq \overline{\mathrm{A}\widetilde{\mathrm{B}}} + \overline{\widetilde{\mathrm{A}}\mathrm{B}}. \tag{4.2}$$

Per arrivare ad un assurdo troveremo molto comodo ricondurci a posizioni relative particolari tra i punti $\mathrm{A}, \mathrm{B}, \widetilde{\mathrm{A}}$ e $\widetilde{\mathrm{B}}$, eventualmente spostandoli in modo che continui a valere la (4.2); è molto importante essere chiari su questo punto: visto che avendo i punti $\mathrm{A}, \mathrm{B}, \widetilde{\mathrm{A}}$ e $\widetilde{\mathrm{B}}$ ci interessa solo trovare un assurdo basandoci sulla (4.2), è lecito cambiare le loro posizioni in modo che

$$\overline{\mathrm{AB}} + \overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}} - \overline{\mathrm{A}\widetilde{\mathrm{B}}} - \overline{\widetilde{\mathrm{A}}\mathrm{B}}$$

diminuisca, e quindi valga a maggior ragione la disuguaglianza; è chiaro però che a questo punto potremo fare soltanto ragionamenti geometrici sulle distanze fra i punti, senza fare riferimento al problema, visto che niente imporrà alle nuove coppie di punti di continuare a stare nel supporto di γ. Diremo dunque *leciti* quegli spostamenti per cui diminuisce $\overline{\mathrm{AB}} + \overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}} - \overline{\mathrm{A}\widetilde{\mathrm{B}}} - \overline{\widetilde{\mathrm{A}}\mathrm{B}}$, e faremo solo quelli. E' utile notare il

Lemma 4.2 *i) Avvicinare* A *(risp.* B,$\widetilde{\mathrm{A}}$,$\widetilde{\mathrm{B}}$*) a* B *(risp.* A,$\widetilde{\mathrm{B}}$,$\widetilde{\mathrm{A}}$*) lungo la congiungente è lecito, così come allontanarlo da* $\widetilde{\mathrm{B}}$ *(risp.* $\widetilde{\mathrm{A}}$,B,A*).*
ii) Con riferimento alla figura 4.1, uno spostamento di B *verso destra lungo la retta r è lecito se $\alpha > \tilde{\alpha}$, verso sinistra se $\alpha < \tilde{\alpha}$, fino a che tali disuguaglianze fra gli angoli restano verificate. Inoltre uno spostamento qualunque è lecito se $\alpha = \tilde{\alpha}$ e* $\overline{\mathrm{AB}} > \overline{\widetilde{\mathrm{A}}\mathrm{B}}$.

Si tratta di proprietà ovvie, che discendono immediatamente da una stima al primo ordine delle variazioni delle lunghezze dei segmenti, ma visto che le utilizzeremo spesso viene comodo scriverle una volta per tutte, per evitare di confonderci. Possiamo occuparci subito della prima delle proprietà di restrizione sugli insiemi:

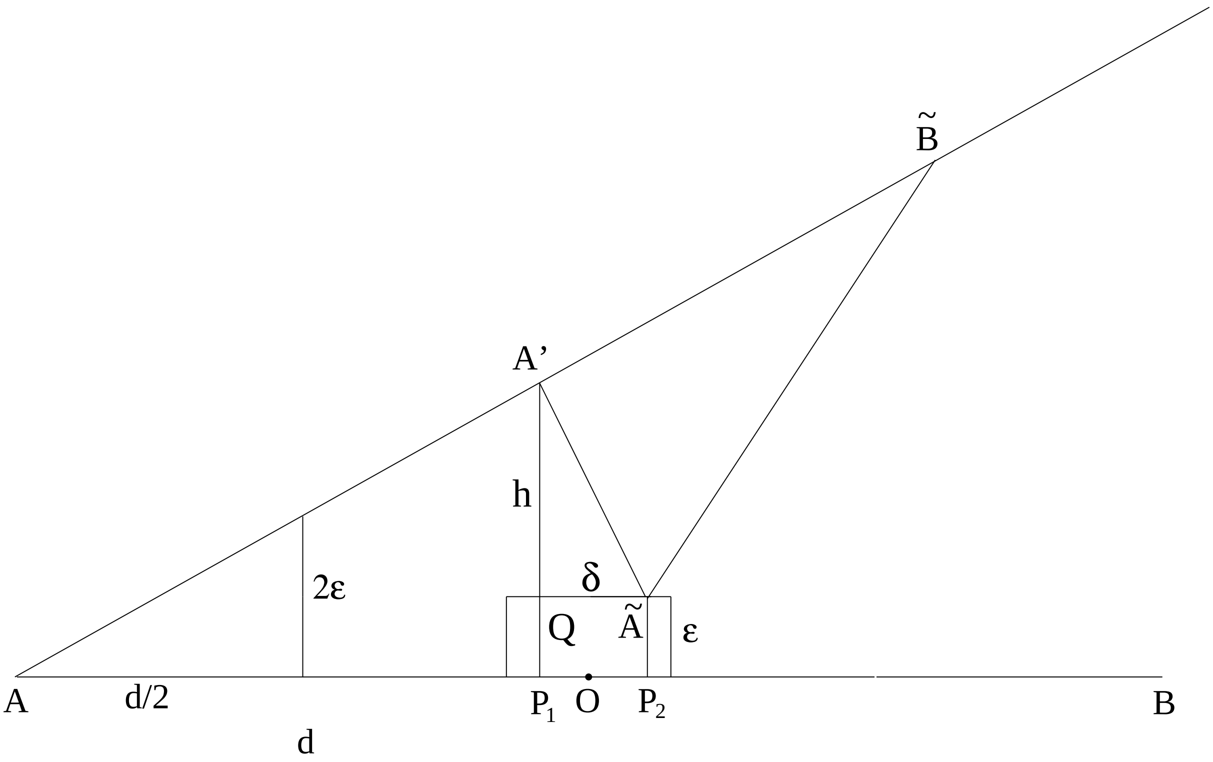

Figura 4.2: Costruzione geometrica per il lemma 4.3

Lemma 4.3 *Dato un punto* O *interno al raggio* AB*, se ϵ è sufficientemente piccolo un raggio* $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ *può intersecare $B(\mathrm{O},\epsilon)$ solo se* $\widetilde{\mathrm{B}}$ *è contenuto in un opportuno cono avente superficie di base minore di $C\epsilon^{n-1}$, dove C non dipende da ϵ.*

Dim : Con riferimento alla figura 4.2, che per simmetria illustra solamente metà della regione, possiamo intanto supporre che O sia il punto medio del segmento AB, eventualmente avvicinando il punto fra i due che è più lontano, cosa lecita grazie al lemma 4.2, e quindi chiamiamo $d = \overline{\mathrm{AO}} = \overline{\mathrm{OB}}$; consideriamo innanzitutto il caso bidimensionale, la generalizzazione seguirà facilmente dopo. Piuttosto che lavorare con una palla risulta più comodo un quadrato, come vedremo, per cui al posto di $B(\mathrm{O},\epsilon)$ pensiamo ad un quadrato centrato in O di semilato ϵ avente un lato ortogonale ad AB: visto che questo è più grande di $B(\mathrm{O},\epsilon)$, se un raggio $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ incontra la palla a maggior ragione incontra il quadrato, per cui questa semplificazione è lecita. Il cono sia quello che a distanza $d/2$ dal vertice, posto in A, abbia apertura 2ϵ, come in figura; è chiaro che lavorando in dimensione due il cono è in effetti un triangolo. Visto che Ω ha diametro limitato, si può maggiorare la lunghezza della base (che nella figura sarebbe un segmento verticale posto all'estremità destra del dominio) con $K\epsilon$, essendo K inversamente proporzionale a d, per cui le dimensioni di questo cono sono compatibili con l'enunciato.

Prendiamo ora per assurdo un raggio $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ che incontri il quadrato pur essendo $\widetilde{\mathrm{B}}$ esterno al cono, e poi avviciniamo tra loro $\widetilde{\mathrm{A}}$ e $\widetilde{\mathrm{B}}$ fino a che $\widetilde{\mathrm{B}}$ non incontra la superficie esterna del cono e $\widetilde{\mathrm{A}}$ non esce dal quadrato; in effetti, visto che il segmento $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ deve incontrare il quadrato, avvicinando $\widetilde{\mathrm{A}}$ si deve prima entrare (a meno che $\widetilde{\mathrm{A}}$ non sia già nell'interno) e poi uscire. Consistentemente con la figura consideriamo come positiva l'ordinata di $\widetilde{\mathrm{B}}$, e cerhiamo di vedere che

— *E' lecito ricondursi al caso in cui* $\widetilde{\mathrm{A}}$ *sia sul lato superiore del quadrato.*

Per fare questo possiamo intanto supporre che anche $\widetilde{\mathrm{A}}$ abbia ordinata positiva; se non è così, infatti, possiamo ribaltarlo rispetto al segmento AB, cosa lecita perché $\overline{\widetilde{\mathrm{A}}\mathrm{B}}$ resta uguale mentre $\overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}}$ diminuisce; si può poi notare che, purché ϵ sia sufficientemente piccolo, l'ordinata di $\widetilde{\mathrm{B}}$ deve essere almeno ϵ, altrimenti $\overline{\mathrm{A}\widetilde{\mathrm{B}}} < \overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}}$ e $\overline{\widetilde{\mathrm{A}}\mathrm{B}} < \overline{\mathrm{AB}}$, contraddicendo la (4.2). E' allora lecito spostare $\widetilde{\mathrm{A}}$ verso l'alto finché non arriva sul lato superiore del quadrato, dal momento che in questo modo $\overline{\widetilde{\mathrm{A}}\mathrm{B}}$ aumenta mentre $\overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}}$ diminuisce.

Si chiami ora A' la proiezione ortogonale di $\widetilde{\mathrm{A}}$ sul lato AB, e poi si faccia in modo che $\widetilde{\mathrm{B}}$ sia oltre A' –ovvero ad ascissa maggiore– eventualmente allontanandolo da A lungo la loro congiungente, consistentemente con il lemma 4.2. Chiamate quindi, come in figura, P_1 e P_2 le proiezioni di A' ed $\widetilde{\mathrm{A}}$ su AB e Q il punto sul segmento $\mathrm{A}'\mathrm{P}_1$ di ordinata ϵ, sia $h = \overline{\mathrm{A}'\mathrm{Q}}$ e $\delta = \overline{\widetilde{\mathrm{A}}\mathrm{Q}}$. Finalmente notiamo che

$$\overline{\mathrm{AA}'} = \sqrt{\overline{\mathrm{AP}_1}^{\,2} + \overline{\mathrm{P}_1\mathrm{A}'}^{\,2}} < \overline{\mathrm{AP}_1} + \frac{\overline{\mathrm{P}_1\mathrm{A}'}^{\,2}}{2\overline{\mathrm{AP}_1}},$$

e allo stesso modo

$$\overline{\widetilde{\mathrm{A}}\mathrm{B}} = \sqrt{\overline{\widetilde{\mathrm{A}}\mathrm{P}_2}^{\,2} + \overline{\mathrm{P}_2\mathrm{B}}^{\,2}} < \overline{\mathrm{P}_2\mathrm{B}} + \frac{\overline{\widetilde{\mathrm{A}}\mathrm{P}_2}^{\,2}}{2\overline{\mathrm{P}_2\mathrm{B}}} = \overline{\mathrm{P}_2\mathrm{B}} + \frac{\epsilon^2}{2\overline{\mathrm{P}_2\mathrm{B}}}.$$

In accordo con la (4.2) dovrebbe poi essere

$$\overline{\widetilde{\mathrm{A}}\mathrm{B}} + \overline{\mathrm{A}\widetilde{\mathrm{B}}} \geq \overline{\mathrm{AB}} + \overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}} > \overline{\mathrm{AB}} + \overline{\mathrm{A}'\widetilde{\mathrm{B}}},$$

e quindi, essendo $\overline{\mathrm{AA}'} = \overline{\mathrm{A}\widetilde{\mathrm{B}}} - \overline{\mathrm{A}'\widetilde{\mathrm{B}}}$,

$$\overline{\mathrm{AB}} = 2d < \overline{\widetilde{\mathrm{A}}\mathrm{B}} + \overline{\mathrm{AA}'} < \overline{\mathrm{P}_2\mathrm{B}} + \frac{\epsilon^2}{2\overline{\mathrm{P}_2\mathrm{B}}} + \overline{\mathrm{AP}_1} + \frac{\overline{\mathrm{P}_1\mathrm{A}'}^{\,2}}{2\overline{\mathrm{AP}_1}} \Longrightarrow \delta < \frac{\epsilon^2}{2\overline{\mathrm{P}_2\mathrm{B}}} + \frac{\overline{\mathrm{P}_1\mathrm{A}'}^{\,2}}{2\overline{\mathrm{AP}_1}}; \quad (4.3)$$

usando ora le similitudini tra i triangoli, si ha che

$$\frac{\delta}{h} = \frac{4\epsilon}{d} \Longrightarrow \delta = \frac{4\epsilon h}{d}, \quad (4.4)$$

per cui la (4.3) diventa

$$\frac{4\epsilon h}{d} < \frac{\epsilon^2}{2\overline{\mathrm{P_2B}}} + \frac{\overline{\mathrm{P_1A'}}^{\,2}}{2\overline{\mathrm{AP_1}}}.$$

Ma $\overline{\mathrm{P_2B}}$ è compreso tra $d-\epsilon$ e $d+\epsilon$, e $\overline{\mathrm{AP_1}}$ tra $d-\epsilon-\delta$ e $d+\epsilon$; allora, visto che grazie a (4.4) si ha $\delta = \mathrm{o}(\epsilon)$, per ϵ sufficientemente piccolo si trova

$$\frac{4\epsilon h}{d} < 0.6\frac{\epsilon^2}{d} + 0.6\frac{\overline{\mathrm{P_1A'}}^{\,2}}{d}. \tag{4.5}$$

D'altra parte $\overline{\mathrm{P_1A'}} = h + \epsilon$ e quindi la (4.5), semplificando il fattore d che compare in ogni membro, porta a

$$4\epsilon h < 0.6\epsilon^2 + 0.6(h+\epsilon)^2 \Longrightarrow 0.6\left(\frac{h}{\epsilon}\right)^2 - 2.8\left(\frac{h}{\epsilon}\right) + 1.2 > 0.$$

Ma se ϵ è sufficientemente piccolo si riesce facilmente a stimare $\overline{\mathrm{A'P_1}}$ dal basso e dall'alto con 2ϵ e 5ϵ rispettivamente, e quindi risulta $\epsilon < h < 4\epsilon$; ma questo porta all'assurdo cercato perché l'equazione di secondo grado in h/ϵ appena trovata ha le radici approssimativamente pari a 4.19 e 0.48.

Risolto così il caso bidimensionale, vediamo come ricondurci ad esso in ogni situazione: intanto definiamo il cono alla stessa maniera, ovvero quello che a distanza $d/2$ dal vertice, posto in A, abbia apertura 2ϵ; anche stavolta, grazie alla limitatezza di Ω, si può maggiorare con $K\epsilon^{n-1}$, essendo K inversamente proporzionale a d^{n-1}, la superficie della base, e quindi anche in questo caso le dimensioni del cono sono compatibili con l'enunciato. Al posto della sfera, inoltre, non prendiamo il quadrato, come nel caso bidimensionale, bensì il cilindro che abbia raggio e semialtezza pari ad ϵ, centrato in O e con l'asse lungo il segmento AB. E' chiaro che se i quattro punti A, B, $\widetilde{\mathrm{A}}$ e $\widetilde{\mathrm{B}}$ sono su uno stesso piano si può applicare lo stesso ragionamento di prima ottenendo la tesi, visto che su quel piano la situazione sarebbe identica a quella che abbiamo già studiato; altrimenti si consideri la possibilità di ruotare $\widetilde{\mathrm{B}}$ intorno al segmento AB, muovendolo dunque su una sfera $(n-2)$-dimensionale: $\overline{\mathrm{A}\widetilde{\mathrm{B}}}$ è sempre lo stesso, per cui è lecito ricondursi al caso in cui $\overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}}$ è minimo, cosa che avviene in uno dei due punti di incontro tra la sfera e il piano per A, B e $\widetilde{\mathrm{A}}$: essendoci quindi ridotti alla situazione bidimensionale già studiata, segue la tesi anche nel caso generale.

■

Per riuscire da questo ad estrarre informazioni sulla dimensione di μ ci serve per il momento un'ipotesi supplementare, che poi saremo in grado di eliminare: consideriamo $\mathcal{W}-\dim(f^+) \geq \mathcal{W}-\dim(f^-)$, cosa possibile vista la simmetria del problema, e formuliamo l'ipotesi (Ip): detta $k = \mathcal{W}-\dim(f^+)$,

$$\exists\, C : \forall x \in \Omega \quad f^+(B(x,\rho)) \leq C\rho^k.$$

Si noti che questa è una richiesta piuttosto forte, anche se nella maggior parte degli esempi semplici è verificata: il fatto che la dimensione minimale di f^+ sia k assicura che in $(f^+)-$quasi nessun punto Λ_{f^+} è minore di k, cioè per ogni $\epsilon > 0$ e per quasi ogni punto x esiste una costante $C(x, \epsilon)$ per cui $f^+(B(x, \rho)) \leq C(x, \epsilon)\rho^{k-\epsilon}$: stiamo invece chiedendo che quella maggiorazione valga per ogni punto x, con $\epsilon = 0$ e per di più sempre con la stessa costante; come si può ben immaginare, non è difficile trovare esempi in cui (Ip) non valga... In compenso grazie a questa ipotesi siamo in grado di trovare un'importante conseguenza al lemma 4.3:

Corollario 4.4 *Se vale l'ipotesi (Ip), per ogni punto x che sia interno a qualche raggio si ha $\Lambda_\mu(x) \geq k$; più precisamente esiste una costante $C(x)$ tale che $\mu(B(x, \rho)) \leq C(x)\rho^k$.*

Dim : Preso un generico punto x interno a qualche raggio di trasporto, si deve calcolare $\mu(B(x, \rho))$; per definizione questo è

$$\mu(B(x, \rho)) = \int_{\Omega \times \Omega} l(\mathrm{AB} \cap B(x, \rho))d\gamma;$$

grazie al lemma 4.3 applicato scambiando i ruoli di f^+ ed f^-, cosa possibile per simmetria, AB non incontra $B(x, \rho)$, e quindi la funzione integranda è nulla, se A non sta in un opportuno cono la cui superficie di base è minore di $K\rho^{n-1}$, purché ρ sia sufficientemente piccolo; chiamato C_ρ questo cono, ricordando che l'intersezione tra un segmento e una sfera non può avere lunghezza maggiore del diametro e che $\pi_{1\#}\gamma = f^+$ si ottiene

$$\mu(B(x, \rho)) = \int_{C_\rho \times \Omega} l(\mathrm{AB} \cap B(x, \rho))d\gamma \leq 2\rho \int_{C_\rho} df^+ = 2\rho f^+(C_\rho).$$

Grazie all'ipotesi (Ip), tuttavia, è chiaro che se Q_ρ è un cubo di lato ρ si ha $f^+(Q_\rho) \leq C(\sqrt{n}/2)^k\rho^k$, visto che il cubo è contenuto in una sfera di raggio $\sqrt{n}\rho/2$. Inoltre il cono C_ρ può essere ricoperto da N/ρ cubi di lato ρ, avendo superficie di base stimata con ρ^{n-1}, e da ciò segue che

$$f^+(C_\rho) \leq \frac{N}{\rho}\, C \left(\frac{\sqrt{n}}{2}\right)^k \rho^k \Longrightarrow \mu(B(x, \rho)) \leq \widetilde{C}\rho^k;$$

per quanto visto finora la disuguaglianza è valida solo per ρ sufficientemente piccolo, ma visto che μ è una misura finita vale per ogni ρ, eventualmente aumentando la costante $\widetilde{C}$; questa maggiorazione, per definizione, assicura $\Lambda_\mu(x) \geq k$.

$\blacksquare$

Ma allora si può ottenere, come si voleva, che la dimensione di μ è maggiore di quelle di f^+ e di f^-, con la

Proposizione 4.5 *Se vale* (Ip)*, allora* $\mathcal{W}-\dim(\mu) \geq max(\mathcal{W}-\dim(f^+), \mathcal{W}-\dim(f^-))$.

Dim : Chiamando al solito $k = \mathcal{W}-\dim(f^+) \geq \mathcal{W}-\dim(f^-)$, per definizione di dimensione minimale si deve mostrare che la dimensione puntuale di μ è almeno k $\mu-$q.o., e quindi equivalentemente che è trascurabile l'insieme $\Psi = \{x : \Lambda_\mu(x) < k\}$; ma grazie al corollario 4.4 in Ψ non vi sono punti interni ad alcun raggio, e quindi i punti che ci sono sono esterni ad ogni raggio, od eventualmente estremi di qualche raggio. Da questo segue che dato un qualsiasi raggio AB la sua intersezione con Ψ consiste al più di due punti, gli estremi, e quindi $l(\mathrm{AB} \cap \Psi) = 0$; usando la formula (3.16) per calcolare $\mu(\Psi)$, allora, si integra in $d\gamma$ una funzione nulla sul supporto di γ, per cui in effetti Ψ è $\mu-$trascurabile, e questo porta alla tesi.

■

Nella dimostrazione di questa proposizione si è visto che $\mu-$quasi ogni punto è interno a qualche raggio: si noti bene che questo vuol dire che è trascurabile non l'insieme dei punti che sono estremi di un raggio, ma quello dei punti che sono *solo* estremi, cioè che non sono interni a nessun raggio; se un punto è estremo di un raggio ed è interno ad un altro, in pratica, per fare delle stime sulla sua dimensione puntuale risulta più comodo utilizzare il secondo piuttosto che il primo.

Mostrato allora il risultato che ci eravamo prefissi nel caso in cui valga (Ip), si riesce a fare a meno di questa ipotesi grazie alle proprietà delle dimensioni sulle misure mostrate nel capitolo 2, arrivando al

Teorema 4.6 $\mathcal{W}-\dim(\mu) \geq max(\mathcal{W}-\dim(f^+), \mathcal{W}-\dim(f^-))$.

Dim : Fissato ϵ, per ogni $m \in \mathbb{N}$ si definisca

$$\Omega_{m,\epsilon} = \left\{ x \in \Omega : \forall\ \rho > 0 \quad \frac{\mu(B(x,\rho))}{\rho^{k-\epsilon}} \leq m \right\};$$

visto che $k - \epsilon < \mathcal{W}-\dim(f^+)$, dalla definizione si ottiene che $\Omega_{m,\epsilon}$, al variare di m, è una successione crescente di insiemi che riempie Ω a meno di un insieme f^+-trascurabile, per cui $f^+ \llcorner \Omega_{m,\epsilon} \underset{m\to\infty}{\longrightarrow} f^+$, dove la convergenza è la forte in $\mathcal{M}(\Omega)$. Chiamato allora $\gamma_{m,\epsilon} = \gamma \llcorner \Omega_{m,\epsilon} \times \Omega$, visto che $\pi_{1\#}\gamma = f^+$ è vero anche che gli $\Omega_{m,\epsilon} \times \Omega$ riempiono $\Omega \times \Omega$ a meno di un insieme $\gamma-$trascurabile, e quindi si ha anche $\gamma_{m,\epsilon} \underset{m\to\infty}{\longrightarrow} \gamma$. Definiamo allora $\mu_{m,\epsilon}$ a partire dalla $\gamma_{m,\epsilon}$ applicando formalmente la (3.15); vista la formula (3.16) che permette di calcolare la misura $\mu_{m,\epsilon}$ di un insieme, il fatto che $\gamma_{m,\epsilon} \underset{m\to\infty}{\longrightarrow} \gamma$ assicura che per ogni misurabile A si ha

$$\mu_{m,\epsilon}(A) \longrightarrow \mu(A). \tag{4.6}$$

Ora il supporto di $\gamma_{m,\epsilon}$ è chiaramente contenuto in quello di γ e quindi è c-ciclicamente monotono, essendo c il costo lineare, visto che questa proprietà passa ai sottoinsiemi; ma allora continua a valere la (4.2) per le coppie di punti in $supp(\gamma_{m,\epsilon})$, e inoltre vale la (Ip) con esponente $k-\epsilon$ e costante m, per costruzione; ne segue che si possono applicare tutti i risultati di questa sezione, che si basavano solo sulla diseguaglianza (4.2), per ottenere che $\mathcal{W}-\dim(\mu_{m,\epsilon}) \geq k-\epsilon$. Sia allora A tale che $\mathcal{H}^{k-2\epsilon}(A)=0$: dal momento che $\mathcal{W}-\dim(\mu_{m,\epsilon}) \geq k-\epsilon$, è $\mu_{m,\epsilon} \ll \mathcal{H}^{k-2\epsilon}$, e dunque $\mu_{m,\epsilon}(A)=0$; grazie alla (4.6), allora, $\mu(A)=0$ e quindi per la genericità di A si è trovato $\mu \ll \mathcal{H}^{k-2\epsilon}$, da cui $\mathcal{W}-\dim(\mu) \geq k-2\epsilon$: per la genericità di ϵ, a questo punto, si ottiene la tesi.

$\blacksquare$

Il fatto che si possano applicare a $\mu_{m,\epsilon}$ i risultati già mostrati è discusso in modo più esauriente nell'osservazione 4.10 e nella discussione che la precede.

4.3 Stima uniforme nel caso di supporti separati

Come abbiamo visto, per fare stime sulla dimensione di μ è sufficiente avere un controllo sulla dimensione puntuale, ma non è indispensabile alcuna omogeneità: l'ipotesi (Ip), in effetti, era soltanto un passaggio intermedio comodo nella dimostrazione, ma in sostanza va benissimo se la disuguaglianza

$$\mu(B(x, \rho)) \leq C\rho^k$$

vale per $\mu-$q.o. x con una costante $C = C(x)$ che dipende da x –anzi, è sufficiente che sia verificata con ogni esponente minore di k. Per poter avere informazioni sulla sommabilità di μ, invece, c'è bisogno di stime uniformi: se ad esempio si sa che $\mu = g\mathcal{H}^n$ con $g \in L^1$ e che $g(x) \leq C$, per dedurne che $g \in L^\infty$ è necessario che C non dipenda da x!

Si vede subito che il lemma 4.3 non ci può essere di grande aiuto in questo nuovo contesto, visto che è in grado di stimare il volume del cono in cui costringe $\widetilde{\mathrm{B}}$ solo con $C\epsilon^{n-1}$ e C è inversamente proporzionale alla potenza $(n-1)-$esima della distanza tra il punto e l'estremo più vicino di un raggio che lo attraversa: quando i punti si avvicinano agli estremi dei raggi, allora, C diverge. Si avverte allora la necessità di un approccio diverso per capire come possano essere situati i raggi $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ che incontrano $B(x, \rho)$, che sia valido anche quando il punto x si avvicini all'estremo di un raggio; sarà conveniente allora affrontare direttamente il caso peggiore, cioè quando il punto scelto sia proprio origine di un raggio: questo ci servirà non per controllare gli effettivi estremi, che tanto sappiamo essere $\mu-$trascurabili, ma perché ci si può comunque ricondurre a questa situazione: se $\mathrm{O} \in \mathrm{AB}$ nulla ci impedisce, grazie al lemma 4.2, di avvicinare tra loro A e B fino a che uno dei due coincida con O. Pensiamo allora di avere un raggio AB e di voler controllare la misura delle sfere centrate in A con la stessa idea del lemma 4.3, cioè dando limitazioni alla posizione delle coppie $(\widetilde{\mathrm{A}}, \widetilde{\mathrm{B}})$ che siano nel supporto di γ e il cui raggio incontri una data sfera centrata in A. Intanto si può intuire subito che si dovrà cercare di capire dove sta $\widetilde{\mathrm{A}}$, cercare limitazioni per $\widetilde{\mathrm{B}}$ è uno sforzo inutile: se ad esempio $f^+ = \delta_{\mathrm{A}}$, vi saranno raggi $\mathrm{A}\widetilde{\mathrm{B}}$ (che certo intersecano $B(\mathrm{A}, \rho)$ qualunque sia ρ) per ogni punto $\widetilde{\mathrm{B}}$ che appartenga al supporto di f^-. Anche una limitazione per $\widetilde{\mathrm{A}}$, comunque, non può essere fatta senza altre ipotesi sulla situazione: nel caso in cui $f^- = \delta_{\mathrm{B}}$, ad esempio, vi sono raggi che uniscono B con ogni $\widetilde{\mathrm{A}} \in supp f^+$, e nel caso in cui $\mathrm{A} = \mathrm{B}$ tutti questi raggi incontrano $B(\mathrm{A}, \rho)$; si vede in effetti che l'ostacolo non è solo l'eventuale vicinanza tra A e B, ma più in generale la possibilità che vi siano, vicino ad A, dei punti del supporto di f^-, anche se distinti da B. Se questo non succede, però, ovvero se esiste un intorno di A che non incontra il supporto di f^-, si può risolvere la questione, come ci accingiamo a mostrare; è opportuno notare che l'ipotesi più semplice che permette di avere un tale intorno è quella che i supporti siano separati. Bisognerà porre attenzione agli spostamenti fatti grazie al lemma 4.2, perché potrebbero avvicinare i punti del supporto di f^+ al supporto di f^- da cui sono inizialmente lontani, effetto del quale si dovrà tener conto. Il risultato fondamentale di questa sezione è allora il

Teorema 4.7 *Esiste una costante K dipendente solo dalla geometria del problema tale che se $\mathrm{A} \in supp(f^+)$ è estremo di un raggio e $d(\mathrm{A}, supp(f^-)) = d > 0$, un raggio $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ può incontrare $B(\mathrm{A}, \epsilon)$ (purché ϵ sia sufficientemente piccolo) solo se $\widetilde{\mathrm{A}}$ è contenuto in un cilindro di superficie di base minore di $K\epsilon^{n-1}/d^{n-1}$.*

<u>Dim</u> : Sia B l'altro estremo del raggio passante per A, o di uno dei raggi passanti per A, nel caso in cui non sia unico; per prima cosa, come si vede nelle figure 4.3 e 4.4 che illustrano la situazione, si può avvicinare B ad A fino a distanza d, visto che inizialmente la distanza doveva essere almeno d per ipotesi. Il cilindro in cui mostreremo dover stare $\widetilde{\mathrm{A}}$ è quello che ha la retta passante per AB come asse e raggio $C\epsilon/d$, dove la costante C, che verrà trovata nel corso della dimostrazione, non dipenderà da A né da ϵ: si noti che questo cilindro ha superficie di base compatibile con l'enunciato.

Si scelga ora un raggio $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ che incontri $B(\mathrm{A}, \epsilon)$, e per prima cosa si avvicini $\widetilde{\mathrm{B}}$ ad $\widetilde{\mathrm{A}}$ finché non arrivi a distanza d da A; a questo punto si consideri per comodità una base ortonormale di $\mathbb{R}^n$ tale che A sia l'origine, che il primo vettore di base punti da A verso B e che il piano formato dai primi due vettori di base contenga anche $\widetilde{\mathrm{B}}$ in modo che abbia seconda coordinata minore o uguale a 0; al posto della palla $B(\mathrm{A}, \epsilon)$, poi, si prenda il cubo centrato nell'origine con le facce normali agli assi coordinati e di semilato ϵ, cosa lecita visto che contiene la sfera, esattamente come fatto nella dimostrazione del lemma 4.3: per ottenere la tesi ci basterà mostrare che è possibile limitare con $C\epsilon/d$ tutte le coordinate di $\widetilde{\mathrm{A}}$ esclusa la prima.

Per quanto riguarda le coordinate dalla terza in poi non ci sono problemi: indichiamo con $x_i(\mathrm{P})$ la i-esima coordinata di P, e sia Q un punto di intersezione fra il cubo e il segmento $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$; se $i \geq 3$, per costruzione si ha $x_i(\widetilde{\mathrm{B}}) = 0$ e $|x_i(\mathrm{Q})| \leq \epsilon$. Inoltre $|\widetilde{\mathrm{B}} - Q| \geq d - \epsilon\sqrt{n}$; per similitudine, allora,

$$\frac{x_i(\mathrm{Q})}{x_i(\widetilde{\mathrm{A}})} = \frac{|\widetilde{\mathrm{B}} - \mathrm{Q}|}{|\widetilde{\mathrm{B}} - \widetilde{\mathrm{A}}|} \Longrightarrow |x_i(\widetilde{\mathrm{A}})| = \frac{|\widetilde{\mathrm{B}} - \widetilde{\mathrm{A}}|}{|\widetilde{\mathrm{B}} - \mathrm{Q}|}|x_i(\mathrm{Q})| \leq \frac{R\epsilon}{d - \epsilon\sqrt{n}},$$

per cui, purché ϵ sia sufficientemente piccolo, si ha la limitazione cercata per tutte le coordinate dalla terza in poi; manca allora solo la seconda. Per controllarla conviene ricondursi alla situazione bidimensionale, cosa abbastanza semplice: visto che $\overline{\mathrm{AB}} = \overline{\mathrm{A}\widetilde{\mathrm{B}}} = d$, infatti, la (4.2) diventa semplicemente

$$\overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}} \leq \overline{\widetilde{\mathrm{A}}\mathrm{B}},$$

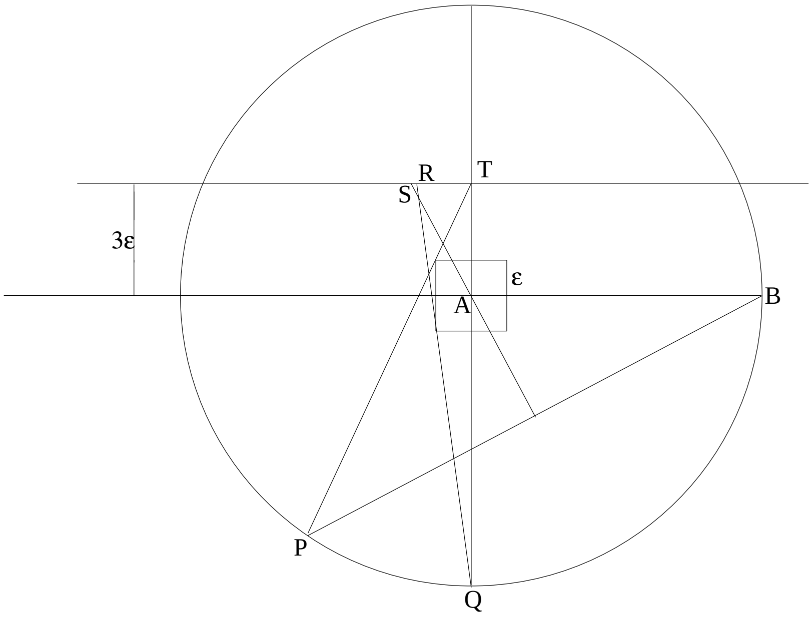

Figura 4.3: Caso i)

che resta vera se proiettiamo $\widetilde{\mathrm{A}}$ sul piano passante per A,B e $\widetilde{\mathrm{B}}$ non facendone variare, in questo modo, la seconda coordinata (se A,B e $\widetilde{\mathrm{B}}$ sono allineati la situazione è già bidimensionale, senza bisogno di spostare nulla). Inoltre, visto che il raggio $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ incontrava il cubo, dopo aver proiettato tutto sul piano resta vero che il segmento $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ –è opportuno chiamarlo segmento, perché non è più un raggio dopo aver spostato $\widetilde{\mathrm{A}}$– incontra il quadrato di centro A e semilato ϵ.

Essendoci così ricondotti al caso bidimensionale, si può notare subito che, immaginando negativa l'ordinata di $\widetilde{\mathrm{B}}$, non ci sono problemi per stimare opportunamente dal basso l'ordinata di $\widetilde{\mathrm{A}}$: con un ragionamento identico a quello fatto per le coordinate dalla terza in poi si trova che

$$x_2(\widetilde{\mathrm{A}}) \geq -\frac{R\epsilon}{d - \epsilon\sqrt{2}} \ ,$$

che come prima sistema la questione pur di avere ϵ sufficientemente piccolo; tutto quello che ancora manca alla tesi, allora, è una opportuna stima dall'alto di $x_2(\widetilde{\mathrm{A}})$, che sarà la parte più laboriosa della dimostrazione; consideriamo tre casi, a seconda delle possibili ascisse di $\widetilde{\mathrm{A}}$ e $\widetilde{\mathrm{B}}$:

i) $x_1(\widetilde{\mathrm{B}}) < 0$ e $x_1(\widetilde{\mathrm{A}}) \leq 0$.

Riferendoci alla figura 4.3, mostreremo che in questo caso $x_2(\widetilde{\mathrm{A}}) \leq 3\epsilon$, ottenendo dunque una stima migliore del necessario, non essendoci d al denominatore. Per fare questo si supponga per assurdo l'esistenza di un raggio $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ per cui non valga questa limitazione;

intanto si avvicini $\widetilde{\mathrm{A}}$ a $\widetilde{\mathrm{B}}$ finché la sua ordinata non sia esattamente pari a 3ϵ: grazie alle ipotesi la sua ascissa resta negativa; perché allora il segmento $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ possa incontrare il quadrato, è necessario che $\widetilde{\mathrm{B}}$ sia un punto della circonferenza lungo l'arco $\overset{\frown}{\mathrm{PQ}}$; inoltre l'ordinata di $\widetilde{\mathrm{A}}$, pur essendo negativa, deve essere maggiore dell'ordinata di R. Abbiamo poi già notato che deve avere luogo la condizione $\overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}} \leq \overline{\widetilde{\mathrm{A}}\mathrm{B}}$; facendo l'asse del segmento $\mathrm{B}\widetilde{\mathrm{B}}$, allora, $\widetilde{\mathrm{A}}$ si deve trovare alla sua sinistra, e quindi la sua ascissa deve essere minore di quella del punto di intersezione tra l'asse di $\mathrm{B}\widetilde{\mathrm{B}}$ e la retta orizzontale di ordinata 3ϵ; questo punto di intersezione, essendo $\widetilde{\mathrm{B}}$ nell'arco $\overset{\frown}{\mathrm{PQ}}$, ha ascissa massima nel caso in cui $\widetilde{\mathrm{B}} = \mathrm{P}$, e in questo caso è il punto S mostrato in figura. Abbiamo dunque visto che l'ascissa di $\widetilde{\mathrm{A}}$ deve essere maggiore di quella di R ma minore di quella di S, e allora avremo l'assurdo cercato mostrando che per ϵ sufficientemente piccolo $x_1(\mathrm{S}) < x_1(\mathrm{R})$, come avviene in figura. Per farlo sono sufficienti alcuni semplici conti: la retta PT, passando per i punti $(-\epsilon, \epsilon)$ e $T \equiv (0, 3\epsilon)$, ha equazione

$$y = 2x + 3\epsilon,$$

e quindi l'ascissa $\bar{x}$ del punto P cresce, per $\epsilon \searrow 0$, fino al caso limite in cui $\epsilon = 0$ per il quale $-\sqrt{d^2 - \bar{x}^2} = 2\bar{x} \Longrightarrow 5\bar{x}^2 = d^2$; allora per continuità si ha $4\bar{x}^2 \leq d^2$, e quindi $\bar{x} \geq -d/2$, purché ϵ sia sufficientemente piccolo; otteniamo allora una maggiorazione per l'ascissa di S considerando $\mathrm{P} \equiv (-d/2, -\sqrt{3}d/2)$; in questo caso la retta PB ha equazione

$$y = \frac{\sqrt{3}}{3}(x - d),$$

e quindi l'asse, dovendo passare per l'origine, ha equazione

$$y = -\sqrt{3}x;$$

ne segue che il punto d'incontro tra questo asse e la retta orizzontale di ordinata 3ϵ è $(-\sqrt{3}\epsilon, 3\epsilon)$, e quindi $x_1(\mathrm{S}) < -\sqrt{3}\epsilon$; per quanto riguarda R, invece, è chiaro per similitudine che

$$\frac{d-\epsilon}{\epsilon} = \frac{d+3\epsilon}{-x_1(\mathrm{R})} \Longrightarrow x_1(\mathrm{R}) = -\frac{d+3\epsilon}{d-\epsilon}\,\epsilon > -\sqrt{3}\epsilon > x_1(\mathrm{S})$$

per ϵ sufficientemente piccolo, da cui la tesi in questo primo caso.

ii) $x_1(\widetilde{\mathrm{B}}) < 0$ e $x_1(\widetilde{\mathrm{A}}) > 0$.

Non c'è bisogno di affrontare questa situazione, in quanto ci si può ricondurre al caso iii) ribaltando $\widetilde{\mathrm{A}}$ e $\widetilde{\mathrm{B}}$ rispetto al diametro verticale: restano infatti invariati $\overline{\mathrm{AB}}$, $\overline{\mathrm{A}\widetilde{\mathrm{B}}}$ e $\overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}}$, mentre aumenta $\overline{\widetilde{\mathrm{A}}\mathrm{B}}$; inoltre $x_2(\widetilde{\mathrm{A}})$ non cambia con questo spostamento.

iii) $x_1(\widetilde{\mathrm{B}}) \geq 0$. Questa situazione è la più scomoda da trattare; intanto ci si può

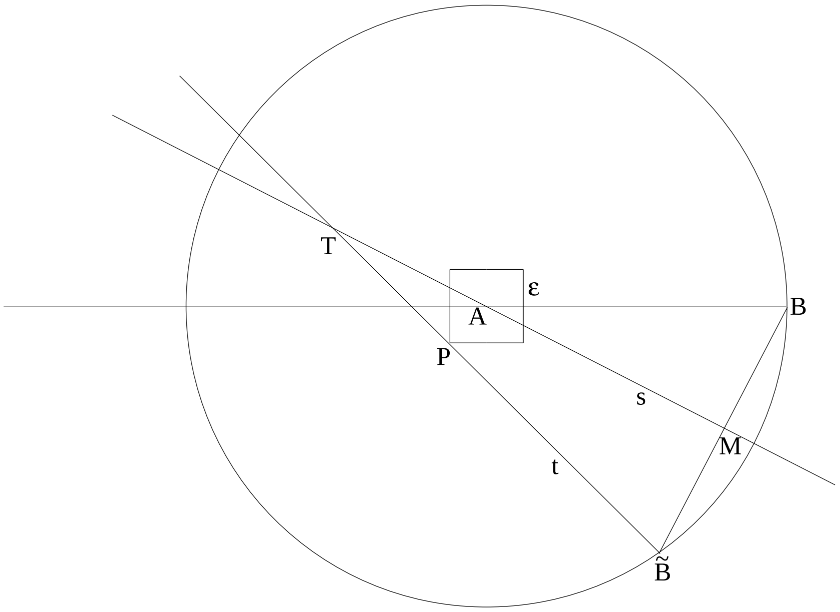

Figura 4.4: Caso iii)

ricondurre al caso in cui

$$x_2(\widetilde{\mathrm{B}}) \leq -5\epsilon;$$

se l'ordinata di $\widetilde{\mathrm{B}}$ è compresa fra -5ϵ e 0, infatti, si può maggiorare l'ordinata di $\widetilde{\mathrm{A}}$ con $12R\epsilon/d$, in maniera del tutto analoga a quanto già fatto per minorare l'ordinata o dare una stima alle coordinate dalla terza in poi. L'idea che seguiremo è illustrata nella figura 4.4, e consiste sostanzialmente nel notare che $\widetilde{\mathrm{A}}$ deve trovarsi a destra della retta t perché il segmento $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ possa incontrare il quadrato, ma a sinistra di s, l'asse del segmento $\mathrm{B}\widetilde{\mathrm{B}}$, perché deve essere $\overline{\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}} \leq \overline{\widetilde{\mathrm{A}}\mathrm{B}}$. Ne segue che l'ordinata di $\widetilde{\mathrm{A}}$ deve essere minore dell'ordinata di T. In questo caso, per fortuna, si può considerare $d = 1$, a meno di similitudine: infatti questo sarebbe stato sbagliato ogni volta che si è usata l'ipotesi di limitatezza di Ω, perché non si sarebbe più potuto stimare il suo diametro sempre con la stessa costante; in questo caso, invece, non serve usare questa ipotesi e quindi, se mostriamo che $x_2(\mathrm{T}) \leq C\epsilon$ per ϵ sufficientemente piccolo nel caso $d = 1$, in generale si avrà per similitudine $x_2(\mathrm{T}) \leq C\epsilon d \leq CR\epsilon$, e quindi la tesi con d generico.

Definiamo allora θ l'angolo $\mathrm{B}\widehat{\mathrm{A}}\mathrm{M}$, che deve essere compreso tra un certo $\theta_0 > 0$ e $\pi/4$, dove il primo angolo tiene conto del fatto che $x_2(\widetilde{\mathrm{B}}) \leq -5\epsilon$. In questo modo si possono scrivere le coordinate dei punti come segue:

$$\begin{aligned} \mathrm{A} &\equiv (0,0) \qquad \mathrm{B} \equiv (1,0) \qquad \widetilde{\mathrm{B}} \equiv (\cos(2\theta), -\sin(2\theta)) \\ \mathrm{M} &\equiv (\cos(\theta), -\sin(\theta)) \qquad \mathrm{P} \equiv (-\epsilon, -\epsilon). \end{aligned}$$

Allora l'equazione della retta t, passante per P e per $\widetilde{\mathrm{B}}$, è

$$y = \frac{-\sin(2\theta) + \epsilon}{\cos(2\theta) + \epsilon}\, x - \epsilon\, \frac{\sin(2\theta) + \cos(2\theta)}{\cos(2\theta) + \epsilon}\ ,$$

mentre l'equazione di s, passante per l'origine A e per M, è

$$y = -\frac{\sin\theta}{\cos\theta}\, x;$$

allora il punto T di intersezione tra s e t è tale che

$$-\frac{\sin\theta}{\cos\theta}\, x = \frac{-\sin(2\theta) + \epsilon}{\cos(2\theta) + \epsilon}\, x - \epsilon\, \frac{\sin(2\theta) + \cos(2\theta)}{\cos(2\theta) + \epsilon}\ .$$

Visto così non è molto utilizzabile, per cui risulta utile definire $\alpha = \sin\theta$; allora le limitazioni su θ permettono di desumere quelle su α: ci basta notare che per ϵ sufficientemente piccolo deve essere $\sqrt{2}/2 \geq \alpha \geq 2\epsilon$. Con questo cambio di variabile l'equazione precedente diventa

$$\frac{-\alpha}{\sqrt{1-\alpha^2}}\, x = \frac{-2\alpha\sqrt{1-\alpha^2} + \epsilon}{1 - 2\alpha^2 + \epsilon}\, x - \epsilon\, \frac{2\alpha\sqrt{1-\alpha^2} + 1 - 2\alpha^2}{1 - 2\alpha^2 + \epsilon}\ ;$$

visto poi che $y \in s$, si ha $y = -\alpha x/\sqrt{1-\alpha^2}$, e quindi l'equazione, raccogliendo opportunamente, diventa

$$\frac{y}{\epsilon} = -\frac{\left(\dfrac{2\alpha\sqrt{1-\alpha^2} + 1 - 2\alpha^2}{1 - 2\alpha^2 + \epsilon} \cdot \dfrac{\alpha}{\sqrt{1-\alpha^2}}\right)}{\left(\dfrac{\alpha}{\sqrt{1-\alpha^2}} + \dfrac{-2\alpha\sqrt{1-\alpha^2} + \epsilon}{1 - 2\alpha^2 + \epsilon}\right)}\ . \tag{4.7}$$

Definendo allora per semplicità $\tau = y/\epsilon$, la tesi risulterà definitivamente provata mostrando che τ è limitata. Innanzitutto notiamo che questa funzione deve essere definita sul triangolo $(0 < \epsilon \leq \sqrt{2}/4\ ;\ 2\epsilon \leq \alpha \leq \sqrt{2}/2)$ ed è continua; è anche ovvio che resta continua estendendola sul segmento $\epsilon = 0$ purché non si annulli nessun denominatore. Ora α è minore di 1, per cui $\sqrt{1-\alpha^2}$ non si annulla in nessun caso, mentre $1 - 2\alpha^2 + \epsilon$ si annulla solo per $\alpha = \sqrt{2}/2$ e $\epsilon = 0$. Si deve ora vedere quando si annulla il denominatore grosso, ma un semplice conto assicura che, a meno che non sia $\alpha = \epsilon = 0$,

$$\frac{\alpha}{\sqrt{1-\alpha^2}} + \frac{-2\alpha\sqrt{1-\alpha^2} + \epsilon}{1 - 2\alpha^2 + \epsilon} = 0 \iff \frac{\alpha}{\sqrt{1-\alpha^2}} = \frac{2\alpha\sqrt{1-\alpha^2} - \epsilon}{1 - 2\alpha^2 + \epsilon}$$

$$\iff \alpha = \epsilon\left(\alpha + \sqrt{1-\alpha^2}\right) \implies \alpha \leq \epsilon\left(\frac{\sqrt{2}}{2} + 1\right) \implies \alpha < 2\epsilon;$$

visto che sappiamo $\alpha \geq 2\epsilon$, si può concludere che il denominatore grosso si annulla solo nel caso $\alpha = \epsilon = 0$. La funzione τ è allora definita e continua su tutto il triangolo chiuso meno i due vertici $(0,0)$ e $(0,\sqrt{2}/2)$; per compattezza, se τ non esplode vicino a quei due punti τ sarà limitata. Se $(\epsilon,\alpha) \longrightarrow (0,\sqrt{2}/2)$, chiamando $\varphi = 1-2\alpha^2+\epsilon$ e $\rho = \alpha/\sqrt{1-\alpha^2}$ si ha che $\varphi \longrightarrow 0$, $\rho \longrightarrow 1$. Inoltre risulta, essendo $\varphi \neq 0$,

$$\tau = -\frac{\left(\dfrac{2\alpha\sqrt{1-\alpha^2}+1-2\alpha^2}{\varphi}\cdot\rho\right)}{\left(\rho+\dfrac{-2\alpha\sqrt{1-\alpha^2}+\epsilon}{\varphi}\right)} = -\frac{\left(2\alpha\sqrt{1-\alpha^2}+1-2\alpha^2\right)\cdot\rho}{\rho\varphi-2\alpha\sqrt{1-\alpha^2}+\epsilon} \longrightarrow 1.$$

Dunque in $(0,\sqrt{2}/2)$ non ci sono problemi, perché τ può addirittura essere estesa in modo continuo. Vicino a $(0,0)$, invece, si può notare che i termini $2\alpha\sqrt{1-\alpha^2}+1-2\alpha^2$, $1-2\alpha^2+\epsilon$ e $\sqrt{1-\alpha^2}$ tendono ad 1; dividendo allora nel membro di destra di (4.7) –che definisce τ– numeratore e denominatore per α, il numeratore tende a -1 e dunque basta mostrare che non si avvicini a zero il denominatore

$$\frac{1}{\sqrt{1-\alpha^2}}+\frac{\epsilon/\alpha-2\sqrt{1-\alpha^2}}{1-2\alpha^2+\epsilon}.$$

Visto poi che $\sqrt{1-\alpha^2}$ e $1-2\alpha^2+\epsilon$ tendono ad 1, basta che $2\sqrt{1-\alpha^2}-\epsilon/\alpha$ non si avvicini ad 1. Ma $2\sqrt{1-\alpha^2} \longrightarrow 2$, mentre $\alpha \geq 2\epsilon$ e quindi $\epsilon/\alpha \leq 1/2$, per cui τ, pur non avendo limite in $(0,0)$, è comunque limitata in un suo intorno e quindi finalmente si ha la tesi.

■

4.4 Stime in L^P

In questa sezione studieremo in che modo le proprietà di sommabilità di f^+ ed f^- assicurino analoghe proprietà su μ; come già discusso nella scorsa sezione, per poter affrontare questo argomento sarà molto utile il teorema 4.7, che permette stime di tipo uniforme sulla dimensione puntuale valide anche agli estremi dei raggi, purché vi sia una certa distanza tra i supporti; vedremo in effetti che l'ipotesi di separazione dei supporti, di cui abbiamo già discusso nel capitolo 1, risulterà molto comoda, ma troveremo anche risultati importanti validi in sua assenza.

Strettamente collegata con le nozioni di dimensione è la questione dell'appartenenza di una misura allo spazio L^1, grazie all'immediata

Osservazione 4.8 $\gamma \in L^1 \Longrightarrow \mathcal{H}-\dim(\gamma) = n$.

Infatti, $\gamma \in L^1 \Longrightarrow \gamma \ll \mathcal{H}^n$, visto che, come ben noto, la misura di Hausdorff $n-$dimensionale coincide con quella di Lebesgue; a questo punto la dimensione di Hausdorff di una misura L^1 è almeno n e l'enunciato segue dal fatto che nessuna misura non nulla può avere dimensione superiore ad n, visto che $\mathcal{H}^k$ è identicamente nulla su $\mathbb{R}^n$ se $k > n$.

Sarebbe tuttavia sbagliato, per una misura, confondere l'appartenenza ad L^1 con la proprietà di avere dimensione massima, in quanto nell'osservazione appena fatta non è valida anche l'implicazione inversa; per rendersene conto basta notare che una misura appartiene ad L^1 se e solo se è assolutamente continua rispetto alla misura di Lebesgue, grazie al teorema di Radon-Nikodym, mentre ha dimensione n se e solo se è assolutamente continua rispetto ad $\mathcal{H}^k$ per ogni $k < n$, cosa leggermente più debole. Per quanto necessaria, questa distinzione è apprezzabile solamente in situazioni estreme, visto che una misura $n-$dimensionale che non sia L^1 deve avere parte singolare contenuta in un insieme che abbia $\mathcal{H}^n-$misura nulla ma $\mathcal{H}^k-$misura infinita per ogni $k < n$...

In questa sezione chiameremo per convenzione f^+ la misura, tra le due date, con le migliori proprietà di sommabilità, in analogia con quanto fatto nella prima sezione di questo capitolo, dove f^+ aveva dimensione maggiore. Per evitare di appesantire la notazione con nuove lettere, chiameremo poi f^+, f^- e μ anche le densità relative alle omonime misure, quando queste saranno assolutamente continue, visto che risulterà comunque chiaro quando ci riferiremo ad una misura e quando ad una funzione. Fatte queste premesse, possiamo mostrare subito la prima importante proprietà:

Teorema 4.9 *Se $f^+ \in L^1$, allora $\mu \in L^1$.*

<u>Dim</u> : Supponiamo innanzitutto che $f^+ \in L^\infty$; allora in particolare vale l'ipotesi (Ip), e quindi grazie al corollario 4.4 per $\mu-$q.o. x (in particolare per ogni x che sta all'interno

di qualche raggio di trasporto) esiste una costante $C(x)$ tale che

$$\mu(B(x,\rho)) \leq C(x)\rho^n.$$

Ma questo assicura che per $\mu-$quasi ogni x è finito $\theta_n^*(\mu,x)$, direttamente dalla definizione (2.4); se μ non fosse L^1, il supporto della parte singolare di μ sarebbe un insieme B trascurabile secondo Lebesgue ma di $\mu-$misura positiva; chiamando allora

$$B_m := \{x \in B : \theta_n^*(\mu,x) \leq m\},$$

si avrebbe che i B_m invadono B a meno di un insieme $\mu-$trascurabile; ora, grazie alla parte i) del teorema 2.2, si otterrebbe che

$$\mu(B_m) \leq 2^n m\mathcal{H}^n(B_m) \leq 2^n m\mathcal{H}^n(B) = 0,$$

e per subadditività si troverebbe $\mu(B) = 0$, assurdo; si può quindi concludere che $\mu \in L^1$.

Nel caso generale si può usare un metodo di approssimazione pressoché identico a quanto fatto nella dimostrazione del teorema 4.6, defininendo $\Omega_m = \{x : f^+(x) \leq m\}$, e quindi

$$f_m^+ := f^+ \llcorner \Omega_m;$$

visto che gli Ω_m invadono Ω a meno di un insieme f^+-trascurabile si possono definire, esattamente come nel teorema 4.6, γ_m e μ_m tali che $\mu_m \longrightarrow \mu$; ogni μ_m, poi, è assolutamente continua grazie a quanto appena mostrato nell'ipotesi $f^+ \in L^\infty$, e quindi passando al limite è assolutamente continua anche μ.

■

Per avere allora l'assoluta continuità della μ ottimale non è necessario che siano assolutamente continue sia f^+ che f^-, anzi è sufficiente che lo sia solo una delle due.

Vale la pena notare che in questo teorema, così come nel 4.6, abbiamo considerato f^+ ristretta ad un sottoinsieme di Ω e poi abbiamo rinominato opportunamente γ e μ; a questo punto abbiamo applicato dei risultati mostrati nell'ipotesi che μ sia ottenuta da γ tramite la (3.15) e γ sia una misura ottimale nel senso di Kantorovich con marginali f^+ ed f^-: avendo però cambiato le misure in gioco, questo non è più assolutamente vero, anche perché a questo punto f^+ ed f^- non hanno più la stessa massa, e quindi non ha nemmeno senso porsi il problema del trasporto ottimale. Nella dimostrazione del teorema 4.6 avevamo aggirato questo ostacolo notando che tutti i risultati applicati utilizzavano in realtà solo l'ipotesi che γ fosse c-ciclicamente monotona, cosa che continua ad essere vera, e lo stesso discorso si può fare anche per il teorema 4.9; pur essendo corretto, questo modo di ragionare è molto poco elegante, e soprattutto estremamente scomodo, visto che obbliga a controllare che nelle dimostrazioni di tutti i risultati applicati si usa soltanto l'ipotesi della c-ciclica monotonia. E' possibile però fare un diverso ragionamento che sistema la questione: basta notare che in effetti la nuova γ è ottimale per un opportuno problema per il quale uno dei dati iniziali sia la f^+ ristretta; grazie a ciò, sarà legittimo applicare i risultati fin qui mostrati senza bisogno di ricorrere a scomodi ragionamenti ad hoc. Dunque è parecchio utile la seguente

Osservazione 4.10 *Sia γ ottimale nel senso di Kantorovich per il problema con dati (f^+, f^-), e sia $\widetilde{\Omega} \subseteq \Omega$ un insieme di Borel; detta allora*

$$\tilde{\gamma} := \gamma \llcorner \widetilde{\Omega} \times \Omega,$$

$\tilde{\gamma}$ risulta ottimale per il problema avente come dati i suoi marginali.

Dim : Definiamo $\tilde{f}^+ = f^+ \llcorner \widetilde{\Omega}$; per poter applicare questa osservazione come discusso sopra, è necessario che sia $\tilde{\gamma}_1 = \tilde{f}^+$, e questo segue immediatamente dal fatto che

$$\tilde{\gamma}_1(A) = \tilde{\gamma}(A \times \Omega) = \gamma(A \times \Omega \cap \widetilde{\Omega} \times \Omega) = \gamma(A \cap \widetilde{\Omega} \times \Omega) = f^+(A \cap \widetilde{\Omega}) = \tilde{f}^+(A).$$

Definiamo a questo punto $\tilde{f}^- = \tilde{\gamma}_2$; allora $\tilde{\gamma}$ è una misura che ammette come marginali $\tilde{f}^+$ e $\tilde{f}^-$, per cui ha senso chiedersi se è ottimale nel senso di Kantorovich. Se non lo fosse, esisterebbe una $\tilde{\alpha}$ di costo totale strettamente minore; definendo ora

$$\alpha = \begin{cases} \tilde{\alpha} & \text{su } \widetilde{\Omega} \times \Omega \\ \gamma & \text{su } {}^c\widetilde{\Omega} \times \Omega \end{cases},$$

è immediato mostrare che α ammette f^+ ed f^- come marginali ed ha costo totale strettamente minore di γ, violando così la sua minimalità; l'assurdo mostra la tesi.

L'unica obiezione potrebbe essere che la nuova γ non ha massa totale unitaria, ma è chiaro che questo non ha influenza sui teoremi fatti finora, perché il problema di Kantorovich ha bisogno solo che le due misure date abbiano la stessa massa, averla posta unitaria è spesso più comodo ma non cambia nulla.

■

Il teorema 4.9 è solo il primo scalino nello studio della sommabilità di μ; nel seguito di questa sezione cercheremo di trovare ipotesi su f^+ ed f^- che assicurino alla misura ottimale μ una maggiore sommabilità; un primo risultato in questo senso che fa uso del teorema 4.7 è il

Lemma 4.11 *Se $f^+ \in L^\infty$, definito $B = \{x : d(x, supp(f^-)) \geq d\}$ la norma L^∞ di μ ristretta a B è finita. In particolare esiste una costante K dipendente solo da Ω tale che*

$$\|\mu\|_{L^\infty(B)} \leq \frac{K}{d^{n-1}} \|f^+\|_\infty .$$

Dim : Sia $x \in B$; a meno di escludere un insieme $\mu-$trascurabile, come già visto, si può supporre x interno ad un raggio AB; avvicinato allora A a B fino ad incontrare x, cosa che sappiamo essere lecita, si ha naturalmente $d(\mathrm{A}, supp(f^-)) \geq d$, visto che a questo punto A e x coincidono e $x \in B$; applicando allora il teorema 4.7 si trova una regione $C(x,\epsilon)$ di volume minore di $K\epsilon^{n-1}/d^{n-1}$ che contiene il primo estremo di tutti i raggi che incontrano $B(x,\epsilon)$. Questo assicura che

$$\mu(B(x,\epsilon)) = \int_{C(x,\epsilon)\times\Omega} l(\mathrm{AB} \cap B(x,\epsilon))d\gamma \leq 2\epsilon \int_{C(x,\epsilon)} df^+ \leq 2K\frac{\epsilon^n}{d^{n-1}}\|f^+\|_\infty.$$

Ma grazie al teorema 4.9 sappiamo già essere $\mu \in L^1$, per cui, a meno di un insieme $\mu-$trascurabile, si può supporre che x sia un punto di Lebesgue. Passando dunque al limite per ϵ che tende a zero si ottiene la stima enunciata, e quindi la tesi.

■

Sempre dal teorema 4.7 si può ricavare un risultato analogo nel caso in cui sia $p < +\infty$:

Lemma 4.12 *Se $f^+ \in L^p$, $p < +\infty$, definito $B = \{x : d(x, supp(f^-)) \geq d\}$ la norma L^p di μ ristretta a B è finita; in particolare*

$$\|\mu\|_{L^p(B)} \leq \frac{K\|f^+\|_p}{d^{\frac{n-1}{p'}}}$$

Dim : Innanzitutto dato $C \in \mathbb{R}$ definiamo $A = \{x \in supp(f^+) : C/2 \leq f^+(x) \leq C\}$.
i) Supponiamo per ora che $supp(f^+) = A$, poi riusciremo ad eliminare questa ipotesi: preso un generico punto $x \in B$, si può supporre che appartenga ad un raggio di trasporto AB, escludendo con questo solo un insieme $\mu-$trascurabile; grazie al teorema 4.7, come già visto, gli $\widetilde{\mathrm{A}}$ per cui un raggio $\widetilde{\mathrm{A}}\widetilde{\mathrm{B}}$ può incontrare $B(x,\epsilon)$ sono contenuti in un insieme Δ di misura minore di $\lambda\epsilon^{n-1}/d^{n-1}$, per costruzione di B. Ma allora

$$\mu(B(x,\epsilon)) \leq 2\epsilon f^+(\Delta) \leq 2\epsilon\ \|f^+\|_\infty\ \mathcal{H}^n(\Delta) \leq \frac{2\lambda C\epsilon^n}{d^{n-1}},$$

da cui, supponendo senza perdita di generalità che x sia un punto di Lebesgue, si ottiene

$$\mu(x) \leq \frac{2\lambda C}{d^{n-1}}.$$

Abbiamo dunque una limitazione della norma L^∞ di μ ristretta a B, poco utile per ora, visto che è subordinata alla costante C e quindi all'ipotesi fatta all'inizio che $supp(f^+) = A$; cerchiamo allora di stimare anche la norma L^1 come segue, ricordando che μ è positiva, che il diametro di Ω è limitato da R e che il primo marginale di γ è f^+:

$$\|\mu\|_{L^1(B)} = \int_B \mu(x)dx = \mu(B) = \int_{\Omega\times\Omega} l(xy \cap B)d\gamma(x,y) \leq R\int_\Omega f^+(x)dx \leq Rf^+(A).$$

In generale, avendo una limitazione in L^1 ed una in L^∞, se ne può dedurre una in L^p: infatti chiaramente la norma L^p risulta massima nel caso in cui μ valga esattamente $\|\mu\|_\infty$ su un insieme di misura tale che venga esaurita la norma L^1, e poi 0 fuori, e dunque si trova immediatamente la stima

$$\|\mu\|_p \leq \|\mu\|_\infty^{1/p'} \|\mu\|_1^{1/p}. \tag{4.8}$$

In questo caso la formula appena trovata assicura

$$\|\mu\|_{L^p(B)} \leq \left(\frac{2\lambda C}{d^{n-1}}\right)^{1/p'} \big(Rf^+(A)\big)^{1/p} \leq \frac{2\lambda^{1/p'} R^{1/p} C}{d^{\frac{n-1}{p'}}} \left(\frac{f^+(A)}{C}\right)^{1/p}. \tag{4.9}$$

Usando inoltre il fatto che, per ipotesi, su A si ha $C/2 \leq f^+(x) \leq C$, si possono anche dare le stime $\|f^+\|_p \geq C/2\ \ \mathcal{H}^n(A)^{1/p}$ e $f^+(A) \leq C\ \mathcal{H}^n(A)$ che, messe insieme, assicurano

$$\|f^+\|_p \geq \frac{C}{2}\left(\frac{f^+(A)}{C}\right)^{1/p};$$

questa disuguaglianza, sostituita nella (4.9), porta a

$$\|\mu\|_{L^p(B)} \leq \frac{4\lambda^{1/p'} R^{1/p}}{d^{\frac{n-1}{p'}}} \|f^+\|_p\ . \tag{4.10}$$

ii) Visto che nella disuguaglianza (4.10) è finalmente sparito ogni riferimento a C e dunque all'ipotesi iniziale, si può scomporre il supporto di f^+ in numerabili regioni A_z, $z \in \mathbb{Z}$, in modo che $A_z = \{x \in supp(f^+) : 2^z < f^+(x) \leq 2^{z+1}\}$; la f^+ ristretta ad A_z permette di definire una μ_z, ed è chiaro che $\mu = \sum_{z\in\mathbb{Z}} \mu_z$. Grazie all'osservazione 4.10, allora, alle misure ristrette si può applicare il ragionamento fatto fin qui ottenendo

$$\|\mu_z\|_{L^p(B)} \leq \frac{4\lambda^{1/p'} R^{1/p}}{d^{\frac{n-1}{p'}}} \|f^+\|_{L^p(A_z)}. \tag{4.11}$$

iii) Purtroppo dalla (4.11) non è affatto immediato concludere la tesi: sommando su $z \in \mathbb{Z}$, in effetti, il membro di sinistra maggiora la norma L^p di μ, per la disuguaglianza triangolare, ma per lo stesso motivo il membro di destra, a parte le costanti moltiplicative, è maggiore della norma L^p di f^+, anziché stimarla; anzi si può facilmente notare che in generale l'ultima disuguaglianza può valere anche in casi in cui $\mu \notin L^p$. E' dunque opportuno a questo punto fare un passo indietro e sfruttare bene la situazione particolare in cui ci troviamo: ricapitolando quanto visto finora si deduce che la norma in L^p di μ è massima se $\mu_z = C_z$ su un insieme Ω_z di misura V_z e gli insiemi Ω_z sono inscatolati. Tuttavia è piuttosto scomodo cercare di scrivere esattamente la norma L^p di μ in questo caso, visto che V_z, la misura di Ω_z, non è decrescente in z –anche se chiaramente tende a

0 per $z \to +\infty$; per questo motivo, allora, nel caso "peggiore" gli Ω_z sono effettivamente inscatolari, ma non esattamente secondo l'indice z.

iv) Si ottiene una maggiorazione per la norma $L^p(B)$ di μ ipotizzando che sia

$$\mu = \sum_{j=-\infty}^{z} C_j$$

su un insieme di misura V_z e che questi insiemi siano disgiunti.

Per capire come mai questo porti ad una maggiorazione può essere utile vedere cosa succede nel caso in cui vi siano due soli insiemi di misure V_1 e V_2 su cui μ_1 e μ_2 valgano rispettivamente C_1 e C_2: detto V il volume dell'intersezione tra i due insiemi, la funzione μ vale C_1 su un insieme di misura $V_1 - V$, C_2 su uno di misura $V_2 - V$ e $C_1 + C_2$ su uno di misura V, mentre la nostra ipotesi è che valga C_1 su un insieme di misura V_1 e $C_1 + C_2$ su uno di misura V_2, ottenendo quindi una sovrastima; basandosi su questa osservazione è immediato mostrare che l'ipotesi fatta porti ad una maggiorazione nel caso di finiti insiemi, ed un banale passaggio al limite assicura l'affermazione fatta in questo punto.

v) Grazie a quanto visto nel punto iv) si può finalmente giungere ad una stima che provi la tesi, ricordando che nella nostra costruzione, a meno di un fattore moltiplicativo, si ha $C_z = 2^z$, e quindi $\sum_{j=-\infty}^{z} C_j \leq 2C_z$; si ha dunque

$$\int_B |\mu(x)|^p \leq \sum_{z=-\infty}^{+\infty} (2C_z)^p V_z = 2^p \sum_{z=-\infty}^{+\infty} C_z^p V_z.$$

D'altra parte la (4.11) limita la norma di μ_z proprio considerando il caso limite in cui valga costantemente C_z su un insieme di misura V_z, e quindi assicura

$$C_z^p V_z \leq \left(\frac{4\lambda^{1/p'} R^{1/p}}{d^{\frac{n-1}{p'}}} \right)^p \int_{A_z} |f^+(x)|^p.$$

Visto che gli insiemi A_z sono disgiunti, allora, si può concludere che

$$\int_B |\mu(x)|^p \leq \left(\frac{8\lambda^{1/p'} R^{1/p}}{d^{\frac{n-1}{p'}}} \right)^p \sum_{z=-\infty}^{+\infty} \int_{A_z} |f^+(x)|^p = \left(\frac{8\lambda^{1/p'} R^{1/p}}{d^{\frac{n-1}{p'}}} \right)^p \int_{supp(f^+)} |f^+(x)|^p$$

$$\Downarrow$$

$$\|\mu\|_{L^p(B)} \leq \frac{8\lambda^{1/p'} R^{1/p}}{d^{\frac{n-1}{p'}}} \|f^+\|_p ,$$

ottenendo così la stima enunciata e quindi la tesi.

■

I due lemmi che abbiamo dimostrato permettono di giungere facilmente ad un bel risultato riguardo alla sommabilità di μ:

Teorema 4.13 *Se $f^+ \in L^q$ e $f^- \in L^p$, con $1 \leq p \leq q \leq +\infty$, e i supporti sono separati allora $\mu \in L^p$; in particolare, detta d la distanza fra i supporti,*

$$\|\mu\|_p \leq \frac{K}{d^{\frac{n-1}{p'}}} \left(\|f^+\|_p \vee \|f^-\|_p\right).$$

<u>Dim</u> : Chiamiamo $B_1 = \{x : d(x, supp(f^-)) \geq d/2\}$ e $B_2 = \{x : d(x, supp(f^+)) \geq d/2\}$; visto che i supporti sono separati da una distanza d, è chiaro che B_1 e B_2 ricoprono Ω; inoltre, essendo Ω limitato, sia f^+ che f^- appartengono a L^p perché $q \geq p$. Applicando allora il lemma 4.11 o il 4.12 a seconda che sia o meno $p = \infty$, si trova

$$\|\mu\|_{L^p(B_1)} \leq \frac{K\|f^+\|_p}{d^{\frac{n-1}{p'}}} \qquad\qquad \|\mu\|_{L^p(B_2)} \leq \frac{K\|f^-\|_p}{d^{\frac{n-1}{p'}}};$$

dal momento che B_1 e B_2 ricoprono Ω si ha $\|\mu\|_p \leq \|\mu\|_{L^p(B_1)} + \|\mu\|_{L^p(B_2)}$ e quindi, sommando le due disequazioni di sopra, si ottiene la tesi.

■

Si noti che l'ipotesi di separazione dei supporti, già incontrata, ci ha permesso di mostrare un risultato alquanto soddisfacente nel caso in cui entrambe le misure date siano sommabili.

A questo punto cambiamo decisamente obiettivo, cercando di mostrare che, purché almeno una fra f^+ ed f^- appartenga ad uno spazio L^q con $q > 1$, anche μ sta in uno spazio L^p con un opportuno $p > 1$. Come già fatto per altri teoremi, cominciamo con un risultato valido se $f^+ \in L^\infty$ e poi, grazie a questo, passiamo al caso generale; si noti che, visto che lo vorremo applicare ad opportune restrizioni di f^+, non supporremo $f^+(\Omega) = 1$.

Lemma 4.14 *Se $p < (2n)'$, allora $f^+ \in L^\infty \Longrightarrow \mu \in L^p$. In particolare*

$$\|\mu\|_p \leq K\ \|f^+\|_\infty^{1/p'}\ f^+(\Omega)^{1/p}.$$

<u>Dim</u> : **i)** Per prima cosa introduciamo la misura μ_r, ricordando che μ era stata definita tramite la

$$\langle \mu, \varphi \rangle := \int_{\Omega\times\Omega} \left(\int_\Omega \varphi(z)\ d\mathcal{H}^1_{xy}(z) \right) d\gamma(x,y); \tag{3.15}$$

dati $(x,y) \in supp(\gamma)$ e $\sigma \geq 0$, intanto, indichiamo con y_σ il punto lungo il segmento xy che abbia distanza σ da y, oppure x se $\overline{xy} < \sigma$; a questo punto la misura μ_r sia data da

$$\langle \mu_r, \varphi \rangle := \int_{\Omega\times\Omega} \left(\int_\Omega \varphi(z) d\mathcal{H}^1_{y_r y_{2r}}(z) \right) d\gamma(x,y); \tag{4.12}$$

come già fatto nella (3.16), è comodo avere una scrittura della μ_r-misura di un insieme, che si ottiene immediatamente dalla (4.12):

$$\mu_r(A) = \int_{\Omega\times\Omega} l(y_r y_{2r} \cap A) d\gamma(x,y).$$

Visto che $l(y_r y_{2r} \cap A) \geq 0$ e che $\sum\{l(y_r y_{2r} \cap A) : r = R, R/2, R/4, \dots\} = l(xy \cap A)$, applicando il teorema di Beppo–Levi si trova subito che $\mu = \mu_R + \mu_{R/2} + \mu_{R/4} + \cdots$; essendo poi $f^+ \in L^\infty$, è anche $\mu \in L^1$, e quindi $\mu_r \in L^1$ per ogni r, visto che $\mu_r \ll \mu \ll \mathcal{H}^n$. Dato x, infine, si può osservare che se esiste δ tale che $\forall y \in supp(f^-)$, $d(x, y) \notin [r - \delta, 2r + \delta]$, allora $\mu_r(x) = 0$.

ii) A questo punto fissiamo r e supponiamo temporaneamente che tutta f^- sia concentrata in un cubo Q di lato ϵ con $\epsilon\sqrt{n} \leq r/2$. Grazie all'ultima osservazione fatta nel punto i), allora, il supporto di μ_r sarà contenuto entro una corona sferica S centrata nel centro del cubo di raggi $r - \epsilon\sqrt{n}/2$ e $2r + \epsilon\sqrt{n}/2$. Per calcolare $\mu_r(S)$, allora, basta notare che chiaramente $l(y_r y_{2r} \cap A) \leq r$ e quindi

$$\mu_r(S) \leq r \ f^-(Q);$$

visto poi che dall'ipotesi $\epsilon\sqrt{n} \leq r/2$ segue che ogni punto di S dista almeno $r/2$ da Q, grazie al lemma 4.11 si ha la stima

$$|\mu_r(x)| \leq \frac{C\|f^+\|_\infty}{r^{n-1}}$$

valida μ_r-q.o.; abbiamo dunque per μ_r una stima in L^1 ed una in L^∞ e questo, grazie a (4.8), permette di osservare che

$$\|\mu_r\|_p \leq \left(\frac{C\|f^+\|_\infty}{r^{n-1}}\right)^{1/p'} \big(r \ f^-(Q)\big)^{1/p} \leq \frac{K \ \|f^+\|_\infty^{1/p'} \ f^-(Q)^{1/p}}{r^{n-1-n/p}},$$

definendo la costante $K = C^{1/p'}$.

iii) Mantenendo ancora fissato r, togliamo l'ipotesi che il supporto di f^- sia contenuto in un unico cubo di lato ϵ con $\epsilon\sqrt{n} \leq r/2$; ricopriamolo invece con $N(r)$ cubi di quel lato, Q_1, Q_2, $\dots Q_{N(r)}$, e definiamo $\mu_{r,i}$ la frazione di μ_r originata da raggi che partano da Q_i, restringendo opportunamente γ come già fatto più volte: ne segue che $\mu_r = \mu_{r,1} + \cdots + \mu_{r,N(r)}$; il punto ii) assicura che

$$\|\mu_{r,i}\|_p \leq \frac{K \ \|f^+\|_\infty^{1/p'} \ f^-(Q_i)^{1/p}}{r^{n-1-n/p}}.$$

Grazie alla ben nota disuguaglianza fra media $p-$esima e media aritmetica si ha

$$\frac{\mu_{r,1} + \cdots + \mu_{r,N(r)}}{N(r)} \leq \left(\frac{\mu_{r,1}^p + \cdots + \mu_{r,N(r)}^p}{Nr}\right)^{1/p} \Longrightarrow \mu_r^p \leq N(r)^{p-1}\left(\mu_{r,1}^p + \cdots + \mu_{r,N(r)}^p\right),$$

che integrando ed estraendo la radice $p-$esima diventa

$$\begin{aligned}\|\mu_r\|_p &\leq N(r)^{1/p'}\left(\|\mu_{r,1}\|_p^p+\cdots+\|\mu_{r,N(r)}\|_p^p\right)^{1/p}\\ &\leq N(r)^{1/p'}\,\frac{K\ \|f^+\|_\infty^{1/p'}}{r^{n-1-n/p}}\ \left(f^-(Q_1)+\cdots+f^-(Q_{N(r)})\right)^{1/p}\\ &\leq N(r)^{1/p'}\ \frac{K\ \|f^+\|_\infty^{1/p'}}{r^{n-1-n/p}}\ f^+(\Omega)^{1/p},\end{aligned}\tag{4.13}$$

visto che $f^+(\Omega)=f^-(\Omega)$.

iv) Grazie a quanto visto nei punti ii) e iii) sappiamo quanto basta su μ_r: bisogna ora ricordare che, come notato nel punto i), $\mu=\mu_R+\mu_{R/2}+\mu_{R/4}+\cdots$. Allora dalla (4.13) segue direttamente la stima enunciata, e quindi alla tesi, pur di mostrare che per $p<(2n)'$ si ha

$$\sum_{i=0}^{+\infty}\frac{N(R/2^i)^{1/p'}}{(R/2^i)^{n-1-n/p}}<+\infty.$$

A questo punto bisogna controllare il comportamento di $N(r)$; ma è chiaro che, per come è definito, è possibile scegliere dei ricoprimenti "ottimali" in modo che a meno di una costante moltiplicativa si abbia $N(r)\leq 1/r^n$; è sufficiente allora mostrare che

$$\begin{aligned}\sum_{i=0}^{+\infty}\frac{1}{(R/2^i)^{n-1-n/p+n/p'}}<+\infty \quad&\Longleftrightarrow\quad \sum_{i=0}^{+\infty}\left(\frac{1}{(1/2)^{n-1-n/p+n/p'}}\right)^i<+\infty \Longleftrightarrow\\ 2^{n-1-n/p+n/p'}<1 \quad&\Longleftrightarrow\quad n-1-n/p+n/p'<0 \Longleftrightarrow p<(2n)',\end{aligned}$$

per cui la tesi risulta provata.

$\blacksquare$

Ottenuto quanto si voleva nel caso $f^+\in L^\infty$, si può subito generalizzare, concludendo questo lavoro con il

Teorema 4.15 *Se $f^+\in L^q$, allora $\mu\in L^p$ per ogni $p<(2n)'\wedge(1+(q-1)/2)$.*

Dim : In maniera simile a quanto già fatto in altre dimostrazioni, definiamo

$$\Omega_j:=\left\{x:|f^+(x)|\in[j-1,j)\right\}$$

e di conseguenza μ_j, in modo che $\mu=\sum_j\ \mu_j$; grazie al lemma 4.14 si ha che, purché sia $p<(2n)'$,

$$\|\mu_j\|_p\leq Kj^{1/p'}\ f^+(\Omega_j)^{1/p}\leq Kj^{1/p'}\left(j\cdot|\Omega_j|\right)^{1/p}\leq Kj\ |\Omega_j|^{1/p}.$$

Definito allora $\lambda_j=|\Omega_j|$, grazie alla disuguaglianza triangolare della norma L^p basta mostrare

$$\sum_j j\ \lambda_j^{1/p}<+\infty.$$

Ma per ipotesi $f^+ \in L^q$, per cui di certo $\sum_j (j-1)^q \ \lambda_j < +\infty$ e quindi

$$\sum_j j^q \ \lambda_j < +\infty; \tag{4.14}$$

evidentemente si avrebbe la tesi se fosse $j \ \lambda_j^{1/p} \leq j^q \ \lambda_j$, o equivalentemente $\lambda_j \geq j^{p'(1-q)}$, ma non è detto che sia così; possiamo però definire

$$\rho_j := \begin{cases} \lambda_j & \text{se } \lambda_j \geq j^{p'(1-q)} \\ 0 & \text{se } \lambda_j < j^{p'(1-q)} \end{cases} \qquad \sigma_j := \begin{cases} 0 & \text{se } \lambda_j \geq j^{p'(1-q)} \\ \lambda_j & \text{se } \lambda_j < j^{p'(1-q)} \end{cases};$$

grazie alla definizione, ci basta chiaramente mostrare separatamente $\sum_j \ j \ \rho_j^{1/p} < +\infty$ e $\sum_j \ j \ \sigma_j^{1/p} < +\infty$. La prima diseguaglianza segue, come notato prima, direttamente dalla (4.14), visto che $\rho_j \geq j^{p'(1-q)}$ per ogni indice j per cui non sia $\rho_j = 0$; per quanto riguarda la seconda, invece, sappiamo che $\sigma_j < j^{p'(1-q)}$, che assicura

$$j \ \sigma_j^{1/p} < j^{1+\frac{p'(1-q)}{p}} = j^{1-\frac{q-1}{p-1}};$$

allora è vero che $\sum_j \ j \ \sigma_j^{1/p} < +\infty$ purché sia

$$1 - \frac{q-1}{p-1} < -1 \iff p < 1 + \frac{q-1}{2},$$

come mostra un semplice conto. Da questo segue la tesi.

■

Appendice A

Qualche risultato di analisi convessa

A.1 Un teorema di dualità

In questa sezione mostreremo un teorema di dualità che estende la Proposizione 2.3 del terzo capitolo in [11]; questo risultato, abbastanza generale, ci sarà molto utile nel capitolo 3. In accordo con le convenzioni di [11], in quanto segue considereremo funzioni a valori in $\overline{\mathbb{R}}$, definendo quindi convessa una funzione g tale che, dati x, y appartenenti al suo dominio e $\lambda \in (0,1)$, si ha

$$g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y) \tag{A.1}$$

ogni volta che sia definito il secondo membro; un modo di vedere la situazione che è equivalente, ma che può risultare più immediato da visualizzare, è definire convessa una funzione per cui valga la (A.1) per ogni scelta di x, y e λ, adottando però la convenzione che

$$-\infty + \infty = +\infty - \infty = +\infty.$$

Questo è quanto fatto, ad esempio, da Castaing e Valadier in [10], e la stessa convenzione verrà mantenuta nel resto di questa sezione. Premettiamo il

Lemma A.1 *Se $g : X \longrightarrow \overline{\mathbb{R}}$, essendo X uno spazio di Banach, è s.c.i. in 0 e convessa, allora $g(0) = g^{**}(0)$; se poi g è limitata in un intorno di 0, esiste $\bar{\xi} \in X^*$ tale che $g^{**}(0) = -g^*(\bar{\xi})$.*

<u>Dim</u> : Come noto, g^{**} è il sup delle funzioni lineari affini minori di g, per cui è vero in generale che $g^{**}(0) \leq g(0)$; per l'altra disuguaglianza si può supporre $g(0) > -\infty$, altrimenti è automaticamente vera. Notiamo allora intanto che se g è convessa lo è anche il suo sopragrafico $\mathcal{G}r(g) = \{(x,t) : t \geq g(x)\}$, come segue subito dalla definizione, e quindi anche la chiusura $\overline{\mathcal{G}r(g)}$. Ora, se fosse $(0, g^{**}(0)) \notin \overline{\mathcal{G}r(g)}$, potrei separarli con un

iperpiano chiuso grazie al teorema di Hahn-Banach, e questo è assurdo perché in questo modo g^{**} risulterebbe minore di una funzione lineare continua minore di g; ne segue $(0, g^{**}(0)) \in \overline{\mathcal{G}r(g)}$. Ora, se fosse $g^{**}(0) < g(0)$, per la semicontinuità inferiore di g si avrebbe

$$g(x) > (g(0) + g^{**}(0))/2$$

in un intorno di 0, il che va contro al fatto che $(0, g^{**}(0)) \in \overline{\mathcal{G}r(g)}$; si è dunque provato che $g(0) = g^{**}(0)$.

Per la seconda parte della tesi, si consideri g limitata in un intorno di 0; dalla definizione si ha che

$$g^{**}(0) = \sup_{\xi \in X^*} -g^*(\xi) \Longrightarrow -g^{**}(0) = \inf_{\xi \in X^*} g^*(\xi).$$

Se g^* fosse debolmente* s.c.i. e coerciva, avrebbe un minimo $\bar{\xi}$ e quindi avremmo ottenuto la tesi: esisterebbe infatti un insieme $K \subseteq X^*$ limitato, e quindi relativamente compatto con la topologia debole*, fuori dal quale $g^* \geq M > \inf_{\xi \in X^*} g^*(\xi)$; ma in generale una funzione semicontinua inferiormente ammette minimo sui compatti, per cui g^* avrebbe un minimo in K, che sarebbe chiaramente anche un minimo globale. Ora $g^*(\xi) = \sup_{x \in X} \{\langle x, \xi \rangle - g(x)\}$, e quindi g^* è per definizione estremo superiore di funzionali debolmente* continui, e quindi è debolmente* s.c.i.; a questo punto è sufficiente mostrare che, grazie all'ipotesi di limitatezza locale di g, si ha anche la coercività. Se supponiamo per assurdo che non valga, si può considerare una successione $\lambda_n \nearrow +\infty$ ed una $\xi_n \in X^*$ con $\|\xi_n\| = 1$ e $g^*(\lambda_n \xi_n) \leq M$. Siano poi $x_n \in X$ tali che $\|x_n\| = 1$ e $\langle \xi_n, x_n \rangle \geq 1 - 1/\lambda_n$. Allora

$$\begin{aligned} g\left(\frac{x_n}{\sqrt{\lambda_n}}\right) &\geq g^{**}\left(\frac{x_n}{\sqrt{\lambda_n}}\right) = \sup_{\xi \in X^*} \left\{\langle \frac{x_n}{\sqrt{\lambda_n}}, \xi \rangle - g^*(\xi)\right\} \\ &\geq \langle \frac{x_n}{\sqrt{\lambda_n}}, \lambda_n \xi_n \rangle - g^*(\lambda_n \xi_n) \geq \sqrt{\lambda_n} - \frac{1}{\sqrt{\lambda_n}} - M \longrightarrow +\infty. \end{aligned}$$

Ma questo non può succedere perché $x_n/\sqrt{\lambda_n} \to 0$ e per ipotesi g è limitata in un intorno di 0. Da questo segue la tesi.

■

Prima di poter mostrare il teorema di dualità ci serve ancora una definizione: dato l'operatore lineare $A : D(A) \longrightarrow Y$ con dominio $D(A)$ denso in X, definiamo

$$D(A^*) := \{\tilde{y} \in Y^* : \exists \tilde{x} \in X^*, \ \langle \tilde{y}, A(x) \rangle = \langle \tilde{x}, x \rangle \ \forall x \in D(A)\}; \tag{A.2}$$

per ogni $\tilde{y} \in D(A^*)$, vi è un unico $\tilde{x}$ secondo la definizione, dal momento che $D(A)$ è denso: si può allora definire l'operatore duale $A^* : D(A^*) \longrightarrow X^*$ che mandi $\tilde{y}$ in $\tilde{x}$. E' immediato notare che A^* è un operatore lineare, ed un'altra proprietà interessante è data dal

Lemma A.2 $\tilde{y} \in D(A^*) \Longleftrightarrow \exists K,\ \langle \tilde{y}, A(x) \rangle \leq K\|x\|\ \forall x \in D(A)$.

Dim : $\Longrightarrow$ è immediato: $\langle \tilde{y}, A(x) \rangle = \langle A^*(\tilde{y}), x \rangle \leq \|A^*(\tilde{y})\|\|x\|$. Per mostrare $\Longleftarrow$, a questo punto, si consideri un elemento $\tilde{y} \in Y^*$ per cui esiste una costante K tale che $\langle \tilde{y}, A(x) \rangle \leq K\|x\|$ per ogni $x \in D(A)$, e si definisca $\varphi : D(A) \longrightarrow \mathbb{R}$ come $\varphi(x) = \langle \tilde{y}, A(x) \rangle$; la funzione φ risulta lineare e, grazie all'ipotesi, continua; si vorrebbe allora estendere il dominio a tutto X. Per fare questo si scelga $x \in X$ ed una successione x_n contenuta in $D(A)$ che converge ad x: chiaramente $\varphi(x_n)$ risulta una successione di Cauchy reale e quindi tende ad un numero α, e lo stesso numero α corrisponde ad ogni possibile successione che converga ad x; definendo allora $\varphi(x) = \alpha$, si può estendere $\varphi : X \longrightarrow \mathbb{R}$ in modo che sia ancora lineare e continua. Ma allora $\varphi \in X^*$, e per definizione $\langle \varphi, x \rangle = \langle \tilde{y}, A(x) \rangle$ per ogni $x \in D(A)$, da cui segue che $\tilde{y} \in D(A^*)$ e $A^*(\tilde{y}) = \varphi$.

■

Fatte queste premesse, veniamo al teorema di dualità:

Teorema A.3 *Siano X e Y due spazi di Banach; siano inoltre date $g : X \longrightarrow \overline{\mathbb{R}}$ continua e convessa, $h : Y \longrightarrow \overline{\mathbb{R}}$ s.c.i. e convessa e $A : D(A) \longrightarrow Y$, un operatore lineare con dominio $D(A) \subseteq X$ denso. Se esiste $x_0 \in D(A)$ tale che $h(A(x_0) + y) + g(x_0)$ è limitata dall'alto in un intorno di $y = 0$, allora*

$$\inf_{x \in D(A)} \{h(A(x)) + g(x)\} = \sup_{\tilde{y} \in D(A^*)} \{-g^*(-A^*(\tilde{y})) - h^*(\tilde{y})\}.$$

Inoltre, se non è $-\infty$, il sup *è un massimo.*

Dim : Si definisca $k : Y \longrightarrow \overline{\mathbb{R}}$ tramite la

$$k(y) = \inf_{x \in D(A)} \{h(A(x) + y) + g(x)\},$$

in modo che il primo membro dell'eguaglianza da mostrare sia $k(0)$.

i) k è convessa.

Siano fissati y_1 e y_2 in Y e $\lambda \in (0, 1)$: dobbiamo mostrare che

$$k(\lambda y_1 + (1 - \lambda)y_2) \leq \lambda k(y_1) + (1 - \lambda)k(y_2).$$

Se $k(y_1)$ e $k(y_2)$ sono finiti, per definizione si possono scegliere x_1 e x_2 tali che

$$k(y_1) > h(A(x_1) + y_1) + g(x_1) - \epsilon \quad ; \quad k(y_2) > h(A(x_2) + y_2) + g(x_2) - \epsilon.$$

Allora risulta, grazie alla linearità di A e alla convessità di g ed h,

$$\begin{aligned}
k(\lambda y_1 + (1-\lambda)y_2) &\leq h\Big(A\big(\lambda x_1 + (1-\lambda)x_2\big) + \lambda y_1 + (1-\lambda)y_2\Big) + g(\lambda x_1 + (1-\lambda)x_2) \\
&\leq h\big(\lambda(A(x_1)+y_1) + (1-\lambda)(A(x_2)+y_2\big) + \lambda g(x_1) + (1-\lambda)g(x_2) \\
&\leq \lambda\big(h(A(x_1)+y_1) + g(x_1)\big) + (1-\lambda)\big(h(A(x_2)+y_2) + g(x_2)\big) \\
&\leq \lambda k(y_1) + (1-\lambda)k(y_2) + \epsilon,
\end{aligned}$$

da cui la disuguaglianza voluta, grazie alla genericità di ϵ, nel caso in cui $k(y_1)$ e $k(y_2)$ siano finiti. Se $k(y_2) = +\infty$ la diseguaglianza è vera a vuoto (si ricordi che $-\infty + \infty = +\infty$); se $k(y_1) = -\infty$, si può supporre allora $k(y_2) \leq C$; usando quindi un punto x_1 tale che

$$h(A(x_1)+y_1) + g(x_1) < -M,$$

con gli stessi conti di sopra si arriva a

$$k(\lambda y_1 + (1-\lambda)y_2) \leq -\lambda M + (1-\lambda)C + \epsilon,$$

e questo assicura $k(\lambda y_1 + (1-\lambda)y_2) = -\infty$. Da tutto ciò segue che la disuguaglianza vale in generale e quindi la convessità è provata.
ii) k è s.c.i. in 0.
La semicontinuità inferiore, per una funzione convessa, è automatica in ogni punto che abbia un intorno in cui la funzione è limitata dall'alto, come si nota facilmente direttamente dalla definizione, per cui basta mostrare che k è limitata dall'alto in un intorno di 0. Per ipotesi esiste un punto $x_0 \in D(A)$ tale che $h(A(x_0)+y) + g(x_0) \leq K$ in un intorno di $y = 0$ in Y; in quell'intorno si ha

$$k(y) \leq h(A(x_0)+y) + g(x_0) \leq K,$$

da cui la limitatezza locale dall'alto in 0 e quindi la semicontinuità inferiore di k in 0.
iii) A questo punto si può valutare k^* come segue:

$$\begin{aligned}
k^*(\tilde{y}) &= \sup_{y\in Y} \{\langle \tilde{y}, y\rangle - k(y)\} \\
&= \sup_{y} \Big\{\langle \tilde{y}, y\rangle - \inf_{x\in D(A)} \{h(A(x)+y) + g(x)\}\Big\} \\
&= \sup_{y} \sup_{x} \{\langle \tilde{y}, y\rangle - h(A(x)+y) - g(x)\} \\
&= \sup_{x} \sup_{y} \{\langle \tilde{y}, A(x)+y\rangle - h(A(x)+y) - g(x) - \langle \tilde{y}, A(x)\rangle\} \\
&= \sup_{x} \{h^*(\tilde{y}) - g(x) - \langle \tilde{y}, A(x)\rangle\}\,.
\end{aligned}$$

Per proseguire il conto bisogna distinguere due casi: se $\tilde{y} \in D(A^*)$, si può continuare con

$$\begin{aligned}
k^*(\tilde{y}) &= \cdots = \sup_{x\in D(A)} \{h^*(\tilde{y}) - g(x) - \langle \tilde{y}, A(x)\rangle\} \\
&= h^*(\tilde{y}) + \sup_{x\in D(A)} \{-g(x) + \langle -A^*(\tilde{y}), x\rangle\} = h^*(\tilde{y}) + \sup_{x\in X} \{\langle -A^*(\tilde{y}), x\rangle - g(x)\} \\
&= h^*(\tilde{y}) + g^*(-A^*(\tilde{y})).
\end{aligned}$$

Si noti che il passaggio in cui il sup su $x \in D(A)$ è divenuto un sup su $x \in X$ è lecito per la semicontinuità superiore di g e la densità di $D(A)$ in X.

Se invece $\tilde{y} \notin D(A^*)$, vogliamo mostrare che $k^*(\tilde{y}) = +\infty$: in effetti, a meno che non sia $g \equiv +\infty$ (e in tal caso la tesi è immediata), si può scegliere $\bar{x} \in X$ tale che $g(\bar{x}) < +\infty$; grazie al lemma A.2, inoltre, si può facilmente trovare una successione $x_n \underset{X}{\longrightarrow} \bar{x}$ contenuta in $D(A)$ tale che $\langle \tilde{y}, A(x_n)\rangle \to -\infty$. Si può allora proseguire il conto di prima con

$$\begin{aligned}
k^*(\tilde{y}) &= \cdots = \sup_{x\in D(A)} \{h^*(\tilde{y}) - g(x) - \langle \tilde{y}, A(x)\rangle\} \\
&\geq h^*(\tilde{y}) + \sup_{n} \{-g(x_n) - \langle \tilde{y}, A(x_n)\rangle\} = h^*(\tilde{y}) + \infty = +\infty,
\end{aligned}$$

ricordando ancora la convenzione $-\infty + \infty = +\infty$.

Calcolato finalmente k^*, è a questo punto immediato il fatto che

$$k^{**}(0) = \sup_{\tilde{y}\in Y^*} -k^*(\tilde{y}) = \sup_{\tilde{y}\in D(A^*)} -k^*(\tilde{y}) = \sup_{\tilde{y}\in D(A^*)} \{-h^*(\tilde{y}) - g^*(-A^*(\tilde{y}))\},$$

che coincide con il secondo membro dell'uguaglianza da mostrare; allora la prima parte della tesi è a questo punto $k(0) = k^{**}(0)$.

iv) Per notare che $k(0) = k^{**}(0)$ e che il secondo è un massimo purché sia diverso da $-\infty$, è ora sufficiente applicare il lemma A.1 a k; per poterlo fare bisogna verificare che, se $k(0) > -\infty$, k è limitata in un intorno di 0; ma la limitatezza dall'alto si è già notata prima, per mostrare che k è s.c.i., mentre quella dal basso è immediata, se $k(0) > -\infty$, proprio grazie alla semicontinuità inferiore in 0.

$\blacksquare$

A.2 Un'uguaglianza per funzioni convesse

In questa sezione mostreremo una comoda uguaglianza valida per funzioni convesse:

Proposizione A.4 *Se $f : X \longrightarrow \mathbb{R}$ è una funzione convessa e con crescita asintotica più che lineare, X è uno spazio di Hilbert e V è un suo sottospazio lineare, coincidono*

$$f_1(x) = \sup_{p \in V} \{p \cdot x - f^*(p)\} \qquad e \qquad f_2(x) = \inf_{r \in V^\perp} \{f(x + r)\}.$$

<u>Dim</u> : Dalla definizione si trova subito che

$$\begin{aligned} f_1(x) &= \sup_{p \in V} \Big\{ p \cdot x - \sup_{z \in X} \{p \cdot z - f(z)\} \Big\} = \sup_{p \in V} \inf_{z \in X} \{p \cdot (x - z) + f(z)\} \\ &= \sup_{p \in V} \inf_{z \in X} \{p \cdot (-z) + f(x + z)\} = \sup_{p \in V} \inf_{z \in X} \{p \cdot z + f(x + z)\}; \end{aligned}$$

conviene allora fissare x e ridefinire f tramite una traslazione di x del suo argomento, in modo da ricondurci a mostrare che

$$\sup_{p \in V} \inf_{z \in X} \{p \cdot z + f(z)\} = \inf_{r \in V^\perp} \{f(r)\}; \tag{A.3}$$

si noti anche che la traslazione non ha alterato la convessità né la crescita asintotica più che lineare di f. Cerchiamo di scrivere meglio il termine di sinistra: visto che il prodotto scalare $p \cdot z$ dipende solo dalla componente di z lungo V è comodo scomporre e definire, dato $q \in V$, $\overline{f}(q) = \inf_{r \in V^\perp} \{f(q + r)\}$; un banale conto assicura che anche $\overline{f} : V \longrightarrow \mathbb{R}$ è convessa e a crescita asintotica più che lineare. Inoltre si ha, fissato $p \in V$,

$$\begin{aligned} \inf_{z \in X} \{p \cdot z + f(z)\} &= \inf_{q \in V} \inf_{r \in V^\perp} \{p \cdot (q + r) + f(q + r)\} = \inf_{q \in V} \inf_{r \in V^\perp} \{p \cdot q + f(q + r)\} \\ &= \inf_{q \in V} \{p \cdot q + \overline{f}(q)\}, \end{aligned}$$

per cui il primo membro della (A.3) diventa

$$\sup_{p \in V} \inf_{q \in V} \{p \cdot q + \overline{f}(q)\} = \sup_{p \in V} \Big\{ - \sup_{q \in V} \{(-p) \cdot q - \overline{f}(q)\} \Big\} = \sup_{p \in V} \{-\overline{f}^*(-p)\} = \overline{f}^{**}(0),$$

ricordando che $\overline{f}$ è definita sul sottospazio V di X; d'altra parte il secondo membro della (A.3) è per definizione $\overline{f}(0)$ e quindi in definitiva basta provare $\overline{f}(0) = \overline{f}^{**}(0)$. Grazie al lemma A.1, allora, è sufficiente mostrare che $\overline{f}$ è s.c.i. in 0 e convessa; ma sappiamo già che $\overline{f}$ è convessa, e com'è noto la convessità assicura anche la continuità globale per funzioni a valori reali (mentre non è vero se sono possibili anche i valori $\pm\infty$).

$\blacksquare$

A.3 Un altro teorema di dualità

In questa sezione mostreremo un altro teorema di dualità; il risultato si potrebbe vedere come caso particolare del teorema A.3, ma le ipotesi che useremo sono più deboli, visto che non chiediamo la continuità di f né la densità del dominio $D(A)$: per comodità considereremo $f \circ A$ come esteso identicamente a $+\infty$ fuori dal dominio $D(A)$, in modo da poter scrivere $(f \circ A)^* : X^* \longrightarrow \mathbb{R}$; ci riferiremo poi ad A^* come ad un operatore che mandi $\eta \in Y^*$ in $\eta \circ A : D(A) \longrightarrow \mathbb{R}$ e scriveremo con un abuso di notazione $\langle A^*(\eta), x\rangle$ per indicare $A^*(\eta)(x)$, anche se $A^*(\eta)$ è lineare ma non necessariamente continuo: nelle ipotesi generali di questa sezione non è detto infatti che si possa dare una definizione dell'operatore A^* come abbiamo fatto nella sezione A.1; se $\xi \in X^*$, infine, scriveremo, con un altro abuso di notazione, $A^*(\eta) = \xi$ nel caso in cui la funzione $A^*(\eta)$ coincida con la restrizione di ξ a $D(A)$: si noti che, nel caso in cui esista $\xi \in X^*$ con $A^*(\eta) = \xi$, non è detto che sia unico. Stabilite queste notazioni, veniamo al

Teorema A.5 *Siano X ed Y due spazi di Banach, $A : D(A) \longrightarrow Y$ un operatore lineare e $f : Y \longrightarrow (-\infty, +\infty]$ convessa e continua in almeno un punto dell'immagine di $D(A)$. Allora*

$$(f \circ A)^*(\xi) = \inf_{\eta \in Y^*} \{f^*(\eta) : A^*(\eta) = \xi\} ;$$

se poi $(f \circ A)^*(\xi) < +\infty$, l' inf *è in realtà un minimo.*

Dim : Supponiamo per comodità che un punto di $D(A)$ su cui f è continua sia l'origine; fissato poi $\xi \in X^*$, definiamo $h : Y \longrightarrow \mathbb{R}$ come

$$h(y) := \inf_{x \in D(A)} \{f(A(x) + y) - \langle \xi, x\rangle\} ;$$

con due semplici conti notiamo ora che

$$(f \circ A)^*(\xi) = \sup_{x \in X} \{\langle \xi, x\rangle - f \circ A(x)\} = \sup_{x \in D(A)} \{\langle \xi, x\rangle - f \circ A(x)\} = -h(0), \tag{A.4}$$

$$h(y) \leq f(A(0) + y) \; - \langle \xi, 0\rangle \; = f(y). \tag{A.5}$$

Mostriamo adesso la convessità di h: in generale, se $\varphi : A \times B \longrightarrow \mathbb{R}$ è convessa, lo è banalmente anche $\tilde{\varphi} : B \longrightarrow \mathbb{R}$ definita come $\tilde{\varphi}(b) = \inf_{a \in A} \varphi(a, b)$; se allora definiamo $\varphi : D(A) \times Y \longrightarrow \mathbb{R}$ come $\varphi(x, y) = f(A(x) + y) - \langle \xi, x\rangle$, si ha che $h = \tilde{\varphi}$ e quindi la convessità di h segue da quella di φ, che d'altra parte discende immediatamente dalla convessità di f e di $\langle \cdot, \xi\rangle$.

Vediamo infine come esprimere convenientemente h^*:

$$\begin{aligned}
h^*(\eta) &= \sup_{y\in Y}\{\langle\eta,y\rangle - h(y)\} = \sup_y \Big\{\langle\eta,y\rangle - \inf_{x\in D(A)}\{f(A(x)+y)-\langle\xi,x\rangle\}\Big\}\\
&= \sup_y\sup_x\{\langle\eta,y\rangle - f(A(x)+y)+\langle\xi,x\rangle\}\\
&= \sup_x\Big\{\langle\xi,x\rangle + \sup_y\{\langle\eta,y\rangle - f(A(x)+y)\}\Big\}\\
&= \sup_x\Big\{\langle\xi,x\rangle - \langle\eta,A(x)\rangle + \sup_y\{\langle\eta,A(x)+y\rangle - f(A(x)+y)\}\Big\}\\
&= \sup_x\{\langle\xi-A^*(\eta),x\rangle\} + f^*(\eta) = \begin{cases} f^*(\eta) & \text{se } A^*(\eta)=\xi\\ +\infty & \text{altrimenti}\end{cases}
\end{aligned}$$

Da questo si ottiene subito

$$h^{**}(0) = -\inf_{\eta\in Y^*}\{f^*(\eta) : A^*(\eta)=\xi\}\,,$$

per cui grazie a (A.4) per avere la tesi basta mostrare che $-h(0) = -h^{**}(0)$ e che si può scrivere $h^{**}(0) = -h^*(\eta)$ purché sia $h(0) > -\infty$; visto che abbiamo già mostrato la convessità di h, usando il lemma A.1 è sufficiente mostrare che h è s.c.i. in 0 ed è limitata in un intorno di 0 se $h(0) > -\infty$; ma visto che grazie a (A.5) $h \le f$ e che f è continua in 0, h risulta limitata dall'alto in un intorno di 0 e questo per una funzione convessa, come già notato, assicura la semicontinuità inferiore. Se poi $h(0) > -\infty$, per la semicontinuità inferiore si trova un intorno in cui h è limitata dal basso, e d'altra parte abbiamo già notato la limitatezza dall'alto.

■

A.4 Commutatività tra sup ed integrale

In questa sezione enunceremo un teorema che assicura, sotto opportune ipotesi, l'uguaglianza tra l'integrale dell'estremo superiore di una famiglia di funzioni e l'estremo superiore degli integrali delle stesse funzioni; in questa sezione indicheremo con Ω uno spazio generico dotato di una misura positiva e finita $\mu \in \mathcal{M}^+(\Omega)$. Innanzitutto ci servono due definizioni: un *integrando normale* è una funzione misurabile $f : \Omega \times \mathbb{R}^m \longrightarrow \overline{\mathbb{R}}$, dove la misurabilità si intende secondo la misura $\mu \otimes \mathcal{H}^n$; un integrando normale f viene detto *concavo* se per ogni x $f(x, \cdot)$ è concavo e semicontinuo superiore. Un sottoinsieme K delle funzioni misurabili da Ω ad $\mathbb{R}^m$ viene infine detto *stabile* se presa una qualsiasi partizione dell'unità $\varphi_0,\ \varphi_1, \ldots,\ \varphi_n$ fatta di funzioni continue che abbiano tutte tranne al più una supporto compatto, e scelte arbitrariamente $f_0,\ f_1, \ldots,\ f_n \in K$, si abbia $\varphi_0 f_0 + \cdots + \varphi_n f_n \in K$; nel caso in cui Ω sia aperto in un R^N si parla di stabilità quando la condizione suddetta sia verificata anche solo nel caso in cui le φ_i siano in $C^\infty(\overline{\Omega})$. Si può allora enunciare il seguente

Teorema A.6 *Sia $j : \Omega \times R^m \longrightarrow [-\infty, +\infty)$ un integrando normale concavo e K un sottoinsieme stabile delle funzioni misurabili a valori in $\mathbb{R}^m$. Indicando con K_0 il sottoinsieme di K formato dalle funzioni u tali che $\int j(x, u(x))\ d\mu$ è finito e ponendo $\Gamma = (\mu - \mathrm{ess}) \bigcup_{u \in K_0} u(\cdot)$, si ha*

$$\sup_{u \in K_0} \int_\Omega j(x, u(x))\ d\mu = \int_\Omega \sup_{z \in \Gamma(x)} j(x, z)\ d\mu.$$

Si tratta del Teorema 1 in [7]; per i nostri scopi, in realtà, sarà sufficiente una formulazione molto più debole, il

Corollario A.7 *Sia $\Omega \subseteq \mathbb{R}^N$ aperto, $j : \Omega \times R^m \longrightarrow [-\infty, +\infty)$ un integrando normale concavo e K un sottospazio vettoriale delle funzioni misurabili a valori in R^m tale che per ogni $u \in K$ si abbia $u\psi \in K$ per ogni $\psi \in C^\infty(\overline{\Omega}) \cap L^\infty$ e $\int j(x, u(x))\ d\mu$ finito. Allora, posto $\Gamma = (\mu - \mathrm{ess}) \bigcup_{u \in K} u(\cdot)$, si ha*

$$\sup_{u \in K} \int_\Omega j(x, u(x))\ d\mu = \int_\Omega \sup_{z \in \Gamma(x)} j(x, z)\ d\mu.$$

<ins>Dim</ins> : Evidentemente la richiesta che $u\psi$ appartenga a K per ogni $u \in K$ e $\psi \in C^\infty(\overline{\Omega}) \cap L^\infty$ è più forte della stabilità, visto che una partizione dell'unità è compresa tra 0 ed 1 e che K è chiuso rispetto alla somma; inoltre il fatto che $\int j(x, u(x))\ d\mu$ sia finito per ogni $u \in K$ assicura $K = K_0$, per cui la tesi segue immediatamente dal teorema appena enunciato.

$\blacksquare$

Appendice B

Sup ed unione $\mu-$essenziali

Consideriamo uno spazio $\Omega \subset\subset \mathbb{R}^m$ dotato di una misura positiva e finita μ, e sia dato un insieme di multifunzioni $\{\Gamma_i : i \in I\}$ a valori in un generico spazio topologico E; in molte situazioni può risultare comodo considerare l'unione delle immagini delle multifunzioni, ma si può notare facilmente che non appena la cardinalità di I supera quella numerabile vi sono effetti indesiderati, dovuti al fatto che un'unione arbitraria di insiemi trascurabili per μ può non essere trascurabile. Un semplice esempio che illustra la situazione può essere il seguente: sia $\Omega = [0,1]$, $E = \mathbb{R}^+$ e μ la misura di Lebesgue; consideriamo l'insieme dei sottografici delle funzioni positive aventi norma L^∞ minore di 1: l'intuizione pretenderebbe che l'unione di questi fosse $[0,1]\times[0,1]$, mentre evidentemente è $[0,1]\times\mathbb{R}^+$, visto che per ogni punto vi sono funzioni che, pur avendo norma L^∞ minore di 1, hanno un valore arbitrariamente alto; un comportamento di questo genere dell'unione è inaccettabile, e rende l'idea dell'utilità di una diversa nozione di unione che tenga opportunamente conto della misura.

Vi è una soluzione accettabile al problema dal momento che, dato un insieme $\{\Gamma_i : i \in I\}$ di multifunzioni a valori chiusi, esiste una multifunzione Γ che rappresenta il minimo –con la relazione d'ordine data dall'inclusione $\mu-$quasi ovunque– delle multifunzioni che sono maggiori di ciascuna delle Γ_i; in pratica Γ ha le due fondamentali proprietà seguenti:

$$\begin{aligned} &\textbf{a)}\ \text{Per ogni indice } i,\ \Gamma_i(x) \subseteq \Gamma(x) \text{ per } \mu-\text{q.o. } x. \\ &\textbf{b)}\ \text{Se per una multifunzione } \widetilde{\Gamma} \text{ vale a), allora } \Gamma(x) \subseteq \widetilde{\Gamma}(x) \text{ per } \mu-\text{q.o. } x. \end{aligned} \tag{B.1}$$

Questa multifunzione Γ è ciò che verrà chiamato unione $\mu-$essenziale delle Γ_i, e grazie alla proprietà b) appena enunciata si vede che si tratta di una buona definizione, visto che Γ risulta fissata almeno $\mu-$quasi ovunque. Questa nozione è stata introdotta da Valadier nel terzo capitolo di [28], dove è considerato prima il caso di multifunzioni a valori convessi compatti e poi la generalizzazione per valori chiusi.

Vediamo ora brevemente un modo leggermente diverso, anche se equivalente, di effettuare la costruzione enunciata, che consiste nell'introdurre prima la nozione di sup essenziale di funzioni a valori reali, poi quella di sup essenziale di funzioni a valori nelle funzioni lipschitziane e infine quella di unione essenziale: ciascuna di queste costruzioni si appoggerà su quelle precedenti e sarà caratterizzata da proprietà analoghe alla (B.1). Iniziamo con le funzioni a valori reali: siano dunque date $f_i : \Omega \longrightarrow [0,1]$, $i \in I$, misurabili, e si consideri il problema

$$\sup \left\{ \int_\Omega f_J \, d\mu : J \subseteq I, \ \#(J) \leq \#(\mathbb{N}) \right\}, \tag{B.2}$$

dove si intende $f_J = \sup_{i \in J} f_i$; si può notare subito come questo sup sia in realtà un massimo: se infatti si considera una successione J_i tale che $\int f_{J_i} \, d\mu \to \sup \int f_J \, d\mu$, è chiaro che $\bar{J} = \bigcup_i J_i$ è ancora numerabile e realizza il sup; definiamo allora sup $\mu-$essenziale delle f_i la $(\mu-\text{ess}) \sup_{i \in I} f_i = f_{\bar{J}}$, e notiamo che vengono facilmente soddisfatte le seguenti proprietà:

$$\begin{aligned} &\textbf{a)} \text{ Per ogni indice } i, \ f_{\bar{J}} \geq f_i \ \mu\text{−quasi ovunque.} \\ &\textbf{b)} \text{ Se per } g \text{ vale a), allora } f_{\bar{J}} \leq g \ \mu\text{−quasi ovunque.} \end{aligned} \tag{B.3}$$

Con riferimento all'esempio dato prima che mostrava l'inopportunità di fare l'unione effettiva, è confortante notare che se le funzioni f_i sono tutte le funzioni misurabili positive minori di 1 quasi ovunque, allora il sup $\mu-$essenziale è dato dalla funzione costantemente uguale ad 1.

Fatto questo primo passo, si può passare a definire il sup essenziale di funzioni a valori nelle funzioni lipschitziane: siano quindi date delle $f_i : \Omega \longrightarrow \mathrm{Lip}_1(\mathbb{R}^n, [0,1])$, misurabili, che ad ogni $x \in \Omega$ associano la funzione $f_{ix}(y)$, $1-$lipschitziana nella variabile y; indicati poi con φ_h gli elementi di un sottoinsieme di $C_0(\mathbb{R}^n)$ numerabile e denso con la norma del sup e detta $\lambda_h < +\infty$ la misura del supporto di φ_h secondo Lebesgue, si definiscano le coordinate $g_{i,h}(x) = \int f_{ix}(y)\varphi_h(y) \, dy$; dal momento che, fissato h, le $g_{i,h}$ sono funzioni da Ω a $[-\lambda, \lambda]$ si può considerare, grazie a quanto appena visto,

$$(\mu - \text{ess}) \sup_{i \in I} g_{i,h} = \sup_{i \in I_h} g_{i,h}, \tag{B.4}$$

con I_h numerabile; è allora numerabile $\bar{I} = \bigcup I_h$, per cui si può definire

$$(\mu - \text{ess}) \sup_{i \in I} f_i := \sup_{i \in \bar{I}} f_i; \tag{B.5}$$

analogamente a prima, indicando per brevità con f il sup $\mu-$essenziale delle f_i si mostrano immediatamente le proprietà:

$$\begin{array}{l} \textbf{a)} \text{ Per ogni indice } i,\ f_x \geq f_{ix} \text{ per } \mu\text{−quasi ogni } x. \\ \textbf{b)} \text{ Se per } g \text{ vale a), allora } f_x \leq g_x \text{ per } \mu\text{−quasi ogni } x. \end{array} \tag{B.6}$$

A questo punto si può finalmente giungere alla nozione di unione essenziale: date delle multifunzioni $\Gamma_i : \Omega \longrightarrow \mathcal{F}(\mathbb{R}^n)$, dove con $\mathcal{F}$ si indica la classe dei chiusi, si definisca $f_{ix} = d(\cdot, \Gamma_i(x)) \wedge 1$: f_i, allora, associa ad ogni punto una funzione $1-$lipschitziana a valori in $[0, 1]$; si ha allora, per quanto già visto,

$$\left((\mu - \text{ess}) \inf_{i \in I} f_i\right)_x = \inf_{i \in \bar{I}} f_{ix} = d(\cdot, \bigcup_{i \in \bar{I}} \Gamma_i(x)), \tag{B.7}$$

per cui ha senso definire l'unione $\mu-$essenziale come l'unione delle immagini delle Γ_i sugli indici appartenenti a $\bar{I}$; detta quindi $\Gamma = (\mu - \text{ess}) \bigcup_{i \in I} \Gamma_i$, si hanno, in analogia con i casi precedenti, le proprietà:

$$\begin{array}{l} \textbf{a)} \text{ Per ogni indice } i,\ \Gamma_i(x) \subseteq \Gamma(x) \text{ per } \mu\text{−quasi ogni } x. \\ \textbf{b)} \text{ Se per } \psi \text{ vale a), allora } \Gamma(x) \subseteq \psi(x) \text{ per } \mu\text{−quasi ogni } x. \end{array} \tag{B.8}$$

Una proprietà che può essere molto comoda dell'unione essenziale, così come del sup essenziale, è che è effettivamente l'unione (o il sup) relativo ad un sottoinsieme numerabile; per concludere, notiamo che si può costruire una multifunzione anche come unione essenziale di funzioni, visto che una funzione è una multifunzione che ammette come valori dei punti, che non sono altro che dei particolari chiusi. Ad esempio la definizione

$$T^p_\mu := (\mu - \text{ess}) \bigcup \left\{\varphi(\cdot) : \varphi \in X^{p'}_\mu\right\}; \tag{2.6}$$

data nella sezione 2.2 è consistente con quanto fatto finora, e le proprietà date in (2.7) sono esattamente quelle in (B.8) scritte per questo caso particolare.

Bibliografia

[1] L. Ambrosio, N. Fusco & D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford University Press (2000).

[2] P. Appell, Le probleme geometrique des déblais et remblais, Memor. des Sciences Math., Acad. de Sciences de Paris, Gauthier–Villars, **27** (1928), 1–34.

[3] J.-D. Benamou & Y. Brenier, A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem, Numer. Math., **84** (2000), 375–393.

[4] G. Bouchitté & G. Buttazzo, Characterization of optimal shapes and masses through Monge–Kantorovich equation, preprint Università di Pisa (2000).

[5] G. Bouchitté, G. Buttazzo & P. Seppecher, Energies with respect to a measure and applications to low dimensional structures, Calc. Var., **5** (1997), 37–54.

[6] G. Bouchitté, G. Buttazzo & P. Seppecher, Shape Optimization Solutions via Monge–Kantorovich Equation, C. R. Acad. Sci. Paris, **324–I** (1997), 1185–1191.

[7] G. Bouchitté & M. Valadier, Integral representation of convex functionals on a space of measures, Journal of Functional Analysis, **50** (1988), 398–420.

[8] Y. Brenier, A Homogenized Model for Vortex Sheets, Arch. Rat. Mech. Anal., **138** (1997), 319–353.

[9] Y. Brenier, A Monge–Kantorovich approach to the Maxwell equations, in preparation.

[10] C. Castaing & M. Valadier, Convex Analysis and Mesaurable Multifunctions, Springer–Verlag (1977).

[11] I. Ekeland & R. Temam, Analyse Convexe et Problèmes Variationnels, Dunod Gauthier–Villars (1976).

[12] L.C. Evans & W. Gangbo, Differential Equations Methods for the Monge–Kantorovich Mass Transfer Problem, Memoirs of the A.M.S., Vol. **137** (1999).

[13] I. Fragalà & C. Mantegazza, On some notions of tangent space to a measure, Proc. Roy. Soc. Edinmurgh, **129A** (1999), 331–342.

[14] W. Gangbo & R.J. McCann, The geometry of optimal transportation, Acta Math., **177** (1996), 113–161.

[15] L.V. Kantorovich, On the transfer of masses, Dokl. Akad. Nauk. SSSR **37** (1942), 227–229.

[16] L.V. Kantorovich, On a problem of Monge, Uspekhi Mat. Nauk. **3** (1948), 225–226.

[17] H.G. Kellerer, Duality theorems for marginal problems, Z. Wahrsch. verw. Gebiete, **67** (1984), 399–432.

[18] G. Monge, Memoire sur la Theorie des Déblais et des Remblais, Hist. de l'Acad. des Sciences de Paris (1781).

[19] F. Poupaud & M. Rascle, Measure Solutions to the Linear Multi-dimensional Transport Equation with non-smooth Coefficients, Comm. PdE, **22** (1997), 301–324.

[20] D. Preiss, Geometry of measures on $\mathbb{R}^n$: distribution, rectifiability and densities, Ann. Math., **125** (1987), 573–643.

[21] S.T. Rachev & L. Rüschendorf, Mass Transportation Problems, Springer–Verlag (1998).

[22] T. Rockafellar, Characterization of the subdifferential of convex functions, Pacific Journal of Math **17** (1966), 497–510.

[23] L. Rüschendorf, On $c-$optimal random variables, Statistics & Probability Letters, **27** (1996), 267–270.

[24] L. Rüschendorf, Frechet bounds and their applications, in: Advances in Probability Distributions with Given Marginals (G. Dall'Aglio e al., eds.), 151–187, Math. Appl., **67**. Kluwer Acad. Publ., Dordrecht (1991).

[25] L. Simon, Lectures on Geometric Measure Theory, Proceedings of the Centre for Mathematical Analysis, Australian National University, Volume **3** (1983).

[26] C. Smith & M. Knott, On Hoeffding–Frechet bounds and cyclice monotone relations, J. Multivariate Anal., **40** (1992), 328–334.

[27] V.N. Sudakov, Geometric Problems in the Theory of Infinite–Dimensional Probability Distributions, Proc. of the Steklov Institute of Mathematics, **141** (1979).

[28] M. Valadier, Multi-applications mesurables a valeurs convexes compactes, Journal de Mathématiques Pures et Appliquées, **50** (1971), 265–297.