Università degli Studi di Padova

Dipartimento di Matematica

Corso di Laurea Triennale in
Matematica

Esistenza di minimi per il problema di Kantorovich

Relatore:
Prof. Roberto Monti

Laureando:
Davide Tonello
Matricola 2042170

Anno Accademico 2023/2024

Indice

Introduzione

I problemi di ottimizzazione sono da sempre fondamentali per l'analisi matematica. Tra questi spicca per numero di applicazioni e di progressi recenti il problema del trasporto ottimo, focalizzato sulla minimizzazione dei costi associati allo spostamento di oggetti. La prima formalizzazione teorica fu proposta da Gaspard Monge nel 1871 e fu successivamente riformulata da Leonid Kantorovich nel 1939. Questa tesi tratterà il problema di Kantorovich discutendo i risultati necessari per la corretta descrizione del problema e per la dimostrazione dell'esistenza di una soluzione ottima.

La trattazione dell'argomento seguirà lo schema sottostante:

1. **Descrizione del problema.** Si introdurranno le definizioni e la notazione necessarie per la formulazione del problema e si enunceranno le condizioni sufficienti per l'esistenza di una soluzione ottima.

2. **Risultati utili.** Si descriveranno i teoremi di anlisi reale e di analisi funzionale necessari per la successive dimostrazioni.

3. **Esistenza di soluzioni ottime.** Si affronterà la dimostrazione del teorema sulle condizioni sufficienti per l'esistenza di minimi.

Capitolo 1

Descrizione del problema

1.1 Notazione

Introduciamo preliminarmente la notazione necessaria.

- $\mathbb{N}$ indica l'insieme dei numeri naturali senza 0, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.
- $\mathbb{R}$ indica l'insieme dei numeri reali e $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$.
- Dato E spazio vettoriale su $\mathbb{R}$, $\|\cdot\|_E$ indica una norma su E.
- Dato $(E, \|\cdot\|_E)$ spazio normato E' indica il suo duale topologico, ossia l'insieme dei funzionali lineari e continui da E ad $\mathbb{R}$.
- Dato (X, τ) spazio topologico $C(X)$ indica l'insieme delle funzioni continue da X ad $\mathbb{R}$.
- Data $f : X \to \overline{\mathbb{R}}$, $\|f\|_u = \sup_{x \in X} |f(x)|$ è la norma della convergenza uniforme.
- $C_b(X) = \{f \in C(X) \mid \|f\|_u < +\infty\}$ è l'insieme delle funzioni continue e limitate.
- $C_0(X) = \{f \in C(X), \mid \forall \epsilon > 0 \, \exists K \subseteq X compatto : |f(x)| < \epsilon \, \forall x \in X \setminus K\}$ è l'insieme delle funzioni continue e "nulle all'infinito".
- Dato (X, d) spazio metrico $B_X(x, r)$ è la palla di centro x e raggio r e $B_X(x, r] = B_X(x, r) \cup \partial B_X(x, r)$ è la corrispondente chiusura.
- Dato (X, τ) spazio topologico $\mathcal{B}(X)$ è la σ-algebra dei boreliani su X.

1.2 Spazi polacchi

Il contesto in cui costruiremo la teoria del trasporto ottimo è quello degli spazi polacchi:

Definizione 1.1. (X, τ) spazio topologico si dice *spazio polacco* se esiste d metrica su X che induce τ tale che lo spazio metrico (X, d) sia completo e separabile.

Il motivo per cui non consideriamo semplicemente spazi metrici è la volontà di mettere in risalto l'aspetto topologico della teoria. Dato uno spazio topologico metrizzabile infatti in generale la metrica che produce la topologia non è unica.

Inoltre a volte sarà utile considerare topologie diverse sugli stessi spazi a seconda delle proprietà che vorremo far emergere. Gli spazi polacchi da questo punto di vista sono molto più flessibili e dunque più adatti allo scopo.

Dato che lavoreremo con funzioni integrali è necessario collegare il contesto topologico con la teoria della misura (su cui si basa la moderna teoria dell'integrazione). Il modo più semplice per farlo su uno spazio polacco (X, τ) è costruire lo spazio misurabile con la σ-algebra generata da τ ovvero la σ-algebra di Borel $\mathcal{B}(X)$.

Definizione 1.2. Sia (X, τ) spazio polacco. $\mathcal{M}(X)$ è l'insieme delle misure con segno di Borel finite su X.
$\mathcal{M}_+(X) = \{\mu \in \mathcal{M}(X) \mid \mu \geq 0\}$ è l'insieme delle misure (positive) di Borel finite su X.
$\mathcal{P}(X) = \{\mu \in \mathcal{M}_+(X) \mid \mu(X) = 1\}$ è l'insieme delle misure di probabilità di Borel su X.

Le misure di Borel finite sugli spazi polacchi hanno una regolarità simile alle misure di Borel su $\mathbb{R}^n$, in particolare:

Teorema 1.1. *Sia (X, τ) spazio polacco e sia $\mu \in \mathcal{M}_+(X)$. Allora μ è regolare dall'interno ovvero:*

$$\mu(A) = \sup \{\mu(K) \mid K \subseteq X compatto, K \subseteq A\} \, \forall A \in \mathcal{B}(X)$$

e μ è regolare dall'esterno ovvero:

$$\mu(A) = \inf \{\mu(U) \mid U \subseteq X aperto, A \subseteq U\} \, \forall A \in \mathcal{B}(X).$$

Dato che la dimostrazione del teorema è abbastanza lunga la divideremo in tre parti:

Lemma 1.2. *Sia (X, τ) spazio polacco e sia $\mu \in \mathcal{M}_+(X)$. Allora per ogni $E \in \mathcal{B}(X)$ e per ogni $\epsilon > 0$ esistono $U \subseteq X$ aperto e $C \subseteq X$ chiuso tali che $C \subseteq E \subseteq U$ e $\mu(U \setminus C) < \epsilon$.*

Dimostrazione. Sia d una metrica su X che induce τ e sia $\epsilon > 0$ fissato.
Sia $C \subseteq X$ chiuso. Dimostriamo che esiste $U \subseteq X$ aperto tale che $C \subseteq U$ e $\mu(U \setminus C) < \epsilon$.
Per ogni $n \in \mathbb{N}$ sia $U_n = \left\{x \in X \mid d(x, C) < \frac{1}{2^n}\right\}$. Abbiamo che per ogni $x \in U_n$ esiste $r_x > 0$ tale che $d(x, C) + r_x < \frac{1}{2^n}$. Allora $B\left(x, r_x\right) \subseteq U_n \, \forall \, x \in U_n$ e dunque U_n è aperto per ogni $n \in \mathbb{N}$.
Ora chiaramente $C = \bigcap_{n \in \mathbb{N}} U_n$ e $U_{n+1} \subset U_n \forall n \in \mathbb{N}$. μ è finita quindi per continuità dall'alto $\mu(C) = \lim_{n\to\infty} \mu(U_n)$. Allora esiste $n^* \in \mathbb{N}$ tale che $\mu(U_{n^*} \setminus C) < \epsilon$ come si voleva.

Dimostriamo ora che la famiglia Σ dei sottoinsiemi di X per cui vale la tesi è una σ-algebra che contiene $\mathcal{B}(X)$.
Sia $A \in \Sigma$ e siano U aperto e C chiuso tali che $C \subseteq A \subseteq U$ e $\mu(U \setminus C) < \epsilon$. Siano $\widetilde{C} = X \setminus U$ e $\widetilde{U} = X \setminus C$. $\mu(\widetilde{U} \setminus \widetilde{C}) = \mu(X) - \mu(C) - \mu(X) + \mu(U) = \mu(U \setminus C) < \epsilon$ e $\widetilde{C} \subseteq A^{\mathsf{c}} \subseteq \widetilde{U}$, allora $A^{\mathsf{c}} \in \Sigma$.
Sia $\{A_n \mid n \in \mathbb{N}\} \subseteq \Sigma$. Per ogni $n \in \mathbb{N}$ siano U_n aperto e C_n chiuso tali che $C_n \subseteq A_n \subseteq U_n$ e $\mu(U_n \setminus C_n) < \frac{\epsilon}{2^n}$. Siano $A = \bigcup_{n\in\mathbb{N}} A_n$, $U = \bigcup_{n\in\mathbb{N}} U_n$ e $C = \bigcup_{n\in\mathbb{N}} C_n$. Allora $C \subseteq A \subseteq U$ e $\mu(U \setminus C) = \mu(\bigcup_{n\in\mathbb{N}} \left\{U_n \cap \bigcap_{k\in\mathbb{N}} {C_k}^{\mathsf{c}}\right\}) \leq \mu(\bigcup_{n\in\mathbb{N}} \{U_n \cap {C_n}^{\mathsf{c}}\}) \leq \sum_{n\in\mathbb{N}} \mu(U_n \setminus C_n) < \sum_{n\in\mathbb{N}} \frac{\epsilon}{2^n} = \epsilon$ e dunque $A \in \Sigma$.
Ciò dimostra che Σ è una σ-algebra. Ora per quanto visto prima $C \in \Sigma$ per ogni C chiuso e perciò $\mathcal{B}(X) = \sigma(\{C \subseteq X \mid C \, chiuso\}) \subseteq \Sigma$. □

Lemma 1.3. ***Ulam***
Sia (X, τ) spazio polacco e sia $\mu \in \mathcal{M}_+(X)$. Allora per ogni $\epsilon > 0$ esiste $K \subseteq X$ compatto tale che $\mu(X \setminus K) < \epsilon$.

Dimostrazione. (Basata su quella proposta in [1]).
Sia d metrica su X che induce τ tale che (X, d) sia completo e separabile. Sia $\epsilon > 0$ fissato.
Sia $\{x_n \mid n \in \mathbb{N}\}$ sottoinsieme denso e numerabile di X. Per ogni $k \in \mathbb{N}$ abbiamo che $X = \bigcup_{i\in\mathbb{N}} B\left(x_i, \frac{1}{2^k}\right]$. Per continuità dal basso $\mu(X) = \lim_{n\to\infty} \mu(\bigcup_{i\in\mathbb{N}}^{n} B\left(x_i, \frac{1}{2^k}\right])$.

Dato che $\mu(X) < \infty$, per ogni $k \in \mathbb{N}$ esiste $n(k) \in \mathbb{N}$ tale che $\mu(\bigcup_{i\in\mathbb{N}}^{n(k)} B\left(x_i, \frac{1}{2^k}\right]) > \mu(X) - \frac{\epsilon}{2^k}$.
Definiamo $K_{n(k)} = \bigcup_{i\in\mathbb{N}}^{n(k)} B\left(x_i, \frac{1}{2^k}\right]$ e $K = \bigcap_{k\in\mathbb{N}} K_{n(k)}$.
$K_{n(k)}$ è unione finita di chiusi, allora K è chiuso perché intersezione di chiusi. X è completo e perciò anche K lo è.
Per ogni $r > 0$ esiste $k \in \mathbb{N}$ tale che $\frac{1}{2^k} < r$, allora $K \subseteq K_{n(k)} = \bigcup_{i\in\mathbb{N}}^{n(k)} B\left(x_i, \frac{1}{2^k}\right] \subseteq \bigcup_{i\in\mathbb{N}}^{n(k)} B\left(x_i, r\right)$.
Dunque K è compatto perché completo e totalmente limitato (nel Capitolo 2 è presente una dimostrazione di questo risultato nell'ambito del Teorema 2.5).
Ora $\mu(X \smallsetminus K) = \mu(\bigcup_{k\in\mathbb{N}} X \smallsetminus K_{n(k)}) \leq \sum_{k\in\mathbb{N}} \mu(X \smallsetminus \bigcup_{i\in\mathbb{N}}^{n(k)} B\left(x_i, \frac{1}{2^k}\right]) < \sum_{k\in\mathbb{N}} \frac{\epsilon}{2^k} = \epsilon$ come si voleva. □

Concludiamo la dimostrazione del Teorema 1.1:

Dimostrazione. Dal Lemma 1.2 sappiamo che se $A \in \mathcal{B}(X)$, per ogni $\epsilon > 0$ esiste $U \subseteq X$ aperto tale che $A \subseteq U$ e $\mu(U \smallsetminus A) < \epsilon$, dunque vale la regolarità dall'esterno.
Inoltre, sempre grazie al Lemma 1.2, esiste $C \subseteq A$ chiuso tale che $\mu(A \smallsetminus C) < \epsilon$.
Basta dimostrare quindi che esiste $K \subseteq C$ compatto tale che $\mu(C \smallsetminus K) < \epsilon$.
Per il Lemma di Ulam 1.3 esiste $H \subseteq X$ compatto tale che $\mu(X \smallsetminus H) < \epsilon$. Sia $K = C \cap H$. K è compatto perché sottoinsieme chiuso di H compatto e $K \subseteq C$ per definiozione. Inoltre $\mu(C \smallsetminus K) = \mu(C \cap H^{\mathrm{c}}) \leq \mu(H^{\mathrm{c}}) < \epsilon$. □

In un contesto più generale $\mathcal{M}_+(X)$ indica l'insieme delle misure di Radon, ovvero l'insieme delle misure localmente finite, regolari dall'esterno e regolari dall'interno sugli aperti (si veda ad esempio [3]). Tuttavia dato che la maggior parte delle volte lavoreremo con misure di probabilità possiamo permetteci di ignorare questa distinzione.

Il Teorema 1.1 dimostra che ogni spazio polacco è anche spazio di Radon (in cui ogni misura di probabilità di Borel è misura di Radon). La trattazione di questo argomento è troppo vasta per lo scopo della tesi, ci basteranno solo alcuni teoremi che enunceremo nel Capitolo 2. L'argomento è affrontato in maniera esaustiva nelle fonti, si consulti [5].

1.3 Problema di Kantorovich

Immaginiamo che un venditore porta a porta possieda N magazzini e un numero potenzialmente molto grande di clienti distribuiti equamente in una città. Supponiamo che il costo del trasporto dei beni dai magazzini ai clienti dipenda per la maggior parte dal tragitto (supponiamo cioè che i fattori indipendenti dal trasporto in sé siano trascurabili). Il venditore vorrebbe ottimizzare il processo in modo da minimizzare tali costi.

Cercando di descrivere matematicamente il problema il venditore si accorge immediatamente di una criticità: come si può esprimere correttamente il fatto che all'inizio i beni siano "localizzati" in N punti mentre alla fine essi siano distribuiti sul territorio?

Sorge poi un'altra domanda: se il numero di clienti è molto elevato esiste un modo per minimizzare i costi totali senza dover ricorrere allo studio di ogni singola spedizione?

Il primo quesito trova risposta nella teoria della misura, in quanto rispetto alle misure naturali per il contesto (misura di Lebsgue su $\mathbb{R}^n$, misura di Hausdorff sugli spazi metrici) esistono sia misure "localizzate" (mutualmente singolari) sia misure "uniformi" (assolutamente continue).

Nel nostro esempio, supponendo che la mappa della città sia descritta da un sottoinsieme limitato di $\mathbb{R}^2$ con topologia euclidea, la misura di partenza potrebbe essere una combinazione di N Dirac, mentre la misura di arrivo potrebbe essere la legge di una distribuzione uniforme (entrambe opportunamente normalizzate).

Assumendo quindi che la massa totale sia 1 (a meno di normalizzazioni) dato $(X, \mathcal{B}(X), \mu)$ spazio di probabilità "di partenza" e $(Y, \mathcal{B}(Y), \nu)$ spazio di probabilità "di arrivo" vorremmo che il trasporto su $(X \times Y, \mathcal{B}(X \times Y))$ fosse descritto in modo da rispettare delle opportune "condizioni marginali" dettate da μ su X e da ν su Y.

Definizione 1.3. Siano X e Y spazi polacchi. Siano $\mu \in \mathcal{P}(X)$ e $\nu \in \mathcal{P}(Y)$. L'insieme dei *piani di trasporto* $\Pi(\mu, \nu)$ è definito da:

$$\Pi(\mu,\nu) = \{\pi \in \mathcal{P}(X \times Y) \mid \pi(A \times Y) = \mu(A)\,\forall\, A \in \mathcal{B}(X), \\ \pi(X \times B) = \nu(B)\,\forall B \, \in \mathcal{B}(Y)\}.$$

In questo modello dati $A \in \mathcal{B}(X)$ e $B \in \mathcal{B}(Y)$ il valore $\pi(A \times Y)$ è la percentuale di massa presente in A prima del trasporto e il valore $\pi(X \times B)$ è la percentuale di massa arrivata in B dopo il trasporto.

Dato che $\Pi(\mu, \nu) \subseteq \mathcal{P}(X \times Y)$ la massa totale trasportata è sempre 1 (il 100% della massa originariamente in X arriva in Y). Le condizioni marginali che definiscono $\Pi(\mu, \nu)$ assicurano che la massa che parte da A e la massa che arriva in B siano sempre rispettivamente $\mu(A)$ e $\nu(B)$.

Andiamo ora a definire i costi del trasporto che vorremmo minimizzare:

Definizione 1.4. Siano X e Y spazi polacchi. Siano $\mu \in \mathcal{P}(X)$ e $\nu \in \mathcal{P}(Y)$. Una funzione $c : X \times Y \to [0, +\infty]$ Borel-misurabile si dice *funzione costo.*
Sia $\pi \in \Pi(\mu, \nu)$. Definiamo il costo associato al piano di trasporto π come

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

L'integrale presente nella definizione di $\mathcal{C}(\pi)$ è la soluzione matematica al secondo quesito che ci eravamo posti inizialmente: $c(x, y)$ può essere pensata come il costo di una singola spedizione dal punto x al punto y, mentre l'integrale dei costi è una buona formalizzazione del "costo totale".

Definizione 1.5. Il *problema di Kantorovich* consiste nel determinare se esiste $\pi^* \in \Pi(\mu, \nu)$ tale che $\mathcal{C}(\pi^*) = \inf \{\mathcal{C}(\pi) \mid \pi \in \Pi(\mu, \nu)\}$.

Per risolvere questo problema useremo una classica strategia denominata "metodo diretto del calcolo delle variazioni". Questa consiste sostanzialmente nel dimostrare che un'opportuna topologia su $\Pi(\mu, \nu)$ permette di applicare il Teorema di Weierstrass (Teorema 2.12).

Prima di tutto diamo la definizione di semicontinuità inferiore che sarà la nostra condizione sufficiente per l'esistenza di una soluzione ottima:

Definizione 1.6. Sia (X, τ) spazio topologico di Hausdorff e sia $f : X \to \mathbb{R}$ funzione. f si dice *semicontinua inferiormente* in $x_0 \in X$ se $f(x_0) \leq \liminf_{x \to x_0} f(x)$. f è *semicontinua inferiormente* se ciò vale per ogni $x_0 \in X$.

I prossimi capitoli saranno dedicati a dimostrare il seguente teorema:

Teorema 1.4. *Siano X e Y spazi polacchi. Siano $\mu \in \mathcal{P}(X)$ e $\nu \in \mathcal{P}(Y)$.*
Se la funzione costo c è semicontinua inferiormente allora il problema di Kantorovich ammette minimo.

Capitolo 2

Risultati preliminari

2.1 Topologie sugli spazi di misure

Come dicevamo nel capitolo precedente ci serve una topologia su $\mathcal{M}(X \times Y)$ che faccia emergere la compattezza di $\Pi(\mu, \nu)$.
Dato che se (X, τ) è spazio polacco allora $\mathcal{M}(X)$ è spazio vettoriale su $\mathbb{R}$, il modo più naturale per dotare $\mathcal{M}(X)$ di una topologia sembrerebbe quello di costruire una norma.

Definizione 2.1. Sia (X, τ) spazio polacco e sia $\mu \in \mathcal{M}(X)$. La *misura variazione totale* associata a μ è data da

$$|\mu|(B) = \sup \left\{ \sum_{i \in \mathbb{N}} |\mu(B_i)| \,\middle|\, B_i \in \mathcal{B}(X) \,\forall\, i \in \mathbb{N},\ B = \bigcup_{i \in \mathbb{N}} B_i,\ B_i \cap B_j = \emptyset \,\forall i \neq j \right\}.$$

$|\mu|$ è una misura, infatti $|\mu| \geq 0$ e $|\mu|(\emptyset)$ banalmente. Inoltre se $\{A_j\}_{j \in \mathbb{N}} \subseteq \mathcal{B}(X)$ è una successione tale che $A_i \cap A_j = \emptyset \,\forall\, i \neq j$ allora la mappa

$$\left\{ \{B_n\}_{n \in \mathbb{N}} \, \text{partizione di Borel di } \bigcup_{j=1}^{\infty} A_j \right\} \to \prod_{j \in \mathbb{N}} \left\{ \{B_n^j\}_{n \in \mathbb{N}} \, \text{partizione di Borel di } A_j \right\}$$

$$\{B_n\}_{n \in \mathbb{N}} \mapsto (\{B_n \cap A_j\}_{n \in \mathbb{N}})_{j \in \mathbb{N}}$$

è una biezione e pertanto

$$|\mu|\left(\bigcup_{j=1}^{\infty} A_j\right) = \sup\left\{\sum_{i=1}^{\infty} |\mu(B_i)| \,\middle|\, \{B_i\}_{i\in\mathbb{N}} \text{ partizione di Borel di } \bigcup_{j=1}^{\infty} A_j\right\} =$$
$$= \sup\left\{\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} \left|\mu(B_i^j)\right| \,\middle|\, \left\{B_i^j\right\}_{i\in\mathbb{N}} \text{ partizione di Borel di } A_j\right\} =$$
$$= \sum_{j=1}^{\infty} \sup\left\{\sum_{i=1}^{\infty} \left|\mu(B_i^j)\right| \,\middle|\, \left\{B_i^j\right\}_{i\in\mathbb{N}} \text{ partizione di Borel di } A_j\right\} = \sum_{j=1}^{\infty} |\mu|(A_j).$$

Proposizione. $\|\cdot\|_{TV} : \mathcal{M}(X) \to [0,+\infty], \mu \mapsto |\mu|(X)$ *è una norma su* $\mathcal{M}(X)$.

Dimostrazione.

1. $\mu \in \mathcal{M}(X)$ è misura con segno finita quindi $\|\mu\|_{TV} = |\mu|(X) \in [0,+\infty)$ per costruzione.

2. $\|\mu\|_{TV} = 0 \iff |\mu(B_i)| = 0$ per ogni $\{B_i \mid i \in \mathbb{N}\}$ partizione di Borel di X $\iff \mu(B) = 0 \,\forall B \in \mathcal{B}(X) \iff \mu$ è la misura nulla.

3. Siano $\mu, \nu \in \mathcal{M}(X) \implies |(\mu+\nu)(B)| \leq |\mu(B)| + |\nu(B)| \,\forall B \in \mathcal{B}(X)$. Allora:

$$|(\mu+\nu)|(X) =$$
$$= \sup\left\{\sum_{i\in\mathbb{N}} |(\mu+\nu)(B_i)| \,\middle|\, B_i \in \mathcal{B}(X) \forall i \in \mathbb{N},\, X = \bigcup_{i\in\mathbb{N}} B_i,\, B_i \cap B_j = \emptyset \,\forall i \neq j\right\} \leq$$
$$\leq \sup\left\{\sum_{i\in\mathbb{N}} |\mu(B_i)| + |\nu(B_i)| \,\middle|\, B_i \in \mathcal{B}(X) \forall i \in \mathbb{N},\, X = \bigcup_{i\in\mathbb{N}} B_i,\, B_i \cap B_j = \emptyset \,\forall i \neq j\right\} \leq$$
$$\leq |\mu|(X) + |\nu|(X).$$

□

Definendo $\mu_+ = \frac{|\mu|+\mu}{2}$ e $\mu_- = \frac{|\mu|-\mu}{2}$ notiamo che (μ_+, μ_-) è la decomposizione di Jordan di μ. Un altro modo per calcolare $\|\mu\|_{TV}$ conoscendo la decomposizione di Jordan di μ quindi è $\|\mu\|_{TV} = \mu_+(X) + \mu_-(X)$.

Teorema 2.1. ***Decomposizione di Jordan***
Data $\mu \in \mathcal{M}(X)$ *esistono uniche* $\mu_+, \mu_- \in \mathcal{M}_+(X)$ *tali che* $\mu = \mu_+ - \mu_-$ *e* μ_+, μ_- *sono mutualmente singolari (ovvero esistono* $A, B \subseteq X$ *tali che* $A \cup B = X$,

$A \cap B = \emptyset$, $\mu_+(E) = 0 \,\forall\, E \subseteq A$, $\mu_+(F) = 0 \,\forall\, F \subseteq B$).
La coppia (μ_+, μ_-) è detta decomposizione di Jordan *di μ.*

Non riportiamo la dimostrazione del teorema perché non è particolarmente utile per la risoluzione del problema di Kantorovich, per approfondimenti si veda [3].

A questo punto sorge un problema tipico degli spazi normati infinito-dimensionali: in generale le palle non sono totalmente limitate. Per ovviare a questo problema l'approccio usuale è indebolire la topologia indotta dalla norma.

Definizione 2.2. Sia $\mu \in \mathcal{M}(X)$. Considero la seguente base del filtro degli intorni di μ:

$$\left\{ \bigcap_{f \in \mathcal{F}} \left\{ \nu \in \mathcal{M}(X) \,\middle|\, \left| \int_X f d\nu - \int_X f d\mu \right| < \epsilon \right\} \mid \epsilon > 0, \mathcal{F} \subseteq C_b(X), |\mathcal{F}| < \infty \right\}.$$

La topologia generata da queste basi dei filtri di intorni è detta *topologia debole* su $\mathcal{M}(X)$.

La definizione è ben posta perché per costruzione le basi sono chiuse per intersezione.
Osserviamo che data $f \in C_b(X)$ la funzione

$$\phi : (\mathcal{M}(X), \|\cdot\|_{TV}) \to \mathbb{R}, \nu \mapsto \int_X f d\nu$$

è continua perché lineare e limitata:

- Se $\mu, \nu \in \mathcal{M}_+(X)$ e $f \in C_b(X)$, $f \geq 0$ allora per ogni ψ funzione semplice tale che $0 \leq \psi \leq f$ si ha che $\int_X \psi d(\mu + \nu) = \sum_{i=1}^N c_i(\mu(A_i) + \nu(A_i)) = \int_X \psi d\mu + \int_X \psi d\nu$ e pertanto

$$\int_X f d(\mu + \nu) = \sup \left\{ \int_X \psi d(\mu + \nu) \,\middle|\, \psi \,\text{semplice}, \, 0 \leq \psi \leq f \right\} =$$
$$= \sup \left\{ \int_X \psi d\mu + \int_X \psi d\nu \,\middle|\, \psi \,\text{semplice}, \, 0 \leq \psi \leq f \right\} = \int_X f d\mu + \int_X f d\nu$$

Più in generale se $\mu, \nu \in \mathcal{M}(X)$ e $f \in C_b(X)$

$$\int_X f d(\mu+\nu) = \int_X (f^+ - f^-) d(\mu_+ - \mu_- + \nu_+ - \nu_-) =$$
$$= \int_X f^+ d(\mu_+ + \nu_+) - \int_X f^- d(\mu_+ + \nu_+) - \int_X f^+ d(\mu_- + \nu_-) + \int_X f^- d(\mu_- + \nu_-) =$$
$$= \int_X f^+ d(\mu_+ - \mu_-) - \int_X f^- d(\mu_+ - \mu_-) + \int_X f^+ d(\nu_+ - \nu_-) - \int_X f^- d(\nu_+ - \nu_-) =$$
$$= \int f d\mu + \int f d\nu$$

cioè ϕ è lineare su $\mathcal{M}(X)$.

- Basta ora controllare che la norma operatoriale di ϕ sia finita:

$$\sup_{\|\mu\|_{TV}=1} \left| \int_X f d\mu \right| \leq \sup_{\|\mu\|_{TV}=1} \left\{ \left| \int_X f d\mu_+ \right| + \left| \int_X f d\mu_- \right| \right\} \leq$$

$$\leq \sup_{\|\mu\|_{TV}=1} \left\{ \int_X |f| d\mu_+ + \int_X |f| d\mu_- \right\} \leq \sup_{\|\mu\|_{TV}=1} \int_X \|f\|_u d|\mu| \leq \|f\|_u < +\infty$$

Dunque dati $\mu \in \mathcal{M}(X)$ ed $\epsilon > 0$,

$$\left\{ \nu \in \mathcal{M}(X) \,\middle|\, \left| \int_X f d\nu - \int_X f d\mu \right| < \epsilon \right\} = \phi^{-1} \left(\int_X f d\mu - \epsilon, \int_X f d\mu + \epsilon \right)$$

è aperto in $(\mathcal{M}(X), \|\cdot\|_{TV})$.
Da ciò deduciamo che la topologia debole su $\mathcal{M}(X)$ è meno fine di quella indotta dalla norma perché gli insiemi usati per definire la base della prima sono aperti nella seconda.

Il motivo per cui scegliamo questa topologia è dovuto al Teorema di Riesz:

Teorema 2.2. ***Rappresentazione di Riesz***
Sia (X, d) spazio metrico separabile e localmente compatto. Allora per ogni $\phi \in C_0(X)'$ esiste un'unica misura con segno finita $\mu \in \mathcal{M}(X)$ tale che

$$< \phi, f >= \int_X f d\mu \, \forall f \in C_0(X).$$

Inoltre la mappa T che associa ad ogni $\phi \in C_0(X)'$ la misura con segno $\mu \in \mathcal{M}(X)$ come sopra è un'isometria suriettiva $T : (C_0(X)', \|\cdot\|_{op}) \to (\mathcal{M}(X), \|\cdot\|_{TV})$.

Per una dimostrazione di questo risultato fondamentale esteso al contesto più generale delle misure complesse si veda [3] oppure [5].

Se lo spazio polacco (X, τ) è compatto gli spazi di funzioni $C_0(X)$ e $C_b(X)$ coincidono. Perciò, a meno di identificare $\mathcal{M}(X)$ con $C_0(X)'$, la topologia debole coincide con la topologia debole* $\sigma(\mathcal{M}(X), C_0(X))$:

Definizione 2.3. Sia $(E, \|\cdot\|)$ spazio normato e sia E' il suo duale. La *topologia debole** su E' è la meno fine topologia per cui le mappe di valutazione $\{J_x : E' \to \mathbb{R}, f \mapsto < f, x > \mid x \in E\}$ sono continue. Essa è denotata con $\sigma(E', E)$.

Proposizione. *Sia $(E, \|\cdot\|)$ spazio normato. Data $f_0 \in E'$ una base del filtro degli intorni di f_0 in $(E'\sigma(E', E))$ è data da*

$$\left\{ \bigcap_{x \in I} \{f \in E' \mid \langle f - f_0, x \rangle < \epsilon\} \middle| \epsilon > 0, I \subseteq E, |I| < \infty \right\}.$$

La dimostrazione di questa proposizione è presente in [2].

Il fatto che, se (X, τ) è compatto, la topologia debole di $\mathcal{M}(X)$ coincida con la topologia debole* $\sigma(\mathcal{M}(X), C_b(X))$ è importante perché le condizioni di compattezza nella topologia debole* sono semplicemente chiusura e limitatezza:

Teorema 2.3. *Banach-Alaoglu-Bourbaki*
Sia $(E, \|\cdot\|)$ spazio normato. $B_{E'}(0, 1]$ è compatta nella topologia debole.*

Dimostrazione. (Basata su quella proposta in [2]).
Sia $\mathbb{R}^E = \{f : E \to \mathbb{R}\}$ dotato della topologia prodotto τ (ossia della topologia debole rispetto alle mappe di proiezione $\pi_x : \mathbb{R}^E \to \mathbb{R}, f \mapsto f(x)$).
Consideriamo $\phi : B_{E'}(0, 1] \to \mathbb{R}^E$ l'inclusione canonica. Dimostriamo che la mappa $\phi : (B_{E'}(0, 1], \sigma(E', E)) \to (\phi(B_{E'}(0, 1]), \tau)$ è un omeomorfismo.
Sia $x \in E$, $\pi_x \circ \phi(f) = f(x) = \langle J_x, f \rangle$ allora $\pi_x \circ \phi$ è continua sulla topologia debole* $\sigma(E', E)$ per definizione, allora $\pi_x \circ \phi$ è continua per ogni $x \in E$, allora ϕ è continua.
Sia $x \in E$, $J_x \circ \phi^{-1}(f) = \langle f, x \rangle = \pi_x(f)$ allora $J_x \circ \phi^{-1}$ è continua sulla topologia prodotto τ per definizione, allora $J_x \circ \phi^{-1}$ è continua per ogni $x \in E$, allora ϕ^{-1}

è continua.
Dunque ϕ è bicontinua ed è omeomorfismo del dominio sull'immagine. Basta perciò dimostrare che $\phi(B_{E'}(0,1])$ è compatto rispetto alla topologia debole.
$\phi(B_{E'}(0,1]) = \{f : E \to \mathbb{R} \mid f\ lineare\} \cap \{f : E \to \mathbb{R} \mid |f(x)| \leq \|x\|_E \, \forall x \in E\}$.

$$\{f : E \to \mathbb{R} \mid f\ lineare\} = \bigcap_{\alpha,\beta \in \mathbb{R};\, x,y \in E} (\pi_{\alpha x + \beta y} - \alpha \pi_x - \beta \pi_y)^{-1}(\{0\})$$

è chiuso perché intersezione di chiusi.

$$\{f : E \to \mathbb{R} \mid |f(x)| \leq \|x\|_E \, \forall x \in E\} = \prod_{x \in E} [-\|x\|_E, \|x\|_E]$$

è compatto perché prodotto di compatti (Teorema di Tychonoff).
Allora $\phi(B_{E'}(0,1])$ è compatto perché intersezione di un chiuso con un compatto.
In conclusione $B_{E'}(0,1]$ è debole*-compatta perché omeomorfa ad un compatto. $\square$

Osserviamo che con una dimostrazione analoga si ottiene la compattezza debole* di $B_{E'}(0,r]$ per ogni $r > 0$. Di conseguenza:

Corollario 2.4. *Sia $(E, \|\cdot\|)$ spazio normato e sia $C \subseteq E'$ limitato e chiuso rispetto alla topologia debole* $\sigma(E', E)$. Allora C è debole*-compatto.*

Dimostrazione. C è sottoinsieme chiuso di $B(0,r]$ per un opportuno $r > 0$. Per Teorema di Banach-Alaoglu-Bourbaki 2.3 $B(0,r]$ è debole*-compatta e pertanto lo è anche C. $\square$

La dimostrazione del Teorema di Riesz 2.2 è troppo impegnativa per lo scopo di questa discussione, soprattutto perché nella prossima sezione dimostreremo delle condizioni necessarie e sufficienti alla compattezza debole indipendenti dalla dualità (Teorema 2.11)
Nonostante ciò era importante sottolineare questo legame per evidenziare le motivazioni implice dietro la scelta della topologia debole per $\mathcal{M}(X)$. Si rinnova l'invito a consultare le fonti per una trattazione completa degli spazi di misure: si veda [3] oppure [5].

2.2 Compattezza sequenziale

Comiciamo definendo la compattezza sequenziale che useremo per il resto del capitolo:

Definizione 2.4. Sia (X, τ) spazio topologico di Hausdorff. X si dice *sequenzialmente compatto* se ogni successione in X ammette sottosuccessione convergente.

Definizione 2.5. Sia (X, τ) spazio topologico. $\mathcal{F} \subseteq X$ si dice *relativamente compatto* (risp. *relativamente sequenzialmente compatto*) se $\overline{\mathcal{F}}$ è compatto in X (risp. sequenzialmente compatto).

Un risultato classico sugli spazi metrici collega la compattezza sequenziale con la usuale compattezza topologica:

Teorema 2.5. ***Caratterizzazione metrica dei compatti***
Sia (X, d) spazio metrico, allora le seguenti asserzioni sono equivalenti:

1. *X è compatto*
2. *Se $A \subseteq X$ è sottoinsieme infinito allora A ha almeno un punto di accumulazione in X*
3. *X è sequenzialmente compatto*
4. *X è completo e totalmente limitato*

Dimostrazione. Dimostriamo che (1) $\implies$ (2).
Supponiamo X compatto e supponiamo per assurdo che esista $A \subseteq X$ sottoinsieme infinito che non ammette punti di accumulazione.
Allora per ogni $x \in X$ esiste $r_x > 0$ tale che $A \cap B(x, r_x) = \{x\}$.
$\{B(x, r_x) \mid x \in X\}$ è ricoprimento aperto di X, allora esistono $x(1), ..., x(n) \in X$ tali che $X \subseteq \bigcup_{i=1}^n B\left(x(i), r_{x(i)}\right)$ e perciò:

$$A = A \cap X = \bigcup_{i=1}^{n} A \cap B\left(x(i), r_{x(i)}\right) \subseteq \{x(1), ..., x(n)\}.$$

Allora A è finito contro l'ipotesi che A sia infinito.

Dimostriamo che (2) $\implies$ (3).
Sia $A = \{x_n\}_{n \in \mathbb{N}} \subseteq X$ successione.

Se $|A| < \infty$, $x_n = x$ frequentemente per un opportuno $x \in A$, allora $\{x_n\}_{n\in\mathbb{N}}$ ammette sottosuccessione costante e perciò convergente.
Se invece A è infinito esso ammette un punto di accumulazione $x \in X$.
Allora per ogni $k \in \mathbb{N}$ esiste $n(k) \in \mathbb{N}$ tale che $x_{n(k)} \in B\left(x, \frac{1}{k}\right)$. A meno di estrarre un opportuno sottoinsieme da $\{n(k)\}_{k\in\mathbb{N}}$ possiamo assumere che $n(k) < n(k+1)\,\forall\, k \in \mathbb{N}$. Allora $\left\{x_{n(k)}\right\}_{k\in\mathbb{N}}$ è sottosuccesione di $\{x_n\}_{n\in\mathbb{N}}$ che converge a x.

Dimostriamo che (3) $\implies$ (4).
Sia $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ succesione di Cauchy e sia $\{x_n(k)\}_{k\in\mathbb{N}} \subseteq \{x_n\}_{n\in\mathbb{N}}$ sottosuccessione convergente a $x \in X$. Sia $\epsilon > 0$ fissato.
Esiste $n^* \in \mathbb{N}$ tale che $d(x_n, x_m) < \epsilon\,\forall\, n, m > n^*$ perché $\{x_n\}_{n\in\mathbb{N}}$ è di Cauchy.
Esiste $k \in \mathbb{N}$ tale che $n(k) > n^*$ e $d(x, x_{n(k)}) < \epsilon$ perché $x_{n(k)} \to x$.
Allora per ogni $n > n^*$, $d(x, x_n) \leq d(x, x_{n(k)}) + d(x_{n(k)}, x_n) < 2\epsilon$.
Pertanto $\{x_n\}_{n\in\mathbb{N}}$ converge a x.
Sia $r > 0$ aribitrario. Sia $x_0 \in X$ e considero $B_1 = B\left(x_0, r\right)$.
Se esiste $x_1 \in X \smallsetminus B_1$ considero $B_2 = B\left(x_1, r\right)$.
Se esiste $x_2 \in X \smallsetminus (B_1 \cup B_2)$ considero $B_3 = B\left(x_2, r\right)$.
Iterativamente costruisco una famiglia di palle $\{B_1, B_2, ...\}$ e una famiglia di punti $\{x_1, x_2, ...\}$ tali che $x_n \notin \bigcup_{i=1}^{n} B_i$.
Se la famiglia $\{x_1, x_2, ...\}$ fosse infinita sarebbe una successione in X che non ammette sottosuccessione convergente perché per costruzione $d(x_n, x_k) \geq r$ $\forall\, n \in \mathbb{N}\,\forall\, k < n$.
Dunque la famiglia $\{x_1, x_2, ...\}$ è finita e per costruzione lo è anche $\{B_1, B_2, ...\}$.
Allora per costruzione $X \subseteq \bigcup_{i=1}^{N} B_i$.
Per arbitrarietà di $r > 0$ dunque X è totalmente limitato.

Dimostriamo che (4) $\implies$ (1).
Dimostriamo preliminarmente che se (X, d) è completo e $\{C_n\}_{n\in\mathbb{N}}$ è una successione decrescente di chiusi tale che $\lim_{n\to\infty} \operatorname{diam}(C_n) = 0$ allora $\bigcap_{n=1}^{\infty} C_n = \{x\}\ \exists!\, x \in X$.
Sia $\epsilon > 0$ fissato. Siano $x_n \in C_n\,\forall\, n \in \mathbb{N}$.
Sia $n^* \in \mathbb{N}$ tale che $\operatorname{diam}(C_{n*}) < \epsilon$.
$\forall\, n, m > n^*$, $n, m \in C_{n^*} \implies d(x_n, x_m) \leq \operatorname{diam}(C_{n^*}) < \epsilon$.
Allora $\{x_n\}_{n\in\mathbb{N}}$ è succesione di Caucy e per completezza essa converge a $x \in X$.
Allora per ogni $k \in \mathbb{N}$, $\{x_n\}_{n>k}$ è succesione di Caucy in C_k chiuso che converge a x, dunque $x \in C_k\,\forall\, k \in \mathbb{N}$, e perciò $\{x\} \subseteq \bigcap_{n=1}^{\infty} C_n$.

Se esistesse $y \in \bigcap_{n=1}^{\infty} C_n$, $y \neq x$ si avrebbe che $\text{diam}(\bigcap_{n=1}^{\infty} C_n) \geq d(x,y) > 0$, ma ciò è impossibile perché $\lim_{n\to\infty} \text{diam}(C_n) = 0$.

Sia ora (X,d) completo e totalmente limitato e supponiamo per assurdo che non sia compatto.
Allora esiste $\{A_j\}_{j\in J}$ ricoprimento aperto di X che non ammette sottoricoprimento finito.
X è totalmente limitato: siano $\left\{B_1^1, ...B_{n(1)}^1\right\}$ palle chiuse di raggio 1 che ricoprono X. Allora esiste $i(1)$ tale che $B_{i(1)}^1$ non è ricoperta da un numero finito di A_j.
$B_{i(1)}^1$ è totalmente limitata: siano $\left\{B_1^2, ..., B_{n(2)}\right\}$ palle chiuse in $(B_{i(1)}^1, d)$ di raggio $\frac{1}{2}$ che la ricoprono. Allora esiste $i(2)$ tale che $B_{i(2)}^2$ non è ricoperta da un numero finito di A_j.
Iterativamente per ogni $k \in \mathbb{N}$ esiste $B_{i(k)}^k$ palla chiusa in $(B_{i(k-1)}^{k-1}, d)$ di raggio $\frac{1}{k}$ che non è ricoperta da un numero finito di A_j.
Per quanto visto prima esiste un unico $x \in X$ tale che $\{x\} = \bigcap_{k\in\mathbb{N}} B_{i(k)}^k$ perché $\lim_{k\to\infty} \text{diam}(B_{i(k)}^k) = \lim_{k\to\infty} \frac{2}{k} = 0$.
Allora esiste $j^* \in J$ tale che $x \in A_{j^*}$ aperto e perciò esiste $r^* > 0$ tale che $B(x, r^*) \subseteq A_{j^*}$.
Allora per ogni $k > \frac{2}{r*}$, $B_{i(k)}^k \subseteq B\left(y, \frac{1}{k}\right] \subseteq B\left(x, \frac{1}{k} + d(x,y)\right] \subseteq B\left(x, \frac{2}{k}\right] \subseteq B(x, r*) \subseteq A_{j*}$, assurdo perché $B_{i(k)}^k$ non è ricoperta da un numero finito di A_j. □

Questo risultato vale anche su spazi topologici con topologia metrizzabile perché, data una metrica che induce la topologia, una successione converge topologicamente se e solo se converge in metrica.

Dato che useremo la compattezza sequenziale nella topologia debole* ci serve la metrizzabilità di quest'ultima:

Teorema 2.6. *Sia E uno spazio normato separabile. Sia $\{x_n\}_{n\in\mathbb{N}}$ sottoinsieme denso e numerabile di $B_E(0,1]$.*
Consideriamo $B_{E'}(0,1]$ e definiamo la seguente metrica:

$$d(f,g) = \sum_{n=1}^{\infty} \frac{1}{2^n} |\langle f - g, x_n \rangle| \ \ \forall f, g \in B_{E'}(0,1].$$

Allora la metrica d induce la topologia debole $\sigma(E', E)$ su $B_{E'}(0,1]$.*

Per la dimostrazione di questo teorema si veda [5].

Osserviamo che allo stesso modo si possono costruire metriche che inducono la topologia debole* su $B_{E'}(0, r]$ per ogni $r > 0$.

Corollario 2.7. *Sia E spazio normato separabile e sia $\{f_n\}_{n\in\mathbb{N}} \subseteq E'$ successione limitata. Allora $\{f_n\}_{n\in\mathbb{N}}$ ammette sottosuccessione debole*-convergente.*

Dimostrazione. Per il Corollario del Teorema di Banach-Alaoglu-Bourbaki 2.3 $\{f_n\}_{n\in\mathbb{N}}$ è relativamente debole*-compatta. Dato che la topologia debole* su insiemi limitati è metrizzabile $\{f_n\}_{n\in\mathbb{N}}$ è anche relativamente debole* sequenzialmente compatta e perciò ammette sottosuccesione convergente in E'. $\square$

Per poter applicare questo teorema nel nostro contesto resta da verificare la separabilità di $(C_0(X), \|\cdot\|_u)$. In generale questo non è possibile, ma lo è nel caso che ci serve, ossia quando topologia debole e topologia debole* su $\mathcal{M}(X)$ coincidono:

Lemma 2.8. *Sia (X, τ) spazio polacco compatto. Allora $(C(X), \|\cdot\|_u)$ è separabile.*

Per dimostrare il lemma useremo il seguente Teorema, la cui dimostrazione è presente in [3]:

Teorema 2.9. ***Stone-Weierstrass***
Sia (X, τ) uno spazio topologico di Hausdorff compatto e sia $\mathcal{A} \subseteq C(X)$ una sottoalgebra che contiene almeno una funzione costante non nulla.
Allora $\mathcal{A}$ è densa in $C(X) \iff \mathcal{A}$ separa i punti, ossia per ogni $x, y \in X$, $x \neq y$, esiste $f \in \mathcal{A}$ tale che $f(x) \neq f(y)$.

Dimostrazione. (Del Lemma 2.8).
Sia d una metrica su X che induce τ tale che (X, d) sia completo e separabile.
Sia $\{x_n\}_{n\in\mathbb{N}}$ un sottoinsieme denso e numerabile di (X, d).
Definiamo $f_n : X \to [0, +\infty]$, $x \mapsto d(x, x_n)$ funzioni continue e denotiamo con $\mathbb{1}$ la funzione costante $\mathbb{1} : X \to \mathbb{R}$, $x \mapsto 1$.
Consideriamo l'insieme dei prodotti finiti di tali funzioni:

$$\mathcal{F} = \left\{ \prod_{i\in I} f_i^{p_i} \,\middle|\, I \subseteq \mathbb{N}, |I| < \infty,\, p_i \in \mathbb{N} \right\}$$

e definiamo lo spazio vettoriale su $\mathbb{Q}$ delle combinazioni lineari in $\mathcal{F} \cup \{\mathbb{1}\}$:

$$\mathcal{A} = \left\{ \sum_{j \in J} q_j F_j + q\mathbb{1} \,\middle|\, J \subseteq \mathbb{N}, |J| < \infty, \, q, q_j \in \mathbb{Q}, \, F_j \in \mathcal{F} \right\}.$$

Per costruzione $\mathcal{A}$ è chiusa per somma e prodotto puntuale di funzioni ed è dunque sottoalgebra di $C(X)$.
Siano $x, y \in X$, $x \neq y$. (X, d) è Hausdorff quindi esistono $B_x, B_y \subseteq X$ palle di raggio r tali che $x \in B_x$, $y \in B_y$, $B_x \cap B_y = \emptyset$. $\{x_n\}_{n\in\mathbb{N}}$ è denso quindi esiste $n \in \mathbb{N}$ tale che $x_n \in B_x$. Allora per costruzione $f_n(x) = d(x, x_n) < r$ e $f_n(y) = d(y, x_n) \geq r$, quindi esiste $f_n \in \mathcal{A}$ tale che $f_n(x) \neq f_n(y)$.
Per Teorema di Stone-Weiestrass allora $\mathcal{A}$ è densa in $C(X)$.
Infine $\{f_n\}_{n\in\mathbb{N}}$ è numerabile quindi per costruzione $\mathcal{F}$ è numerabile e perciò anche $\mathcal{A} =< \mathcal{F} \cup \{\mathbb{1}\} >_{\mathbb{Q}}$ è numerabile. Per definizione quindi $C(X)$ è separabile. $\square$

In generale la topologia debole e la topologia debole* su $\mathcal{M}(X)$ non coincidono (se X non è localmente compatto non si può applicare il Teorema di Riesz e la topologia debole* non è nemmeno definibile). Ci servono quindi delle condizioni di compatteza più generali del Teorema di Banach-Aloglu-Bourbaki.

Spesso si possono trovare delle condizioni necessarie e sufficienti per cui un insieme limitato risulta essere anche totalmente limitato. Il prototipo di questa strategia è il Teorema di Ascoli-Arzelà (per una trattazione di questo teorema si veda il libro di W. Rudin [5]).
Nel nostro contesto tali condizioni necessarie e sufficienti sono date dal Teorema di Prokhorov. Prima di tutto però enunciamo e dimostriamo un Lemma che ci servirà anche nel Capitolo 3:

Lemma 2.10. *Sia (X, τ) spazio polacco e sia $f : X \to [0, +\infty]$ semicontinua inferiormente. Sia $\{\mu_n\}_{n\in\mathbb{N}} \in \mathcal{P}(X)$ che converge debolmente a μ. Allora*

$$\int_X f d\mu \leq \liminf_{n\to\infty} \int_X f d\mu_n.$$

Dimostrazione. *(Basata su quella proposta in [1]).*
Sia d una metrica su X che induce τ tale che (X, d) sia completo e separabile.
Per ogni $k \in \mathbb{N}$ definiamo $f_k(x) = \inf_{y\in X} \{\min \{f(y), k\} + kd(x, y)\}$.
Ovviamente per ogni $k \in \mathbb{N}$ si ha che $0 \leq f_k \leq f_{k+1} \leq \min \{f, k\}$.

Per ogni $y \in X$ la funzione $\min\{f(y), k\} + kd(x, y)$ è lipschitziana con costante di Lipschitz $\frac{1}{k}$. Allora la funzione $f_k(x)$ è l'estremo inferiore puntuale di una famiglia di funzioni equi-lipschitziane ed è pertanto lipschitziana con costante di Lipschitz minore o uguale a $\frac{1}{k}$.

Sia ora $x \in X$ fissato e supponiamo senza perdere di generalità che $f(x) < +\infty$ (altrimenti per costruzione $f_k(x) \uparrow +\infty$). Per ogni $k \in \mathbb{N}$ sia $x_k \in X$ tale che $\min\{f(x_k), k\} + kd(x, x_k) \leq f_k(x) + \frac{1}{k}$ (ciò è lecito perché $f_k(x) \leq \min\{f, k\} \leq f(x) < +\infty \,\forall\, k \in \mathbb{N}$).

Necessariamente allora $\lim_{k\to\infty} x_k = x$.

Inoltre dato che $\min\{f(x_k), k\} \leq \min\{f(x_k), k\} + kd(x, x_k) \leq f_k(x) + \frac{1}{k}$ e dato che f è semicontinua inferiormente, per k sufficientemente grandi si ha che $f(x) - \frac{1}{k} \leq f(x_k) \leq \min\{f(x_k), k\} \leq f_k(x) + \frac{1}{k}$. Allora $f(x) \leq f_k(x) + \frac{2}{k}$ per k sufficientemente grandi e pertanto $f(x) \leq \lim_{k\to\infty} f_k(x)$ (il limite eiste perché la successione $\{f_k(x)\}_{k\in\mathbb{N}}$ è monotona).

Dato che $f_k(x) \leq f(x) \,\forall\, k \in \mathbb{N}$ ciò dimostra che $f_k \uparrow f$ puntualmente in X.

Ora per ogni $k \in \mathbb{N}$, $\liminf_{n\to\infty} \int_X f d\mu_n \geq \liminf_{n\to\infty} \int_X f_k d\mu_n = \int_X f_k d\mu$ per convergenza debole perché $f_k \in Lip_b(X) \implies f_k \in C_b(X)$.

Allora $\liminf_{n\to\infty} \int_X f d\mu_n \geq \sup_{k\in\mathbb{N}} \int_X f_k d\mu = \lim_{k\to\infty} \int_X f_k d\mu = \int_X f d\mu$ per Teorema di convergenza monotona. □

Teorema 2.11. ***Prokhorov***

Sia (X, τ) spazio polacco e sia $\mathcal{F} \subseteq \mathcal{P}(X)$.

Allora $\mathcal{F}$ è relativamente sequenzialmente compatto nella topologia debole di $\mathcal{M}(X)$ $\iff$ $\mathcal{F}$ è equi-stretto, cioè per ogni $\epsilon > 0$ esiste $K \subseteq X$ compatto tale che $\mu(X \setminus K) < \epsilon$.

Dimostrazione. (*Basata su quella proposta in [1]*).

Dimostriamo che se $\mathcal{F}$ è equi-stretto allora $\mathcal{F}$ è relativamente sequenzialmente compatto.

Per regolarità (Teorema 1.1) per ogni $\mu \in \mathcal{F}$, $\mu(X) = \sup\{\mu(K) \mid K \subseteq X \textit{ compatto}\}$.

Dunque esiste una successione di compatti $\{K_k \subseteq X\}_{k\in\mathbb{N}}$ tali che $K_k \subseteq K_{k+1}$ e

$$\lim_{k\to\infty} \omega_k = \lim_{k\to\infty} \sup_{\mu\in\mathcal{F}} \mu(X \setminus K_k) = 0.$$

Sia ora $\{\mu_n\}_{n\in\mathbb{N}}$ successione in $\mathcal{F}$.

Sia $k \in \mathbb{N}$ e consideriamo la successione di misure $\{\mu_n^k\}_{n\in\mathbb{N}} \subseteq \mathcal{M}_+(K_k)$ data da $\mu_n^k(A) = \mu_n(A \cap K_k)$. Dato che $\sup_{n\in\mathbb{N}} \mu_n^k(K_k) \leq 1$, per quanto visto prima (Co-

rollario 2.7) esiste $\nu_k \in \mathcal{M}_+(K_k)$ ed esiste $\left\{\mu^k_{n(l)}\right\}_{l\in\mathbb{N}}$ sottosuccessione di $\left\{\mu^k_n\right\}_{n\in\mathbb{N}}$ che converge debolmente a ν_k.
Sia $\left\{\mu_{n(l(1))}\right\}_{l(1)\in\mathbb{N}}$ sottosuccessione di $\left\{\mu_n\right\}_{n\in\mathbb{N}}$ tale che $\left\{\mu^1_{n(l(1))}\right\}_{l(1)\in\mathbb{N}}$ converge debolmente a ν_1 in $\mathcal{M}_+(K_1)$.
Sia $\left\{\mu_{n(l(2))}\right\}_{l(2)\in\mathbb{N}}$ sottosuccessione di $\left\{\mu_{n(l(1))}\right\}_{l(1)\in\mathbb{N}}$ tale che $\left\{\mu^2_{n(l(2))}\right\}_{l(2)\in\mathbb{N}}$ converge debolmente a ν_2 in $\mathcal{M}_+(K_2)$ e $\left\{\mu^1_{n(l(2))}\right\}_{l(2)\in\mathbb{N}}$ converge debolmente a ν_1 in $\mathcal{M}_+(K_1)$.
Iterativamente possiamo costruire $\left\{\mu_{n(l(m))}\right\}_{l(m)\in\mathbb{N}}$ sottosuccessione di $\left\{\mu_n\right\}_{n\in\mathbb{N}}$ tale che $\left\{\mu^k_{n(l(m))}\right\}_{l(m)\in\mathbb{N}}$ converge debolmente a ν_k in $\mathcal{M}_+(K_k)$ per ogni $k < m$.
Per arbitrarietà di $m \in \mathbb{N}$ esiste $\left\{\mu_{n(p)}\right\}_{p\in\mathbb{N}}$ sottosuccessione di $\left\{\mu_n\right\}_{n\in\mathbb{N}}$ tale che $\left\{\mu^k_{n(p)}\right\}_{p\in\mathbb{N}}$ converge debolmente a ν_k in $\mathcal{M}_+(K_k)$ (e per estensione in $\mathcal{M}_+(X)$) per ogni $k \in \mathbb{N}$.
In particolare dato che $K_k \subseteq K_{k+1}$ abbiamo che $\nu_k(B) = \lim_{p\to\infty} \mu^k_{n(p)}(B) = \lim_{p\to\infty} \mu^{k+1}_{n(p)}(B \cap K_k) = \nu_{k+1}(B \cap K_k)$ dove "lim" si riferisce alla convergenza debole. Quindi per ogni $B \in \mathcal{B}(X)$, $\nu_k(B) = \nu_{k+1}(B \cap K_k) \leq \nu_{k+1}(B)$.
Definiamo $\nu(B) = \sup_{k\in\mathbb{N}} \nu_k(B)\, \forall B \in \mathcal{B}(X)$.
Per costruzione $\nu \geq 0$ e $\nu(\emptyset) = 0$.
Se $A, B \in \mathcal{B}(X)$, $A \cap B = \emptyset$ abbiamo che $\nu(A \cup B) = \sup_{k\in\mathbb{N}} \nu_k(A \cup B) = \sup_{k\in\mathbb{N}} \{\nu_k(A) + \nu_k(B)\} = \nu(A) + \nu(B)$.
Se $\{A_j\}_{j\in\mathbb{N}} \subseteq \mathcal{B}(X)$ abbiamo che $\nu(\bigcup_{j=1}^{\infty}) = \sup_{k\in\mathbb{N}} \nu_k(\bigcup_{j=1}^{\infty}) \leq \sup_{k\in\mathbb{N}} \sum_{j=1}^{\infty} \nu_k(A_j) \leq \sum_{j=1}^{\infty} \sup_{k\in\mathbb{N}} \nu_k(A_j) = \sum_{j=1}^{\infty} \nu(A_j)$ perché $\nu_k \leq \nu_{k+1}$.
Allora ν è finitamente additiva e σ-subadditiva e perciò $\nu \in \mathcal{M}_+(X)$.
Dato che $1 - \omega_k \leq \mu_{n(p)}(X) - \mu_{n(p)}(X \setminus K_k) = \mu^k_{n(p)}(X) \leq 1$, per ogni $p \in \mathbb{N}$ abbiamo che $1 - \omega_k \leq \nu_k(X) \leq 1$ e perciò $\nu(X) = 1$, ossia $\nu \in \mathcal{P}(X)$.
Dimostriamo che $\left\{\mu_{n(p)}\right\}_{p\in\mathbb{N}}$ converge debolmente a ν. Sia $\phi \in C_b(X)$,

$$\left|\int_X \phi d\mu_{n(p)} - \int_X \phi d\nu\right| \leq$$

$$\leq \left|\int_X \phi d\mu_{n(p)} - \int_X \phi d\mu^k_{n(p)}\right| + \left|\int_X \phi d\mu^k_{n(p)} - \int_X \phi d\nu_k\right| + \left|\int_X \phi d\nu_k - \int_X \phi d\nu\right| \leq$$

$$\leq \left|\int_X \phi d\mu^k_{n(p)} - \int_X \phi d\nu_k\right| + \int_x |\phi| d(\mu_{n(p)} - \mu^k_{n(p)}) + \int_X |\phi| d(\nu - \nu_k) \leq$$

$$\leq \left|\int_X \phi d\mu^k_{n(p)} - \int_X \phi d\nu_k\right| + \|\phi\|_u (\mu_{n(p)} - \mu^k_{n(p)})(X) + \|\phi\|_u (\nu - \nu_k)(X) =$$

$$= \left| \int_X \phi d\mu^k_{n(p)} - \int_X \phi d\nu_k \right| + \|\phi\|_u (\mu_{n(p)}(X \smallsetminus K_k) + \nu(X \smallsetminus K_k)) \leq$$

$$\leq \left| \int_X \phi d\mu^k_{n(p)} - \int_X \phi d\nu_k \right| + 2\|\phi\|_u \omega_k.$$

Allora

$$\limsup_{p\to\infty} \left| \int_X \phi d\mu_{n(p)} - \int_X \phi d\nu \right| \leq 2\|\phi\|_u \omega_k \xrightarrow[k\to\infty]{} 0.$$

Dimostriamo ora che se $\mathcal{F}$ è relativamente sequenzialmente compatto allora $\mathcal{F}$ è equi-stretto.
Dimostriamo preliminarmente che se $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(X)$ converge debolmente a $\mu \in \mathcal{P}(X)$ allora per ogni $C \subseteq X$ chiuso $\limsup_{n\to\infty} \mu_n(C) \leq \mu(C)$.
Sia $C \subseteq X$ chiuso e consideriamo $f(x) = \chi_A(x)$ con χ_A funzione caratteristica di $A = X \smallsetminus C$ aperto. Osserviamo che f è semicontinua inferiormente.
Allora grazie al Lemma 2.10 $\mu(C) = \int_X (1 - f(x)) d\mu(x) \geq 1 - \liminf_{n\to\infty} \int_X f d\mu_n = \limsup_{n\to\infty} (1 - \mu_n(A)) = \limsup_{n\to\infty} \mu_n(X \smallsetminus A) = \limsup_{n\to\infty} \mu_n(C)$.
Sia ora $\epsilon > 0$ fissato e sia $\{x_i\}_{i\in\mathbb{N}} \subseteq X$ un sottoinsieme denso e numerabile.
Per dimostrare che $\mathcal{F}$ è equi-stretto è sufficiente dimostrare che per ogni $j \in \mathbb{N}$ esiste $k_j \in \mathbb{N}$ tale che $\mu(X \smallsetminus \bigcup_{i=1}^{k_j} B\left(x_i, \frac{1}{j}\right)) \leq \frac{\epsilon}{2^j}\, \forall \mu \in \mathcal{F}$ in quanto posto $K = \bigcap_{j=1}^{\infty} \bigcup_{i=1}^{k_j} B\left(x_i, \frac{1}{j}\right]$ esso è compatto (come visto nella dimostrazione del Lemma di Ulam 1.3) e per ogni $\mu \in \mathcal{F}$ vale

$$\mu(X \smallsetminus K) \leq \mu\left(\bigcup_{j=1}^{\infty} \left(X \smallsetminus \bigcup_{i=1}^{k_j} B\left(x_i, \frac{1}{j}\right] \right) \right) \leq \sum_{j=1}^{\infty} \frac{\epsilon}{2^j} = \epsilon.$$

Supponiamo per assurdo che esista $j_0 \in \mathbb{N}$ tale che

$$\forall\, k \in \mathbb{N}\, \exists\, \mu_k \in \mathcal{F} : \mu_k \left(X \smallsetminus \bigcup_{i=1}^{k_j} B\left(x_i, \frac{1}{j}\right) \right) > \frac{\epsilon}{2^{j_0}}.$$

Per la relativa sequenziale compattezza di $\mathcal{F}$ la successione $\{\mu_k\}_{k\in\mathbb{N}}$ ammette sottosuccessione $\left\{\mu_{k(n)}\right\}_{n\in\mathbb{N}}$ che converge debolmente a $\mu \in \mathcal{P}(X)$.
Per quanto visto prima per ogni $k \in \mathbb{N}$ vale

$$\mu\left(X \smallsetminus \bigcup_{i=1}^{k} B\left(x_i, \frac{1}{j_0}\right) \right) \geq \limsup_{n\to\infty} \mu_{n(k)} \left(X \smallsetminus \bigcup_{i=1}^{k} B\left(x_i, \frac{1}{j_0}\right) \right) \geq \frac{\epsilon}{2^{j_0}}$$

e pertanto $\mu\left(X \setminus \bigcup_{i=1}^{\infty} B\left(x_i, \frac{1}{j_0}\right)\right) \geq \frac{\epsilon}{2^{j_0}}$, ma ciò è assurdo perché
$X = \bigcup_{i=1}^{\infty} B\left(x_i, \frac{1}{j_0}\right)$ per densità di $\{x_i\}_{i\in\mathbb{N}}$. $\square$

In realtà la distinzione tra compattezza e compattezza sequenziale nel Teorema di Prokhorov è superflua in quanto se (X, τ) è spazio polacco la topologia debole su $\mathcal{P}(X)$ è metrizzabile. La dimostrazione di questo risultato è molto tecnica ed è reperibile sui testi specifici della teoria del trasporto ottimo (ad esempio [7]).

Concludiamo questo capitolo con la dimostrazione del Teorema di Weierstrass (fondamentale per il metodo diretto del calcolo delle variazioni):

Teorema 2.12. ***Weierstrass***
Sia (X, τ) spazio topologico di Hausdorff sequenzialmente compatto e sia $f : X \to \mathbb{R}$ semicontinua inferiormente. Allora f ammette minimo in X.

Dimostrazione. Sia $I = \inf_{x\in X} f(x)$.
Basta dimostrare che esiste $x^* \in X$ tale che $f(x^*) = I$.
Sia $\{x_n\}_{n\in\mathbb{N}}$ successione in X tale che $\lim_{n\to\infty} f(x_n) = I$.
Se $x_n = x^*$ definitivamente allora $f(x^*) = I$.
Altrimenti esiste $\left\{x_{n(k)}\right\}_{k\in\mathbb{N}}$ sottosuccessione di $\{x_n\}_{n\in\mathbb{N}}$ convergente a $x^* \in X$ e tale che $x_{n(k)} \neq x^* \,\forall\, k \in \mathbb{N}$.
Allora $f(x^*) \leq \liminf_{y\to x^*} f(y) \leq \lim_{k\to\infty} f(x_{n(k)}) = I$, pertanto $f(x^*) = I$. $\square$

Capitolo 3

Esistenza di soluzioni ottime

3.1 Dimostrazione del teorema

Ricordiamo che il nostro obiettivo è dimostrare il Teorema 1.4. Vorremmo applicare il Teorema di Weierstrass 2.12 alla funzione

$$\mathcal{C} : \Pi(\mu, \nu) \to \mathbb{R},\ \pi \mapsto \int_{X\times Y} c(x, y) d\pi(x, y).$$

Per fare ciò dobbiamo dimostrare che il dominio $\Pi(\mu, \nu)$ è compatto e la funzione $\mathcal{C}$ è semicontinua inferiormente sotto un'opportuna topologia. Consideriamo la topologia su $\Pi(\mu, \nu)$ indotta dalla topologia debole su $\mathcal{M}(X \times Y)$:

Lemma 3.1. $\Pi(\mu, \nu)$ *dotato della topologia debole è compatto.*

Dimostrazione. (Basata su quella proposta in [1]).
Ricordiamo che

$$\begin{aligned}\Pi(\mu, \nu) = \{\pi \in \mathcal{P}(X \times Y) \mid \pi(A \times Y) = \mu(A) \,\forall\, A \in \mathcal{B}(X),\\ \pi(X \times B) = \nu(B) \,\forall\, B \in \mathcal{B}(Y)\}.\end{aligned}$$

Sia $\pi \in \Pi(X \times Y)$. Sia $\phi(x) \in C_b(X)$ vista come funzione in $X \times Y$ indipendente da y. Allora ogni approssimazione con funzioni semplici di ϕ può essere presa in modo che gli insiemi misurabili che definiscono tali funzioni semplici siano del tipo $A \times Y$ per opportuni $A \in \mathcal{B}(X)$.
Dato che per ogni $A \in \mathcal{B}(X)$, $\pi(A \times Y) = \mu \otimes \nu(A \times Y)$, grazie al Teorema di

convergenza monotona si ha che

$$\int_{X\times Y}\phi(x)d\pi(x,y)=\int_{X\times Y}\phi(x)d\mu\otimes\nu(x,y)=\int_X\phi(x)d\mu(x)\ \forall\phi\in C_b(X).$$

Analogamente

$$\int_{X\times Y}\psi(y)d\pi(x,y)=\int_{X\times Y}\psi(y)d\mu\otimes\nu(x,y)=\int_Y\psi(y)d\nu(y)\ \forall\psi\in C_b(Y).$$

Siano $\phi\in C_b(X)$ e $\psi\in C_b(Y)$, allora le mappe
$F_\phi:\mathcal{P}(X\times Y)\to\mathbb{R}$, $\pi\mapsto\int_{X\times Y}\phi d\pi$ e $G_\psi:\mathcal{P}(X\times Y)\to\mathbb{R}$, $\pi\mapsto\int_{X\times Y}\psi d\pi$
per definizione sono continue rispetto alla topologia debole di $\mathcal{P}(X\times Y)$. Pertanto

$$\Pi(\mu,\nu)=\bigcap_{\phi\in C_b(X)}F_\phi^{-1}\left(\left\{\int_X\phi d\mu\right\}\right)\cap\bigcap_{\psi\in C_b(Y)}G_\psi^{-1}\left(\left\{\int_Y\psi d\nu\right\}\right)$$

è chiuso nella topologia debole di $\mathcal{P}(X\times Y)$.

Dimostriamo ora che l'insieme di misure di probabilità $\Pi(\mu,\nu)$ è equi-stretto e perciò debolmente compatto grazie al Teorema di Prokhorov 2.11.
Sia $\epsilon>0$ fissato. Per il Lemma di Ulam 1.3 esistono $H\subseteq X$ e $K\subseteq Y$ compatti tali che $\mu(X\smallsetminus H)<\frac{\epsilon}{2}$ e $\nu(Y\smallsetminus K)<\frac{\epsilon}{2}$.
Allora $H\times K\subseteq X\times Y$ è compatto e per ogni $\pi\in\Pi(\mu,\nu)$, $\pi(X\times Y\smallsetminus H\times K)\leq\pi((X\smallsetminus H)\cup(Y\smallsetminus K))\leq\mu(H)+\nu(K)<\epsilon$. □

Occupiamoci ora della semicontinuità inferiore di $\mathcal{C}$:

Lemma 3.2. *Se la funzione costo $c:X\times Y\to[0,+\infty]$ è semicontinua inferiormente (rispetto alla topologia prodotto su $X\times Y$) allora $\mathcal{C}$ è semicontinua inferiormente rispetto alla topologia debole.*

Dimostrazione. (Basata su quella proposta in [1]).
Possiamo applicare il Lemma del Teorema di Prokhorov 2.10 a $f=c$ perché $X\times Y$ con la topologia prodotto è spazio polacco (basta prendere come metrica $d((x_1,y_1),(x_2,y_2))=d_X(x_1,y_1)+d_Y(x_2,y_2)$).
Sia $\{\pi_n\}_{n\in\mathbb{N}}\subseteq\mathcal{P}(X\times Y)$ che converge debolmente a $\pi\in\mathcal{P}(X\times Y)$. Si ha che $\mathcal{C}(\pi)=\int_{X\times Y}cd\pi\leq\liminf_{n\to\infty}\int_{X\times Y}cd\pi_n=\liminf_{n\to\infty}\mathcal{C}(\pi_n)$.
Allora per definizione $\mathcal{C}$ è semicontinua inferiormente rispetto allla topologia debole di $\mathcal{P}(X\times Y)$. □

Teorema 3.3. *Siano X e Y spazi polacchi. Siano $\mu \in \mathcal{P}(X)$ e $\nu \in \mathcal{P}(Y)$.*
Se c è semicontinua inferiormente allora il problema di Kantorovich ammette minimo.

Dimostrazione. Sappiamo che $\Pi(\mu, \nu)$ è compatto (Lemma 3.1) e $\mathcal{C} : \Pi(\mu, \nu) \to \mathbb{R}$ è semicontinua inferiormente (Lemma 3.2).
Per Teorema di Weierstrass 2.12 dunque esiste $\min \{\mathcal{C}(\pi) \mid \pi \in \Pi(\mu, \nu)\}$ ossia esiste una soluzione ottima per il problema di Kantorovich. □

3.2 Considerazioni finali

La dimostrazione che abbiamo fornito non è costruttiva, pertanto non permette di calcolare effettivamente il minimo (ne dimostra soltanto l'esistenza). Ciò può risultare problematico qualora si volessero risolvere quantitativamente dei problemi legati al trasporto ottimo.
Sfruttiamo quest'ultima sezione per discutere di possibili soluzioni per questa criticità.

Dato che l'insieme delle soluzioni ammissibili $\Pi(\mu, \nu)$ è convesso è ragionevole pensare che il problema di Kantorovich ammetta un problema duale. In effetti una decrizione del problema duale venne fornita dallo stesso Kantorovich nelle sue pubblicazioni: si veda [4].
La formalizzazione del problema duale sfrutta la teoria dell'integrazione nello spazio L^1 ed è quindi opinabilmente più "intuitiva" di quella del problema originario. Tuttavia la dimostrazione del Teorema di dualità è assai più sofisticata dell'approccio diretto usato in questa tesi. Per approfondimenti si consulti [1].

Un'alternativa allo studio del problema generale è la restrizione al calcolo della soluzione ottima per una funzione costo fissata. Ad esempio il teorema di Kantorovich-Rubinstein tratta il caso in cui la funzione costo è una metrica.
Per una trattazione esaustiva di questo teorema, nonché del caso $c(x, y) = |x - y|^2$ si confrontino le fonti già citate con il libro di C. Villani: [8].

Infine un'ulteriore opzione in determinati contesti potrebbe essere quella di aggiungere vincoli all'insieme delle soluzioni ammissibili. Ad esempio nell'ambito della finanza matematica è utile considerare solo le misure di probabilità π che siano anche "misure martingala" (ovvero tali che determinate funzioni verifichino la condizione di martingalità per queste misure). Una discussione più approfondita degli aspetti concreti del trasporto ottimo è reperibile nel libro di F. Santambrogio: [6].

Bibliografia

[1] Luigi Ambrosio, Elia Brué, and Daniele Semola. *Lectures on Optimal Transport*. Springer Nature Switzerland AG, 2021.

[2] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer New York, 2010.

[3] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

[4] Leonid Kantorovitch. On the translocation of masses. *Management science*, 5(1):1–4, 1958.

[5] Walter Rudin. *Real and Complex Analysis*. Mathematics series. McGraw-Hill, 1987.

[6] Filippo Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*. Progress in Nonlinear Differential Equations and Their Applications. Springer International Publishing, 2015.

[7] Cédric Villani. *Optimal Transport: Old and New*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2008.

[8] Cédric Villani. *Topics in optimal transportation*, volume 58. American Mathematical Soc., 2021.