

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

CORSO DI LAUREA MAGISTRALE IN MATEMATICA



UNIVERSITÀ DI PISA

TESI DI LAUREA MAGISTRALE

Trasporto ottimo e problemi di evoluzione per sistemi di particelle

CANDIDATO:

Simone Di Marino

RELATORE:

Prof. Luigi Ambrosio

CONTRORELATORE:

Prof. Maurizio Pratelli

ANNO ACCADEMICO 2010/2011

Indice

Introduzione	1
1 Trasporto ottimo in una dimensione	3
1.1 Nozioni preliminari	3
1.1.1 Funzioni di distribuzione e pseudo-inversa	4
1.2 Problema di Monge	5
1.3 Problema di Kantorovich	7
1.3.1 Problema di Kantorovich in \mathbb{R}	9
2 Flussi gradiente in spazi di Hilbert	13
2.0.2 Richiami di analisi convessa	18
2.1 La soluzione discreta e il rapporto tra EVI e GF	21
2.2 Proprietà della soluzione discreta	24
2.3 Regolarità ed estensione di S_t a $\overline{D(F)}$	27
2.3.1 Proprietà di stabilità	31
3 Sticky Particle System	35
3.1 Proprietà di \mathcal{K}	38
3.2 Il flusso gradiente come soluzione del SPS	44
3.2.1 Descrizione nel caso di misura iniziale discreta	44
3.2.2 SPS discreto è un flusso gradiente	46
3.3 Approssimazione con dati discreti	48
3.4 Condizioni di entropia	53
3.4.1 Vanishing viscosity	53
3.4.2 Soluzione entropica di SPS	54
3.5 SPS e leggi di conservazione scalare	55
A Teoria della misura	57
B Curve assolutamente continue in spazi di Hilbert	61
C Inviluppo convesso di una funzione	63

D	Equazione di Hamilton-Jacobi	67
D.1	Soluzioni viscosse di Hamilton-Jacobi	67
D.2	Prima e seconda formula di Hopf	68

Introduzione

In questa tesi tratteremo alcuni aspetti delle applicazioni della teoria del trasporto ottimo di massa a problemi d'evoluzione. La tesi è divisa in tre parti.

La prima è dedicata alla teoria del trasporto ottimo di massa, dove studieremo in particolare la situazione in cui lo spazio ambiente è \mathbb{R} ; in questo caso lo spazio delle misure di probabilità aventi momento secondo finito, munito della distanza di Wasserstein, risulterà essere isometricamente isomorfo al cono \mathcal{K} delle funzioni essenzialmente crescenti nello spazio di Hilbert $L^2(0, 1)$.

Questa caratterizzazione permette di studiare problemi relativi a misure e loro proprietà metriche, tramite lo studio del cono \mathcal{K} : l'esempio più interessante è forse quello dei flussi gradiente, introdotti nel secondo capitolo. Essendo possibile una definizione che dipende unicamente dalla distanza, flussi gradiente nello spazio di Wasserstein (che è uno spazio metrico, a priori senza strutture ulteriori) si possono tradurre dunque in flussi gradiente in $L^2(0, 1)$, molto più semplici da studiare grazie alla natura Hilbertiana di quest'ultimo spazio.

L'applicazione trattata nel terzo capitolo riguarda lo studio delle equazioni di Eulero con pressione nulla, attraverso la dinamica di adesione. Questo sistema di equazioni, chiamato anche *Sticky Particle System* (in breve SPS), si presenta nella forma:

$$\begin{cases} \partial_t \rho_t + \partial_x(v_t \rho_t) = 0 \\ \partial_t(v_t \rho_t) + \partial_x(v_t^2 \rho_t) = 0, \end{cases}$$

dove consideriamo come spazio ambiente proprio lo spazio di Wasserstein $\mathcal{P}_2(\mathbb{R})$. Nel caso di dati iniziali discreti, il sistema si evolve come un insieme di particelle libere di muoversi; nel momento in cui avviene un urto, esso è totalmente anelastico, quindi le particelle restano unite (da cui l'aggettivo *sticky*).

Un'accurata analisi di \mathcal{K} permette, partendo dalle soluzioni del sistema discreto nel caso di particelle *sticky*, di costruire un semigruppò d'evoluzione in un opportuno spazio delle fasi, ammettendo una misura iniziale in $\mathcal{P}_2(\mathbb{R})$ e velocità iniziale in L^2 . Di fondamentale importanza è l'osservazione che questa evoluzione, nel caso del sistema discreto prima del primo urto, è una geodetica nello spazio di Wasserstein, ed è dunque, similmente a quanto accade per varietà Riemanniane, flusso gradiente per la funzione $-\frac{1}{2}W_2^2$, l'opposto della distanza di Wasserstein al quadrato, e si può dunque studiare come flusso gradiente in $L^2(0, 1)$, che ha proprietà molto buone di convergenza.

L'ultima parte è dedicata ad una interessante dualità notata da Brenier e Grenier tra le soluzioni di SPS e la soluzione di una legge di conservazione scalare associata ai dati iniziali, la cui dimostrazione sarà più chiara alla luce dei risultati ottenuti nel corso della tesi.

Capitolo 1

Trasporto ottimo in una dimensione

In questo capitolo affrontiamo il problema del trasporto ottimo in una dimensione, specializzandoci poi al caso in cui il costo dello spostamento è quadratico $c(x, y) = (x - y)^2$, più facile da studiare in quanto potremo associarci anche una struttura Hilbertiana che sarà comoda nel seguito.

Prima di enunciare il problema del trasporto ottimo abbiamo bisogno di alcune nozioni preliminari di teoria della misura.

1.1 Nozioni preliminari

Definiamo qui il concetto basilare per tutta la teoria cioè il trasporto, o push forward, di misure (enunciato in generalità tra spazi metrici):

Definizione 1.1. *Sia $T : X \rightarrow Y$ una mappa di Borel. Essa induce la mappa $T_{\#} : \mathcal{M}_+(X) \rightarrow \mathcal{M}_+(Y)$ così definita:*

$$T_{\#}\mu(E) = \mu(T^{-1}(E)) \quad \forall E \in \mathfrak{B}(Y).$$

È immediato verificare che effettivamente $T_{\#}\mu$ è una misura finita su Y . Si può inoltre notare che $\#$ è un funtore covariante nel senso che $(T \circ S)_{\#} = T_{\#} \circ S_{\#}$. In realtà per definire $T_{\#}\mu$ è sufficiente che T sia μ -misurabile, ma con una definizione del genere ci sarebbero dei problemi nella composizione di mappe: infatti la composizione di funzioni misurabili non è necessariamente misurabile, mentre la composizione di funzioni di Borel è di Borel.

Da questa definizione discende la formula di cambiamento di variabile:

Lemma 1.2 (Formula di cambiamento di variabile). *Sia φ di Borel positiva. Allora:*

$$\int_Y \varphi dT_{\#}\mu = \int_X \varphi \circ T d\mu.$$

Dimostrazione. Basta dimostrarlo per φ funzione caratteristica poiché l'enunciato segue per linearità e convergenza monotona. Ma se $\phi(y) = 1_E(y)$, con E di Borel, l'equazione diventa esattamente la definizione di $T_{\#}\mu$ data in precedenza. \square

1.1.1 Funzioni di distribuzione e pseudo-inversa

Data una misura $\mu \in \mathcal{P}(\mathbb{R})$, definisco $F_\mu : \mathbb{R} \rightarrow [0, 1]$ come la sua funzione di distribuzione, cioè $F_\mu(x) = \mu((-\infty, x])$. Ovviamente F_μ è crescente e continua a destra, per la σ -additività della misura.

Ora costruiamo una pseudo-inversa di F_μ che chiamiamo G_μ . Non possiamo sperare in una inversa vera e propria, poiché non sono assicurate né l'iniettività né la suriettività della F_μ (basta pensare a $F_{\delta_0}(x)$ che risulta essere la funzione di Heaviside). Tuttavia, sfruttando la monotonia possiamo definire per ogni $t \in (0, 1)$:

$$G_\mu(t) = \inf\{x | F_\mu(x) \geq t\},$$

che risulta essere una funzione crescente e continua a sinistra.

Lemma 1.3. *Data $\mu \in \mathcal{P}(\mathbb{R})$ e definendo F_μ e G_μ come sopra si ha:*

- (i) $F_\mu(G_\mu(t)) \geq t$, con uguaglianza sse $t \in \text{Im}(F_\mu)$;
- (ii) $G_\mu(F_\mu(x)) \leq x$, con uguaglianza sse $x \in \text{Im}(G_\mu)$;
- (iii) $(G_\mu)_\# \mathcal{L}^1 = \mu$.

Dimostrazione. (i) Questo punto discende dal fatto che F_μ è continua a destra. Infatti per definizione di inf, esistono $x_n \downarrow G_\mu(t)$ tali che $F_\mu(x_n) \geq t$, dunque per la continuità a destra della funzione di distribuzione si ha che

$$F_\mu(G_\mu(t)) \geq t$$

(in particolare l'estremo inferiore è un minimo). Se poi t_0 è nell'immagine di F_μ , cioè se esiste x_0 tale che $F_\mu(x_0) = t_0$, allora per definizione vale $G_\mu(t_0) \leq x_0$ e quindi, per la crescenza di F_μ , si ha

$$F_\mu(G_\mu(t_0)) \leq F_\mu(x_0) = t_0$$

il che, unito alla disuguaglianza di prima ci dà l'uguaglianza voluta.

- (ii) Dalla definizione di G_μ si ha $G_\mu(F_\mu(x)) = \inf\{y | F_\mu(y) \geq F_\mu(x)\} \leq x$.

(iii) Per dimostrare l'uguaglianza delle due misure, basta testarle su una famiglia che genera tutti i boreliani, ad esempio gli intervalli della forma $(-\infty, x]$. Quindi consideriamo

$$\tilde{F}(x) := G_{\mu\#} \mathcal{L}^1((-\infty, x]) = \mathcal{L}^1(G_\mu^{-1}((-\infty, x]))$$

ed è sufficiente dimostrare che $\tilde{F}(x) = \mu((-\infty, x]) = F_\mu(x)$. Ora, poiché G_μ è crescente e continua a sinistra, $G_\mu^{-1}((-\infty, x])$ è della forma $(0, t_0]$.

Quindi $\tilde{F}(x) = t_0 = \max\{t | G_\mu(t) \geq x\}$; in particolare $G_\mu(\tilde{F}(x)) \leq x$ e $G_\mu(t) > x$ se $t > t_0$. Dal punto (ii) deduciamo $F_\mu(x) \leq \tilde{F}(x)$; per la crescenza di F_μ e per il punto (i), si ha

$$F_\mu(x) \geq F_\mu(G_\mu(\tilde{F}(x))) \geq \tilde{F}(x)$$

e dunque, come volevamo, $F_\mu = \tilde{F}$. □

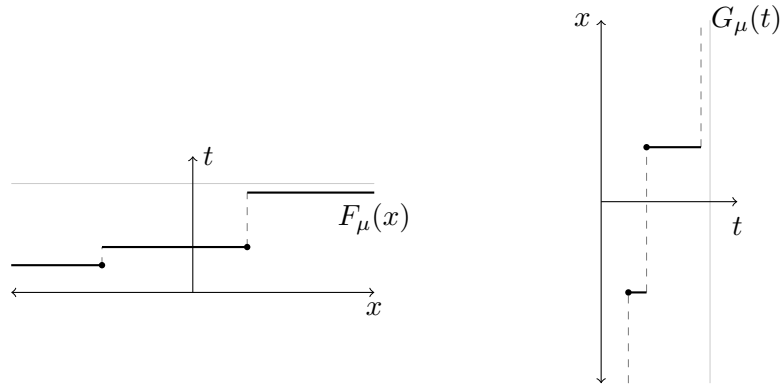


Figura 1.1: Esempio di funzione di distribuzione di una misura e sua pseudoinverta nel caso in cui μ sia una misura discreta; in generale i salti di F_μ corrispondono ad atomi della misura mentre i salti di G_μ corrispondono ad intervalli la cui misura è nulla.

Inoltre la corrispondenza tra misure di probabilità in \mathbb{R} e funzioni crescenti e continue a sinistra è biunivoca: infatti è facile verificare che se $G(t) : (0, 1) \rightarrow \mathbb{R}$ è una funzione crescente e continua a sinistra, allora chiamando $\mu = G_{\#} \mathcal{L}^1$ trovo che $G_\mu = G$ in $(0, 1)$. Dal punto (iii) del lemma precedente e dalla formula del cambiamento di variabili ottengo che

$$f(x) \in L^1(\mu) \iff f(G_\mu(t)) \in L^1(0, 1)$$

e in particolare alle misure μ con momento secondo finito corrispondono funzioni G_μ in $L^2(0, 1)$, e per quanto detto, c'è una corrispondenza biunivoca con il sottoinsieme delle funzioni che hanno un rappresentante crescente e continuo a sinistra, che chiameremo per comodità $L_c^2(0, 1)$, o per brevità, \mathcal{K} .

Nel seguito sarà importante una particolare funzione di trasporto tra 2 misure, l'unica a meno di insiemi μ -trascurabili che sia crescente.

Se imponiamo infatti che T sia crescente allora se $\nu = T_{\#} \mu$ si deve avere:

$$F_\mu(T(x)) = F_\nu(x)$$

formula che identifica la nostra T . Basterebbe ora avere un'inversa destra di F_μ e sarebbe una tale T , e un'inversa destra si ha quando $Im(F_\mu) = (0, 1)$ cioè quando la nostra misura è senza atomi. In questo caso si ha dunque:

$$T(x) = G_\mu(F_\nu(x))$$

1.2 Problema di Monge

Il problema di Monge, datato 1781, è stato storicamente il primo problema sul trasporto di massa e, in terminologia moderna risulta espresso così: siano dati X, Y sottoinsiemi di \mathbb{R}^2

e siano date $f : X \rightarrow [0, \infty)$ e $g : Y \rightarrow [0, \infty)$ (da considerarsi le distribuzioni di massa all'interno degli insiemi X e Y) tali che:

$$\int_X f(x)dx = \int_Y g(y)dy = 1.$$

Il problema di Monge consiste, informalmente, in: “trovare, tra tutte le mappe che trasportano f su g , quella per la quale il costo di trasporto è minimo”, dove il costo di trasporto è inteso in senso fisico e quindi proporzionale allo spostamento.

Riformulando il problema con le nostre notazioni consideriamo le misure di probabilità $\mu = f\mathcal{L}^2$ e $\nu = g\mathcal{L}^2$ e allora Monge chiede di minimizzare il funzionale:

$$\mathcal{C}(T) = \int_X |T(x) - x|d\mu(x).$$

In questo problema di minimo T varia tra tutte le funzioni che “trasportano f su g ” cioè tale che $T_{\#}\mu = \nu$. Ci sono due difficoltà evidenti in questo problema:

- non è ovvio a priori se esistano funzioni di trasporto T ammissibili;
- meno ovvio è che esistano T ottimali (infatti la prima dimostrazione completa è di Evans-Gangbo nel '96).

In generale, il problema di Monge astratto si può riformulare con costi qualunque, misure anche non assolutamente continue rispetto alla misura di Lebesgue o anche in spazi metrici qualunque, non necessariamente euclidei.

Problema di Monge astratto. Siano $\mu \in \mathcal{P}(X)$ e $\nu \in \mathcal{P}(Y)$. Inoltre sia $c : X \times Y \rightarrow [0, +\infty]$ una mappa di Borel. Ci proponiamo di studiare la quantità:

$$\inf_{T_{\#}\mu=\nu} \int_X c(x, T(x)) d\mu(x)$$

e di dire se è un minimo o meno, cioè se esistono mappe ottimali.

Illustriamo ora meglio il caso concreto che è quello che ci riguarderà più da vicino: quello in cui $X = Y = \mathbb{R}$.

Questo è il caso più completo in cui sostanzialmente si sa dire tutto. Infatti vale il seguente teorema, che dimostreremo nel corso di questo primo capitolo:

Teorema 1.4. Sia μ priva di atomi (cioè $\mu(\{x\}) = 0$ per ogni $x \in \mathbb{R}$). Allora valgono le seguenti:

- (1) $\exists!$ T crescente tale che $T_{\#}\mu = \nu$;
- (2) se $c(x, y) = \varphi(y - x)$ e φ è convessa, T è ottimale;

(3) se φ è strettamente convessa¹ allora T è l'unica mappa ottimale.

Osserviamo che questi risultati sono abbastanza “sharp” nel senso che:

- l'ipotesi che μ sia senza atomi è necessaria poiché se ad esempio $\mu = \delta_0$ e $\nu = 1_{[0,1]} \mathcal{L}^1$ allora non esistono mappe ammissibili, infatti $T_{\#} \delta_0 = \delta_{T(0)}$;
- la stretta convessità è necessaria per l'unicità: prendiamo ad esempio il problema di Monge e quindi $c(x, y) = |x, y|$, cioè $\varphi(t) = |t|$, caso in cui la funzione è convessa ma non strettamente convessa. Consideriamo le misure $\mu = 1_{[0,M]} \mathcal{L}^1$ e $\nu = 1_{[1,M+1]} \mathcal{L}^1$ e come mappe di trasporto prendiamo (nella definizione delle mappe, qui e in seguito, ovviamente è sufficiente descriverle sul supporto della misura di partenza):

$$T_1(x) = x + 1 \quad T_2(x) = \begin{cases} x + M & x \in [0, 1] \\ x & x \in [1, M] \end{cases}$$

È chiaro che queste mappe sono distinte e trasportano μ in ν . Si può anche facilmente vedere che hanno uguale costo, infatti per entrambe esso è esattamente M .

Inoltre è facile mostrare che questo è il costo ottimale: prendiamo una qualunque mappa T che trasporta μ in ν e, sfruttando la disuguaglianza $|T(x) - x| \geq T(x) - x$ otteniamo:

$$\begin{aligned} \mathcal{C}(T) &= \int_{\mathbb{R}} |T(x) - x| d\mu \geq \int_{\mathbb{R}} (T(x) - x) d\mu = \int_{\mathbb{R}} T(x) d\mu - \int_{\mathbb{R}} x d\mu = \\ &= \int_{\mathbb{R}} x d\nu - \int_{\mathbb{R}} x d\mu = \frac{1}{2}((M+1)^2 - 1 - M^2) = M. \end{aligned}$$

È dunque dimostrata la non unicità nel caso in cui φ è convessa ma non strettamente convessa.

Tuttavia questo teorema pone un limite, che è quello delle masse diffuse; noi vorremmo avere una teoria generale, che ci permetta di trattare anche casi in cui la misura di partenza contiene atomi. Motivati da questo passiamo al problema di Kantorovich, che al posto di mappe di trasporto si prefigura di avere piani di trasporto, nell'idea che posso prendere la massa in un punto e dividerla per farla andare in vari punti.

1.3 Problema di Kantorovich

Questo problema, posto da Kantorovich nel 1941, sembra che sia stato pensato dall'autore senza pensare alla precedente questione di Monge. Come prima, enunciamo la formulazione generale, nella quale si indaga la quantità:

$$\min_{\substack{\pi \text{ piano} \\ \text{di trasporto}}} \int_{X \times Y} c(x, y) d\pi(x, y)$$

¹Una funzione f si dice strettamente convessa se $f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$ ogniqualvolta $\lambda \in (0, 1)$ e $x \neq y$ (cioè esclusi i casi banali)

Qui la misura π è chiamata *piano di trasporto* se $(\pi_x)_\# \pi = \mu$ e $(\pi_y)_\# \pi = \nu$, ove $\pi_x(x, y) = x$ è la proiezione sulla prima coordinata mentre $\pi_y(x, y) = y$ è la proiezione sulla seconda coordinata. L'insieme dei piani di trasporto è dunque:

$$\Gamma(\mu, \nu) = \{\gamma \in \mathcal{P}(X \times Y) \mid (\pi_x)_\# \gamma = \mu, \quad (\pi_y)_\# \gamma = \nu\}$$

Come abbiamo detto prima, questo ci permette di distribuire la massa presente in un punto su tutto lo spazio Y . Per mostrare questo ci avvaliamo del teorema di disintegrazione, dimostrato in [AFP]:

Teorema. (di disintegrazione) *Sia dato un piano di trasporto γ , in $\mathcal{P}(X \times Y)$, tra μ e ν . Allora per μ -quasi ogni x in X esiste γ_x , misura di probabilità su Y , tale che $\gamma = \mu \otimes \gamma_x$, cioè:*

$$\gamma(E) = \int_X \gamma_x(\pi_y(E)) d\mu(x) \quad \forall E \text{ insieme } \gamma \text{ misurabile.}$$

Le misure γ_x rappresentano proprio il modo di dividere la massa presente nel punto x per spargerla su Y .

Elenchiamo subito i vantaggi, come motivazione del perché conviene studiare questo problema invece che quello di Monge:

- (1) $\Gamma(\mu, \nu) \neq \emptyset$, infatti ad esempio $\mu \times \nu \in \Gamma(\mu, \nu)$;
- (2) $\Gamma(\mu, \nu)$ è convesso (e compatto rispetto ad un'opportuna topologia debole) e la quantità da studiare è lineare in π , che ci consente una formulazione duale con tecniche di analisi convessa;
- (3) l'esistenza del minimo è "facile", per esempio esiste almeno un piano ottimale se $c(x, y)$ è s.c.i. in $X \times Y$.

Inoltre si può notare che le mappe di trasporto sono in un certo senso "incluse" nei piani di trasporto. Data una mappa di trasporto T infatti possiamo considerare la misura su $X \times Y$ data dal push-forward di μ secondo $Id \times T$, cioè $\gamma_T = (Id \times T)_\# \mu$. Infatti, dalla definizione di push-forward:

$$\int_{X \times Y} h(x, y) d\gamma(x, y) = \int_X h(x, T(x)) d\mu(x)$$

per ogni h che sia γ -misurabile. Ma allora, scegliendo come h funzioni di una sola variabile, otteniamo:

$$\begin{aligned} \int_{X \times Y} f(x) d\gamma(x, y) &= \int_X f(x) d\mu(x) \\ \int_{X \times Y} g(y) d\gamma(x, y) &= \int_X g(T(x)) d\mu(x) = \int_Y g(y) d\nu(y) \end{aligned}$$

e queste due uguaglianze ci dicono proprio che γ è un piano di trasporto.

Nel teorema di disintegrazione al piano di trasporto derivato dalla mappa $T(x)$ corrispondono delle misure disintegrate γ_x che sono delte di dirac in $T(x)$. Inoltre il piano di trasporto γ indotto da $T(x)$ è ovviamente una misura concentrata sul grafico di T ; osserviamo col seguente teorema che possiamo anche stabilire quando un piano di trasporto è indotto da una mappa:

Teorema 1.5. *Sia $\pi \in \Gamma(\mu, \nu)$ concentrata su un grafico π -misurabile Γ . Allora esiste T di Borel tale che $\pi = (Id \times T)_{\#}\mu$.*

Prima di dimostrare questo teorema però ci serve richiamare un lemma di teoria della misura, dimostrato in appendice:

Lemma 1.6. *Sia $\mu \in \mathcal{M}_+(X)$. Preso un boreliano E e un $\varepsilon > 0$ esistono un compatto K e un aperto A con $K \subseteq E \subseteq A$ tali che $\mu(A \setminus K) \leq \varepsilon$*

Dimostrazione. (Teorema 1.5) Per definizione di insieme misurabile, esistono $\Gamma_0 \subseteq \Gamma \subseteq \Gamma_1$ con $\Gamma_0, \Gamma_1 \in \mathfrak{B}(X \times Y)$ tali che $\pi(\Gamma_1 \setminus \Gamma_0) = 0$. Prendiamo ora una successione numerabile crescente di compatti $K_n \subseteq \Gamma_0$ tali che $\pi(\Gamma_0 \setminus \bigcup_n K_n) = 0$ (si usa il lemma precedente).

Sia $C_n = \pi_x(K_n)$; allora per ogni $x \in C_n$ esiste una sola $y = f_n(x)$ tale che $(x, y) \in K_n$ poiché K_n è sottoinsieme di un grafico. La f_n così costruita è continua su C_n perché ha il grafico chiuso (che è K_n). Inoltre le f_n si incollano bene dal fatto che $K_n \subseteq K_{n+1}$. Questo ci permette di definire:

$$T : \bigcup_n C_n \rightarrow Y \quad := \quad T|_{C_n} \equiv f_n.$$

La funzione così costruita risulta essere di Borel e la posso estendere a tutto X ponendola costantemente uguale a y_0 nei punti che non sono in $\bigcup_n C_n$.

Ma ora abbiamo che $T(x) = y$ in $\bigcup_n K_n$ e quindi π -quasi ovunque in $X \times Y$. Questo ci permette di concludere che per ogni φ che sia di Borel:

$$\begin{aligned} \int_{X \times Y} \varphi(x, y) d\pi(x, y) &= \int_{X \times Y} \varphi(x, T(x)) d\pi(x, y) = \int_X \varphi(x, T(x)) d\mu(x) = \\ &= \int_{X \times Y} \varphi(x, y) d(Id \times T)_{\#}\mu(x, y) \end{aligned}$$

Cioè che effettivamente $\pi = (Id \times T)_{\#}\mu$ per una mappa boreliana T . □

1.3.1 Problema di Kantorovich in \mathbb{R}

Daremo ora un teorema generale per l'ottimalità del piano di trasporto. Faremo poi vedere che in particolare $(G_\mu, G_\nu)_{\#}\mathcal{L}^1$ è ottimale per costi convessi e in particolare per tutti i costi della forma $c(x, y) = |x - y|^p$ con $p \geq 1$. Se inoltre la misura μ è priva di atomi questo piano di trasporto sarà proprio quello indotto dalla mappa di trasporto vista in precedenza $T(x) = G_\mu(F_\nu(x))$ che dunque sarà ottimale nel problema più generale di Kantorovich e quindi anche nel problema di Monge. Diamo però prima una definizione, che servirà in seguito:

Definizione 1.7. *Data $c(x, y)$ una funzione a valori reali si dice che un insieme $A \subseteq X \times Y$ è c -ciclicamente monotono se per ogni n -upla di punti $(x_i, y_i) \in A$ e ogni permutazione $\sigma \in \mathcal{S}_n$*

vale:

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

Nella definizione possiamo anche richiedere che la disuguaglianza sia vera anche solo per permutazioni cicliche, le altre si ottengono da queste come composizione ‘disgiunta’ e quindi la disuguaglianza sarà vera *a fortiori*. Possiamo enunciare ora il teorema di ottimalità. Chiamiamo d’ora in poi $\Gamma_0(\mu_0, \mu_1)$ l’insieme dei piani ottimali e $\mathcal{C}_c(\mu_0, \mu_1)$ il minimo costo cioè

$$\mathcal{C}_c(\mu_0, \mu_1) = \int_{X \times Y} c(x, y) d\pi(x, y) \quad \forall \pi \in \Gamma_0(\mu_0, \mu_1).$$

Teorema 1.8. *Sia $\gamma \in \Gamma(\mu_0, \mu_1)$, dove il costo preso in considerazione è continuo, finito e inoltre vale la maggiorazione $c(x, y) \leq f(x) + g(y)$ con $f \in L^1(\mu_0)$ e $g \in L^1(\mu_1)$. Se γ è concentrata in un insieme c -ciclicamente monotono, allora γ è ottimale.*

Dimostrazione. Consideriamo \mathcal{C} , l’insieme su cui γ è concentrata, che per ipotesi è c -ciclicamente monotono; usando il lemma di Ulam (richiamato in appendice) possiamo supporre che \mathcal{C} sia σ -compatto. È facile vedere che μ_0 è concentrata in $\tilde{X} = \pi_x(\mathcal{C})$ e μ_1 è concentrata similmente in $\tilde{Y} = \pi_y(\mathcal{C})$. Ora mi basta trovare due funzioni $\varphi(x)$ e $\psi(y)$ tali che $\varphi(x) + \psi(y) \leq c(x, y)$ su tutto $\tilde{X} \times \tilde{Y}$, dove l’uguaglianza però vale almeno su tutto \mathcal{C} . A questo punto avrei finito poichè preso qualunque altro piano di trasporto γ_1 , (che sarà sicuramente concentrato su $\tilde{X} \times \tilde{Y}$), avrei:

$$\begin{aligned} \int_{X \times Y} c(x, y) d\gamma_1(x, y) &= \int_{\tilde{X} \times \tilde{Y}} c(x, y) d\gamma_1(x, y) \geq \\ &\geq \int_{\tilde{X} \times \tilde{Y}} (\varphi(x) + \psi(y)) d\gamma_1(x, y) = \\ &= \int_{\tilde{X}} \varphi(x) d\mu_0(x) + \int_{\tilde{Y}} \psi(y) d\mu_1(y) = \\ &= \int_{\tilde{X} \times \tilde{Y}} (\varphi(x) + \psi(y)) d\gamma(x, y) = \\ &= \int_{X \times Y} c(x, y) d\gamma(x, y) \end{aligned}$$

da cui l’ottimalità; attenzione che per avere che l’integrale della somma è uguale alla somma degli integrali sono necessarie la misurabilità e l’integrabilità di φ e ψ ; in realtà ci bastano integrabilità di φ^+ e ψ^+ per evitare una forma indeterminata del tipo $\infty - \infty$.

Costruiamo la funzione $\varphi(x)$; fisso un punto $(x_0, y_0) \in \mathcal{C}$ tale che $f(x_0)$ e $g(y_0)$ siano entrambi finiti e definisco:

$$\varphi(x) := \inf \{ c(x, y_n) - c(x_n, y_n) + \dots + (x_1, y_0) - c(x_0, y_0) \}$$

Dove l’estremo inferiore è preso su tutte le collezioni di punti $(x_i, y_i) \in \mathcal{C}$. Innanzitutto è facile verificare che la funzione risultante è s.c.s. e dunque misurabile. Ora ho $\varphi(x_0) \geq 0$ proprio perchè \mathcal{C} è c -ciclicamente monotono, inoltre prendendo una sola coppia di punti

$(x_1, y_1) = (x_0, y_0)$ risulta $\varphi(x_0) \leq 0$ e dunque $\varphi(x_0) = 0$.

Nella definizione di $\varphi(x)$ supponiamo ora di fissare sempre l'ultimo punto della n -upla come $(x', y') \in \mathcal{C}$. Allora risulterà:

$$\varphi(x) \leq c(x, y') - c(x', y') + c(x', y_{n-1}) - c(x_{n-1}, y_{n-1}) + \dots + (x_1, y_0) - c(x_0, y_0)$$

prendendo poi l'estremo inferiore al variare degli altri punti, riconosciamo all'interno del secondo membro la definizione di $\varphi(x')$ e quindi ottengo:

$$\varphi(x) \leq c(x, y') - c(x', y') + \varphi(x').$$

Le funzioni che soddisfano questo tipo di disuguaglianze sono dette c -concave. Considerando ora $(x', y') = (x_0, y_0)$ ottengo in particolare $\varphi(x) \leq c(x, y_0) - c(x_0, y_0) + \varphi(x_0) \leq a(x) + b(y_0) - c(x_0, y_0)$ che ci dà dunque l'integrabilità di φ^+ , inoltre scegliendo $x = x_0$ ottengo $\varphi(x') > -\infty$ quando $x' \in \tilde{X}$.

Definisco poi $\psi(y) := \inf_{x \in \tilde{X}} (c(x, y) - \varphi(x))$ in modo che, per definizione, $\psi + \varphi \leq c$ su $\tilde{X} \times \tilde{Y}$. Come prima ψ , essendo un estremo inferiore di funzioni continue, risulta essere s.c.s e quindi misurabile. Prendo ora $(x_1, y_1) \in \mathcal{C}$ e noto che per ogni $x \in \tilde{X}$ ho:

$$c(x_1, y_1) - \varphi(x_1) \leq c(x, y_1) - \varphi(x)$$

per la disuguaglianza di c -concavità per φ . Quindi l'estremo inferiore nella definizione di $\psi(y_1)$ è in realtà un minimo raggiunto proprio da $x = x_1$ e dunque $\psi(y_1) + \varphi(x_1) = c(x_1, y_1)$. Inoltre nella definizione della ψ prendendo x_0 come punto *test* ottengo $\psi(y) \leq c(x_0, y) - \varphi(x_0) \leq b(y) + a(x_0) - \varphi(x_0)$ e dunque anche ψ^+ risulta integrabile. \square

Teorema 1.9. *Consideriamo ora $X = Y = \mathbb{R}$. Sia $\gamma = (G_{\mu_0}, G_{\mu_1})_{\#} \mathcal{L}^1$. Allora γ risulta concentrata sull'immagine $\mathcal{C} = (G_{\mu_0}, G_{\mu_1})([0, 1])$ che è c -monotono se $c(x, y) = \varphi(x - y)$ con φ convessa e continua. In particolare $\gamma \in \Gamma_0(\mu_0, \mu_1)$.*

Dimostrazione. Per definizione se $B \cap \mathcal{C} = \emptyset$ allora $\gamma(B) = \mathcal{L}^1((G_{\mu_0}, G_{\mu_1})^{-1}(B)) = \mathcal{L}^{-1}(\emptyset) = 0$. Successivamente consideriamo un costo $c(x, y) = \varphi(x - y)$; vogliamo ora dimostrare che \mathcal{C} è c -monotono. Per far vedere questo si prendono n punti $(x_i, y_i) = (G_{\mu_0}(t_i), G_{\mu_1}(t_i))$ nel grafico e si mostra che, per ogni permutazione $\sigma \in S_n$:

$$\sum_{i=1}^n \varphi(x_i - y_i) \leq \sum_{i=1}^n \varphi(x_i - y_{\sigma(i)}).$$

Dimostriamolo per induzione: il passo base è $n = 2$. Supponendo che $t_1 > t_2$ otterrei $x_1 \geq x_2$ e similmente $y_1 \geq y_2$ per la crescenza della G . Ora chiamo $x_i - y_i = z_i$ e $x_i - y_{\sigma(i)} = w_i$. So ora che $z_1 + z_2 = w_1 + w_2$; se $w_1 = w_2$ allora se $x_1 > x_2$ necessariamente $y_2 > y_1$ che è assurdo e quindi ho $x_1 = x_2$ e $y_1 = y_2$. se invece $w_1 \neq w_2$ allora so che esiste un t tale che $z_1 = tw_1 + (1-t)w_2$ e $z_2 = tw_2 + (1-t)w_1$. inoltre vale per la disuguaglianza di riarrangiamento:

$$\begin{aligned} w_1^2 + w_2^2 &= x_1^2 + x_2^2 + y_1^2 + y_2^2 - 2(x_1 y_{\sigma(1)} + x_2 y_{\sigma(2)}) \geq \\ &\geq x_1^2 + x_2^2 + y_1^2 + y_2^2 - 2(x_1 y_1 + x_2 y_2) = z_1^2 + z_2^2 \end{aligned}$$

Ma $z_1^2 + z_2^2 = w_1^2 + w_2^2 - 2t(1-t)(w_1 - w_2)^2$ e dunque $0 \leq t \leq 1$ da cui posso concludere la tesi:

$$\begin{aligned} \varphi(w_1) + \varphi(w_2) &= t\varphi(w_1) + (1-t)\varphi(w_2) + t\varphi(w_2) + (1-t)\varphi(w_1) \geq \\ &\geq \varphi(tw_1 + (1-t)w_2) + \varphi(tw_2 + (1-t)w_1) = \varphi(z_1) + \varphi(z_2) \end{aligned}$$

Ora il passo induttivo: Supponiamo la tesi sia vera per $k \leq n$ e vogliamo dimostrarlo per $k = n + 1$. Prendiamo la nostra $n + 1$ -upla e assumiamo che t_{n+1} sia il massimo dei tempi considerati; data la nostra $\sigma \in S_{n+1}$ costruiamo una $\hat{\sigma} \in S_n$ in questo modo $\hat{\sigma}(i) = \sigma(i)$ se $\sigma(i) \neq n + 1$ e $\hat{\sigma}(i) = \sigma(\sigma(i))$ altrimenti (in pratica saltiamo il numero $n + 1$). Applichiamo la nostra ipotesi induttiva ottenendo:

$$\sum_{i=1}^n \varphi(x_i - y_i) \leq \sum_{i=1}^n \varphi(x_i - y_{\hat{\sigma}(i)}).$$

Ora applico la mia ipotesi induttiva con $n = 2$ alle coppie $x_{n+1} \geq x_{\sigma^{-1}(n+1)}$ e $y_{n+1} \geq y_{\sigma(n+1)}$ ottenendo (chiamo $k = \sigma^{-1}(n + 1)$):

$$\varphi(x_{n+1} - y_{n+1}) + \varphi(x_k - y_{\hat{\sigma}(k)}) \leq \varphi(x_k - y_{\sigma(k)}) + \varphi(x_{n+1} - y_{\sigma(n+1)})$$

che insieme alla disugaglianza di prima conclude anche il caso $k = n + 1$. \square

Vogliamo ora considerare il costo $c(x, y) = (x - y)^2$ che sicuramente soddisfa le ipotesi del teorema 1.9. Notiamo che $(x - y)^2 \leq 2x^2 + 2y^2$ e dunque se μ e ν sono misure in $\mathcal{P}_2(\mathbb{R})$, esse soddisfano le ipotesi del teorema 1.8. Considerando quindi tutto otteniamo che il piano di trasporto γ definito poco fa è ottimale per due misure qualunque in $\mathcal{P}_2(\mathbb{R})$ se il costo del trasporto è $(x - y)^2$.

Questo costo minimo può essere considerato come una quantità che misura quanto sono diverse μ e ν . La chiameremo

$$W_2^2(\mu, \nu) = \int_{\mathbb{R}^2} (x - y)^2 d\gamma(x, y) = \int_0^1 (G_\mu(t) - G_\nu(t))^2 dt.$$

Analizziamo meglio questa quantità: prendendo le radici quadrate, otteniamo $W_2(\mu, \nu) = \|G_\mu - G_\nu\|_2$ e dunque W_2 è una distanza, poiché $\|\cdot\|_2$ è una norma. Possiamo munire $\mathcal{P}_2(\mathbb{R})$ di questa distanza per renderlo uno spazio metrico:

Teorema 1.10. *La mappa $G: (\mathcal{P}_2(\mathbb{R}), W_2) \rightarrow \mathcal{K}, \|\cdot\|_2$, definita da*

$$G: \mu \mapsto G_\mu$$

è un'isomorfismo tra spazi metrici.

Dimostrazione. La biiettività tra i due insiemi l'abbiamo già vista nella prima sezione di questo capitolo, mentre il fatto che sia una isometria discende dal fatto che, per l'osservazione fatta in precedenza: $W_2(\mu, \nu) = \|G_\mu - G_\nu\|_2$. \square

Notiamo che, allo stesso modo, possiamo stabilire un isomorfismo di spazi metrici tra $(\mathcal{P}_p(\mathbb{R}), W_p)$ (definito allo stesso modo dello spazio di Wasserstein ma con il costo $c(x, y) = |x - y|^p$) e $(L_c^p(0, 1), \|\cdot\|_p)$ purchè $1 \leq p < +\infty$.

Capitolo 2

Flussi gradiente in spazi di Hilbert

I flussi gradiente sono delle particolari equazioni differenziali ordinarie, del tipo:

$$\begin{cases} \dot{x}(t) = -\nabla F(x(t)) \\ x(0) = \bar{x} \end{cases}$$

dove $x(t)$ è una curva in uno spazio di Hilbert H , mentre F è una funzione data da questo spazio di Hilbert a valori reali. Sono chiamate anche curve di massima pendenza: immaginiamo infatti che $F(x)$ rappresenti l'altitudine del terreno nel punto x , e supponiamo poi di voler raggiungere il punto più basso, allora è conveniente (almeno localmente), in ogni punto seguire il percorso dove la pendenza è massima e la condizione è proprio l'equazione differenziale scritta prima, a meno di riparametrazioni.

Questa definizione ha senso solo nel momento in cui F è derivabile, tuttavia si può generalizzare con una nozione più debole di differenziale, il *sottodifferenziale*, che ad esempio abbiamo quando F è una funzione convessa, anche non regolare.

Definizione 2.1. *Sia data $F : H \rightarrow \mathbb{R} \cup \{+\infty\}$ convessa e semicontinua inferiormente. Indichiamo con $D(F)$ l'insieme degli $x \in H$ dove la funzione assume valore finito; per ogni $x \in D(F)$ si definisce*

$$\partial F(x) = \{p \in H \mid F(y) \geq F(x) + \langle p, y - x \rangle \quad \forall y \in H\},$$

il *sottodifferenziale* di F nel punto x . Indicheremo in seguito con $D(\partial F)$ l'insieme dei punti x per cui $\partial F(x) \neq \emptyset$.

Una proprietà interessante e banalmente vera di $\partial F(x)$ è il fatto di essere un insieme convesso chiuso; in particolare esiste, ed è unico, il vettore di minima norma che indicheremo con $\partial^o F(x)$. Ricordiamo un'ulteriore caratterizzazione del sottodifferenziale, cioè $p \in \partial F(x)$ se e solo se

$$\liminf_{t \downarrow 0} \frac{F(x + tv) - F(x) - t\langle v, p \rangle}{t} \geq 0 \quad \forall v \in H.$$

Senz'altro se $p \in \partial F(x)$ allora quel limite è sempre maggiore o uguale a zero in quanto limite di quantità non negative. Supponiamo invece che p soddisfi questa disuguaglianza; per

qualunque y , considero $v = y - x$. So che, per la convessità, vale

$$F(x + v) - F(x) \geq \frac{F(x + tv) - F(x)}{t} \quad \forall 0 < t < 1$$

Sottraendo ad entrambi i membri la quantità $\langle v, p \rangle$, e facendo il limite per t che tende a 0, troviamo

$$F(x + v) - F(x) - \langle v, p \rangle \geq \liminf_{t \downarrow 0} \frac{F(x + tv) - F(x) - t\langle v, p \rangle}{t} \geq 0$$

e dunque effettivamente, come volevamo dimostrare:

$$F(y) \geq F(x) + \langle y - x, p \rangle \quad \forall y \in H.$$

Ritornando al problema iniziale, possiamo anche rilassare le ipotesi sulla curva, richiedendo solo l'assoluta continuità invece che la derivabilità ovunque.

Definizione 2.2. Una curva $u : I \rightarrow H$ si dice assolutamente continua, e si indica $u \in \mathcal{AC}(I; H)$, se esiste una funzione $g \geq 0$ in $L^1(I; \mathbb{R})$ tale che

$$\|u(b) - u(a)\| \leq \int_a^b g(t) dt \quad \forall a, b \in I, \quad a \leq b$$

Se $H = \mathbb{R}$ è noto che una funzione è assolutamente continua se e solo se è derivabile quasi ovunque e vale il teorema fondamentale del calcolo integrale. Questa proprietà si mantiene valida per uno spazio di Hilbert qualsiasi, quindi se $u \in \mathcal{AC}(I; H)$ allora è derivabile fortemente quasi ovunque e la derivata $u' \in L^1(I, H)$ è tale che $u(b) - u(a) = \int_a^b u'(t) dt$ (in particolare nella definizione di assoluta continuità possiamo prendere $g = \|u'\|$). La dimostrazione di questa proprietà è riportata in appendice.

Se poi richiediamo una integrabilità maggiore della funzione g , ad esempio $g \in L^p$, allora otteniamo lo spazio $\mathcal{AC}^p(I; H)$. Questo oggetto, per $p > 1$ si comporta bene con successioni convergenti puntualmente, infatti:

Teorema 2.3. Sia $\{u_n\}$ una successione di curve in $\mathcal{AC}^p((0, 1); H)$ (con $p > 1$), tali che $\|u'_n\|_p \leq C$. Allora se $u_n(t) \rightarrow u(t)$ per ogni $t \in (0, 1)$, anche la curva u appartiene a $\mathcal{AC}^p((0, 1); H)$

Dimostrazione. Sia q l'esponente coniugato di p . La condizione $\|u'_n\|_p \leq C$ assicura che $\{|u'_n|\}$ è una successione limitata in $L^p(0, 1) = (L^q(0, 1))^*$ e, a meno di estrarre sottosuccessioni, possiamo assumere sia debole- $*$ convergente ad una $g \in L^p$. Considerando ora la disuguaglianza

$$\|u_n(b) - u_n(a)\| \leq \int_a^b |u'_n(t)| dt = \langle \mathbb{1}_{(a,b)}, |u'_n| \rangle$$

e passando al limite ottengo

$$\|u(b) - u(a)\| \leq \langle \mathbb{1}_{(a,b)}, g \rangle = \int_a^b g(t) dt.$$

Cioé, come voluto, $u \in \mathcal{AC}^p((0, 1); H)$.

□

Il problema in più generalità dunque diviene $u' \in -\partial F(u)$ per quasi ogni $t > 0$. Enunciamo subito il teorema centrale che ci proponiamo di dimostrare in questo capitolo:

Teorema 2.4. *Data una funzione convessa s.c.i. $F : H \rightarrow \mathbb{R} \cup \{+\infty\}$ e un punto $x_0 \in \overline{D(F)}$ esiste ed è unica una curva $u \in \mathcal{AC}_{loc}((0, +\infty); H)$ tale che sia soddisfatta l'inclusione differenziale*

$$\dot{u}(t) \in -\partial F(u(t))$$

per quasi ogni $t > 0$ e valga la condizione iniziale $\lim_{t \rightarrow 0} u(t) = x_0$.

Prima di dimostrare questo teorema vogliamo illustrare un paio di esempi di flussi gradiente, per capire l'importanza dell'argomento:

Esempio 1. (*Equazione del calore*) Sia $H = L^2(\mathbb{R}^n)$ e sia F definita in questo modo:

$$F(u) = \begin{cases} \frac{1}{2} \int_{\mathbb{R}^n} |Du(x)|^2 dx & \text{se } u \in H^1(\mathbb{R}^n) \\ +\infty & \text{altrimenti.} \end{cases}$$

Questa funzione soddisfa le ipotesi del teorema poiché è convessa e semicontinua inferiormente, per classici risultati sugli spazi di Sobolev. Inoltre il dominio di F , cioè H^1 , è denso in L^2 e dunque la condizione iniziale può essere qualunque.

Per il teorema 2.4 quindi abbiamo esistenza ed unicità per il flusso gradiente associato ad F . Vogliamo ora scoprire qualcosa di più su questo flusso gradiente: per farlo analizziamo il sottodifferenziale di F . Prendiamo allora u e v nel dominio di F , dunque in H^1 , e assumiamo che p sia un elemento di $\partial F(u)$. Allora vale:

$$F(u + \varepsilon v) \geq F(u) + \varepsilon \langle p, v \rangle$$

Mandando ε a zero (e considerandolo sia positivo che negativo), svolgendo i calcoli otteniamo:

$$\langle Du, Dv \rangle = \langle p, v \rangle,$$

dunque, per definizione, p non è altro che $-\Delta u$, il laplaciano distribuzionale, che esiste se e solo se $u \in H^2$. In particolare abbiamo che $D(\partial F) = H^2$ e per ogni $u \in H^2$ il sottodifferenziale consta di un unico elemento, $-\Delta u$. L'esistenza di un flusso gradiente si riduce dunque all'esistenza di una soluzione a:

$$\begin{cases} u_t(x, t) = \Delta u(x, t) & \text{per quasi ogni } t \geq 0 \\ u(x, 0) = f(x), \end{cases}$$

che è proprio l'equazione del calore. Grazie ai risultati di regolarità presenti a fine capitolo si ha $u(\cdot, t) \in H^2$ per ogni $t > 0$ e con ulteriori raffinamenti, grazie alla linearità dell'equazione, si può facilmente dedurre $u \in H^k$ per ogni k e dunque $u \in C^\infty$, come si sa dalla teoria classica.

Consideriamo ora un altro flusso gradiente, considerando la stessa F , ma come spazio ambiente scegliamo $H_0^1(B(0, 1))$, dove prendiamo il prodotto scalare:

$$\langle u, v \rangle = \int_{|x| \leq 1} Du(x) \cdot Dv(x) dx.$$

Seguendo i medesimi calcoli di prima, si giunge a $\partial F(u) = u$ e quindi l'equazione risultante corrispondente al flusso gradiente è

$$\begin{cases} u_t(x, t) = -u(x, t) & \text{per quasi ogni } t \geq 0 \\ u(x, 0) = f(x), \end{cases}$$

ed è banalmente risolta da $u(x, t) = e^{-t}f(x)$.

Questo esempio mostra che, oltre alla scelta della funzione, è molto importante anche lo spazio considerato e il relativo prodotto scalare, avendo come risultati equazioni d'evoluzione completamente differenti. Infatti se ad esempio prendiamo sempre lo spazio H_0^1 ma corredato del prodotto scalare (equivalente al precedente):

$$\langle u, v \rangle_C = \int_{|x| \leq 1} Du \cdot Dv \, dx + C \int_{|x| \leq 1} u \cdot v \, dx \quad C \geq 0$$

l'equazione d'evoluzione risultante è:

$$u_t = (\Delta - C)^{-1} \Delta u.$$

Esempio 2. Come vedremo nel seguito, in realtà esistono definizioni puramente metriche di flusso gradiente, ad esempio la formulazione (EVI) o quella (EDI) (vedi [DGMT], [DG], [A]). Questo permette di estendere la nozione di flusso gradiente anche a spazi metrici generici, come ad esempio lo spazio di Wasserstein $(\mathcal{P}_2(\mathbb{R}^n), W_2)$. La trattazione formale che segue è tratta dalle note [San10].

Dalla teoria del trasporto ottimo in \mathbb{R}^n , sappiamo che, date μ, ν misure assolutamente continue rispetto alla misura di Lebesgue, esiste una mappa semiconcava ϕ tale che $x - \nabla \phi(x) = T(x)$ è la mappa di trasporto ottimo tra μ e ν . Questa mappa, detta potenziale di Kantorovich, svolge un po' il ruolo di "spostamento" duale, infatti si ha che:

$$-\frac{1}{2} \frac{d}{d\varepsilon} W_2^2(\mu + \varepsilon \chi, \nu)|_{\varepsilon=0} = \int \phi d\chi$$

per ogni misura χ per cui esiste ε_0 tale che $\mu + \varepsilon \chi$ rimane una misura di probabilità positiva, per $\varepsilon \leq \varepsilon_0$. Allo stesso modo posso definire il "differenziale" δF di un funzionale sulle misure, come quella funzione che soddisfa:

$$\frac{d}{d\varepsilon} F(\mu + \varepsilon \chi)|_{\varepsilon=0} = \int \delta F d\chi,$$

dove χ soddisfa le stesse assunzioni di prima. Seguendo poi la costruzione della (EVI) mediante spostamenti discreti descritta dopo:

$$\mu_{n+1}^\tau = \arg \min_{\rho \in \mathcal{P}_2} \left\{ F(\rho) + \frac{W_2^2(\rho, \mu_n^\tau)}{2\tau} \right\}.$$

Immaginando di fare una "variazione prima", cioè ponendo la derivata uguale a zero, otteniamo:

$$\int \left(\frac{\phi_n}{\tau} + \delta F \right) d\chi = 0,$$

che, notando che χ deve avere media nulla, permette di dedurre che:

$$-\frac{\nabla\phi_n}{\tau} = \nabla\delta F$$

Ora la quantità a sinistra la possiamo considerare come una velocità, in quanto è esattamente

$$-\frac{\nabla\phi_n}{\tau}(x) = \frac{T_n(x) - x}{\tau}$$

dove T_n è la mappa di trasporto tra μ_{n+1}^τ e μ_n^τ . Ricordando poi l'equazione di continuità:

$$v_t = \lim_{h \rightarrow 0} \frac{T_t^{t+h} - Id}{h} \implies \partial_t \mu_t + \nabla \cdot (v_t \mu_t) = 0$$

dove T_t^{t+h} è la mappa che trasporta μ_t in μ_{t+h} , tutto ciò suggerisce che μ_t è flusso gradiente di F secondo la distanza di Wasserstein se soddisfa la seguente equazione:

$$\partial_t \mu_t - \nabla \cdot (\mu_t \nabla \delta F(\mu_t)) = 0.$$

Ci sono molte possibili scelte per il funzionale F , ma quello probabilmente più interessante è quello dell'energia interna:

$$F(\rho) = \begin{cases} \int_{\mathbb{R}^n} f(\lambda) dx & \text{se } \rho = \lambda \mathcal{L} \\ +\infty & \text{altrimenti} \end{cases},$$

nel quale caso si verifica che $\delta F(\lambda \mathcal{L}) = f'(\lambda)$. Scegliendo $f(y) = y \ln(y)$, si ottiene quella che è chiamata entropia rispetto alla misura di Lebesgue. Calcolando la derivata, otteniamo $f'(y) = \ln(y) + 1$, dunque in particolare $\nabla f'(\lambda) = \frac{\nabla \lambda}{\lambda}$, da cui l'equazione evolutiva per il flusso gradiente in questo caso diviene, formalmente:

$$\partial_t \mu_t = \nabla \cdot (\mu_t \nabla \delta F(\mu_t)) = \nabla \cdot \left(\mu_t \frac{\nabla \mu_t}{\mu_t} \right) = \Delta \mu_t,$$

cioè nuovamente l'equazione del calore. Notiamo che questo flusso gradiente si può costruire in qualsiasi spazio $\mathcal{P}(X)$ con X spazio metrico, permettendo dunque di generalizzare la nozione di *Heat Flow* anche in spazi metrici.

Scegliendo in modo diverso il funzionale F , possiamo descrivere come flussi gradiente nello spazio di Wasserstein le soluzioni dell'equazione di Fokker-Plank, o anche più complicate equazioni differenziali evolutive di tipo non locale e non lineare (vedi [AGS08]).

Ritorniamo ora alla dimostrazione del teorema 2.4: l'approccio classico di Brezis (cit.) consiste nel sostituire l'operatore multivoco ∂F con uno più regolare $(\partial F)_\tau$, per poi dimostrare che esistono le soluzioni u_τ approssimate e che tendono ad una soluzione dell'equazione di partenza. Invece qui si userà un approccio basato su una approssimazione discreta, lo schema di Eulero implicito, che ha il vantaggio di essere più stabile in quanto la mappa ricorsiva risulta 1-lipschitziana; inoltre è correlato ad un problema variazionale.

La dimostrazione verrà articolata nei seguenti punti:

- (i) dimostrazione dell'unicità;
- (ii) costruzione della soluzione discreta e equivalenza tra essere flusso gradiente (GF) e risolvere la disuguaglianza d'evoluzione variazionale (EVI);
- (iii) proprietà delle soluzioni discrete e convergenza ad \bar{u} ;
- (iv) \bar{u} è in effetti soluzione di (GF);
- (v) estensione della soluzione e regolarità.

Unicità. Siano date $u, v \in \mathcal{AC}_{loc}((0, +\infty); H)$ che soddisfano l'inclusione differenziale del teorema. Considero $f(t) := \frac{1}{2}\|u(t) - v(t)\|^2$: essa è una funzione localmente assolutamente continua in $(0, +\infty)$, in quanto composizione di una funzione assolutamente continua e di una funzione localmente Lipschitziana. Se t_0 è poi un punto di derivabilità per u e v e in cui queste ultime soddisfano l'inclusione differenziale, avrò che anche f è derivabile in questo punto e vale:

$$f'(t_0) = \langle u'(t_0) - v'(t_0), u(t_0) - v(t_0) \rangle.$$

Ora, poiché $u'(t_0) \in -\partial F(u(t_0))$ e $v'(t_0) \in -\partial F(v(t_0))$, per definizione di sottodifferenziale ho:

$$\begin{aligned} F(v(t_0)) &\geq F(u(t_0)) - \langle u'(t_0), v(t_0) - u(t_0) \rangle \\ F(u(t_0)) &\geq F(v(t_0)) - \langle v'(t_0), u(t_0) - v(t_0) \rangle. \end{aligned}$$

Sommando queste due disuguaglianze ottengo

$$0 \geq \langle u'(t_0) - v'(t_0), u(t_0) - v(t_0) \rangle = f'(t_0)$$

Quindi, poiché f è assolutamente continua, ne risulta che è decrescente in $(0, +\infty)$ e, per continuità, anche in $[0, +\infty)$.

Abbiamo dunque mostrato che $\|u(t) - v(t)\| \leq \|u_0 - v_0\|$ e dunque il flusso gradiente è una contrazione, in particolare se ho una soluzione, essa è unica. \square

Prima di passare alla costruzione della soluzione discreta richiamiamo alcune nozioni di analisi convessa che ci serviranno in seguito.

2.0.2 Richiami di analisi convessa

In questa sezione $F : H \rightarrow \mathbb{R} \cup \{+\infty\}$ denoterà sempre una funzione convessa semicontinua inferiormente. Abbiamo già richiamato la definizione di sottodifferenziale e del suo elemento di norma minima $\partial^\circ F$; per brevità di notazione chiameremo

$$|\partial F|(x) := \|\partial^\circ F(x)\| \quad \forall x \in D(\partial F)$$

e poi estendiamo $|\partial F|(x) = +\infty$ se x non è in $D(\partial F)$.

Teorema 2.5. *Sia F una funzione s.c.i. convessa. Consideriamo poi il grafico del sottodifferenziale di F , definito come:*

$$\Gamma(\partial F) = \{(x, p) \in H \times H \text{ tali che } p \in \partial F(x)\}.$$

Allora $\Gamma(\partial F)$ è fortemente-debolmente chiuso, cioè

$$p_n \in \partial F(x_n), \quad x_n \rightarrow x, \quad p_n \rightharpoonup p \quad \implies \quad p \in \partial F(x)$$

Dimostrazione. Prendiamo le disuguaglianze di sottodifferenziale per x_n :

$$F(y) \geq F(x_n) + \langle p_n, y - x_n \rangle \quad \forall y \in H.$$

Ora, considerando $n \rightarrow \infty$ e utilizzando la semicontinuità inferiore di F , otteniamo:

$$\begin{aligned} F(y) &\geq \liminf_{n \rightarrow \infty} F(x_n) + \lim_{n \rightarrow \infty} \langle p_n, y - x_n \rangle \\ &\geq F(x) + \langle p, y - x \rangle \end{aligned} \quad \forall y \in H$$

e dunque la tesi, per definizione di sottodifferenziale. \square

Ricordiamo qui alcuni fatti che discendono da questo teorema che ci saranno utili nel seguito:

Corollario 2.6. *Data F una funzione convessa s.c.i., allora valgono le seguenti affermazioni:*

- 1) *la funzione $|\partial F| : H \rightarrow \mathbb{R} \cup \{+\infty\}$ è semicontinua inferiormente;*
- 2a) *se $x_n \rightarrow x$ e $|\partial F|(x_n) \leq C$ allora $x \in D(\partial F)$;*
- 2b) *se $x_n \rightarrow x$ e $|\partial F|(x_n) \rightarrow |\partial F|(x)$, allora $\partial^\circ F(x_n) \rightarrow \partial^\circ F(x)$ fortemente.*

Dimostrazione. Prendiamo una successione $x_n \in D(\partial F)$ che tende a x , e denotiamo con $p_n = \partial^\circ F(x_n)$. Se $|\partial F|(x_n) \rightarrow \infty$ non c'è nulla da dimostrare, altrimenti vuol dire che, a meno di sottosuccessioni, $\|p_n\| = |\partial F|(x_n) \leq C$. Ma allora, di nuovo a meno di sottosuccessioni per Banach-Alaoglu, $p_n \rightharpoonup p$, e, per il teorema 2.5, $p \in \partial F(x)$ (abbiamo quindi dimostrato il punto 2a).

Per la semicontinuità inferiore della norma nella topologia debole inoltre si ha $|\partial F|(x) = \|p\| \leq C$, e quindi $|\partial F|$ risulta effettivamente semicontinua inferiormente.

Se valesse anche la condizione di continuità presente al punto 2b, si avrebbe che $\|p\| = \|\partial^\circ F(x)\|$, da cui per l'unicità dell'elemento di minima norma p sarebbe proprio $\partial^\circ F(x)$. Inoltre

$$p_n \rightharpoonup p \quad \text{e} \quad \|p_n\| \rightarrow \|p\|$$

e quindi, per risultati classici negli spazi di Hilbert, $p_n \rightarrow p$ fortemente. \square

Lemma 2.7. (Linearità del sottodifferenziale) *Data F funzione convessa e data G una funzione convessa Gateaux-derivabile in $x \in H$, si ha che*

$$\partial(F + G)(x) = G'(x) + \partial F(x)$$

Dimostrazione. Questo segue dalla definizione alternativa di sottodifferenziale, cioè $p \in \partial F(x)$ se e solo se

$$\liminf_{t \downarrow 0} \frac{F(x + tv) - F(x) - t\langle v, p \rangle}{t} \geq 0 \quad \forall v \in H.$$

Applicando questa definizione alla funzione $F + G$ nel punto x si ha $p + G'(x) \in \partial(F + G)(x)$ se e solo se

$$\begin{aligned} & \liminf_{t \downarrow 0} \frac{F(x + tv) + G(x + tv) - F(x) - G(x) - \langle y - x, p + G'(x) \rangle}{\|y - x\|} \geq 0 \\ & \liminf_{t \downarrow 0} \left\{ \frac{F(x + tv) - F(x) - t\langle v, p \rangle}{t} + \frac{G(x + tv) - G(x) - t\langle v, G'(x) \rangle}{t} \right\} \geq 0 \end{aligned}$$

ora il secondo addendo nel membro di sinistra tende a 0 per la derivabilità di G e dunque $p + G'(x) \in \partial(F + G)(x)$ se e solo se

$$\liminf_{t \downarrow 0} \frac{F(x + tv) - F(x) - t\langle v, p \rangle}{t} \geq 0,$$

ciò per definizione $p \in \partial F(x)$, come volevasi dimostrare. \square

Lemma 2.8. *Dato un punto $x \in D(F)$ esiste una successione $\{x_n\} \subseteq D(\partial F)$ tale che:*

- (i) $x_n \rightarrow x$;
- (ii) $\|\partial^\circ F(x_n)\| \cdot \|x_n - x\| \rightarrow 0$;
- (iii) $F(x_n) \rightarrow F(x)$.

Dimostrazione. Innanzitutto ricordiamo che una funzione convessa ha crescita dal basso almeno lineare, cioè

$$F(y) \geq c + \langle v, y - x \rangle \quad \forall y \in H$$

per un qualche numero reale c e un vettore $v \in H$: per dimostrarlo basta applicare Hahn-Banach geometrico alla chiusura del sopragrafico di F e un punto esterno ad esso. Supponiamo ora $x \in D(F)$; usando le idee del capitolo successivo definiamo:

$$x_n = \arg \min_{y \in D(F)} \{F(y) + n\|y - x\|^2\}.$$

Esiste ed è unico x_n in cui viene realizzato il minimo e $\partial F(x_n) \neq \emptyset$ poiché vi appartiene il vettore $2n(x - x_n)$ (per maggiori dettagli si può consultare la sezione riguardante la costruzione della soluzione discreta, dove questo procedimento è spiegato in dettaglio). Ci basta ora dimostrare che $x_n \rightarrow x$ e che $F(x_n) \rightarrow F(x)$: usiamo che

$$F(x_n) + n\|x_n - x\|^2 \leq F(x).$$

Ignorando il secondo addendo che è positivo, otteniamo $F(x_n) \leq F(x)$ e dunque $\limsup F(x_n) \leq F(x)$. Usando invece la stima dal basso di F :

$$F(x) \geq c + \langle v, x_n \rangle + n\|x_n - x\|^2 \geq c - \|v\|\|x_n - x\| + n\|x_n - x\|^2$$

da cui, risolvendo la disequazione per $\|x_n - x\|$, ottengo

$$\|x_n - x\| \leq \frac{\|v\|}{2n} + \sqrt{\frac{\|v\|^2}{4n^2} + \frac{F(x) - c}{n}} \leq \frac{1}{\sqrt{n}}A(x)$$

dove A è una costante che dipende solo da x e non da n . Dunque è dimostrato il primo punto: $x_n \rightarrow x$. Inoltre, usando la semicontinuità inferiore e la disuguaglianza trovata prima otteniamo:

$$\liminf_{n \rightarrow \infty} F(x_n) \geq F(x) \geq \limsup_{n \rightarrow \infty} F(x_n)$$

e quindi il limsup e il liminf coincidono, cioè $F(x_n) \rightarrow F(x)$.

Per il secondo punto dobbiamo raffinare la stima dal basso della funzione F : per trovare una migliore approssimazione, prendiamo in considerazione il punto $(x, F(x) - \varepsilon)$ e un iperpiano che lo separa dal grafico di F . Con questa scelta avremo

$$F(y) \geq c_\varepsilon + \langle v_\varepsilon, y - x \rangle \quad \forall y \in H$$

dove $c_\varepsilon \geq F(x) - \varepsilon$. Usando questa stima nella disuguaglianza per $\|x_n - x\|$:

$$\|x_n - x\| \leq \frac{\|v_\varepsilon\|}{2n} + \sqrt{\frac{\|v_\varepsilon\|^2}{4n^2} + \frac{\varepsilon}{n}}$$

e quest'ultima quantità, per $n \geq N_\varepsilon$, è minore o uguale a $\frac{2\sqrt{\varepsilon}}{\sqrt{n}}$. Dunque otteniamo che

$$\|x_n - x\| = o\left(\frac{1}{\sqrt{n}}\right)$$

e, sapendo che $2n(x_n - x) \in \partial F(x_n)$,

$$\|\partial^\circ F(x_n)\| \cdot \|x_n - x\| \leq \|2n(x_n - x)\| \cdot \|x_n - x\| = 2n\|x_n - x\|^2 = o(1),$$

come volevasi dimostrare. □

2.1 La soluzione discreta e il rapporto tra EVI e GF

Discretizziamo il tempo con passi di lunghezza $\tau > 0$ che poi manderemo a 0 per recuperare la soluzione del problema continuo. Dato $y_0 \in D(F)$ costruiamo la successione

$$y_0, y_1, y_2, \dots, y_n$$

$$y_{i+1} = \arg \min_{x \in D(F)} \left\{ \frac{1}{2\tau} \|y_i - x\|^2 + F(x) \right\}.$$

Questa definizione è ben posta poiché la funzione di cui dobbiamo trovare il minimo è una funzione strettamente convessa e coerciva, e dunque ha un minimo in un unico punto, che sarà y_{i+1} . Quindi possiamo scrivere:

$$\begin{aligned} \frac{1}{2\tau} \|y_i - y_{i+1}\|^2 + F(y_{i+1}) &\leq \frac{1}{2\tau} \|y_i - x\|^2 + F(x) \\ \langle y_i, x - y_{i+1} \rangle + \frac{1}{2} \|y_{i+1}\|^2 + \tau F(y_{i+1}) &\leq \|x\|^2 + \tau F(x) \\ y_i &\in (Id + \tau \partial F)(y_{i+1}) \end{aligned}$$

dove abbiamo usato la linearità del sottodifferenziale. In particolare tutti i passaggi sono invertibili e dunque ciò dà l'invertibilità di $(Id + \tau \partial F)$, che ci permette di scrivere $y_{i+1} = (Id + \tau \partial F)^{-1}(y_i)$.

Quest'ultima scrittura è anche quella dello schema di Eulero implicito: in un contesto discreto possiamo scrivere la derivata come rapporto incrementale e quindi l'inclusione differenziale diventerebbe

$$\frac{y_i - y_{i+1}}{\tau} \in \partial F(y_i).$$

Questa è quella che si chiama costruzione esplicita, in quanto come dato per ∂F a destra prendiamo y_i , già conosciuto. È più conveniente, in termini di stabilità, prendere come dato y_{i+1} e ottenere il cosiddetto schema implicito di Eulero:

$$\begin{aligned} \frac{y_i - y_{i+1}}{\tau} &\in \partial F(y_{i+1}) \\ y_i &\in (Id + \tau \partial F)(y_{i+1}). \end{aligned}$$

Ritroviamo dunque la caratterizzazione di prima e possiamo pensare y_i come la soluzione del flusso gradiente al tempo $i\tau$.

Inoltre possiamo così ritrovare la proprietà di contrazione vista nel caso continuo. Date due successioni costruite secondo questo metodo a partire da x_0 e y_0 , per le proprietà del sottodifferenziale già usate nel dimostrare l'unicità:

$$\begin{aligned} \langle (y_i - y_{i+1}) - (x_i - x_{i+1}), y_{i+1} - x_{i+1} \rangle &\geq 0 \\ \langle y_i - x_i, y_{i+1} - x_{i+1} \rangle &\geq \|y_{i+1} - x_{i+1}\|^2 \end{aligned}$$

ma per Cauchy-Schwarz si ha

$$\|y_i - x_i\| \cdot \|x_{i+1} - y_{i+1}\| \geq \langle y_i - x_i, y_{i+1} - x_{i+1} \rangle \geq \|y_{i+1} - x_{i+1}\|^2$$

e dunque, se $y_{i+1} \neq x_{i+1}$ allora, semplificando $\|y_{i+1} - x_{i+1}\|$ otteniamo

$$\|y_{i+1} - x_{i+1}\| \leq \|y_i - x_i\|$$

d'altronde, se $y_{i+1} = x_{i+1}$, necessariamente $\|y_i - x_i\| \geq 0 = \|y_{i+1} - x_{i+1}\|$. Dunque si ha la contrazione lungo le soluzioni.

Riconduciamo ora il problema di trovare una soluzione al flusso gradiente (GF) al trovare una soluzione per la disuguaglianza variazionale di evoluzione (EVI). Questa formulazione è più comoda in quanto permette più facilmente di controllare le proprietà di stabilità di questo problema.

Definizione 2.9. *Data una curva $u \in \mathcal{AC}_{loc}((0, +\infty); H)$ e una funzione convessa F , si dice che u è una soluzione variazionale di evoluzione per F con dato iniziale $u_0 = \lim_{t \rightarrow 0} u(t)$ se $\forall y \in H$ vale quasi ovunque (in t) la disuguaglianza*

$$\frac{1}{2} \frac{d}{dt} \|y - u(t)\|^2 + F(u(t)) \leq F(y).$$

Se u è un flusso gradiente ovviamente soddisfa (EVI). Infatti sappiamo che per quasi ogni $t > 0$, la velocità della curva $u'(t)$ appartiene a $-\partial F(u(t))$ e quindi, per la disuguaglianza di sottodifferenziale si ha:

$$F(y) \geq F(u(t)) - \langle y - u(t), u'(t) \rangle \quad \forall y \in H.$$

Ma $2\langle y - u(t), u'(t) \rangle$ è proprio la derivata temporale di $\|y - u(t)\|^2$ e dunque la tesi, cioè se la curva u è un flusso gradiente allora soddisfa anche (EVI).

Ora facciamo vedere che vale anche il contrario, cioè che se u è soluzione della (EVI) allora è anche soluzione di (GF): innanzitutto consideriamo $H_1 = \overline{\text{span}(u((0, +\infty)))}$ che risulta separabile in quanto immagine continua di uno spazio separabile. Consideriamo ora

$$\begin{aligned} S : H &\rightarrow H_1 \times \mathbb{R} \\ x &\mapsto (\pi(x), F(x)) \end{aligned}$$

e $\mathcal{F} = \text{Im}(S)$ che, essendo un sottoinsieme di uno spazio separabile, sarà separabile anch'esso. Quindi esiste $\{x_n\} \subseteq H$ tale che $\{S(x_n)\}$ è denso in \mathcal{F} .

Definiamo poi $\mathcal{N}(y)$ come l'insieme dei tempi in cui non vale la disuguaglianza (EVI) fissato y .

Considero ora $\mathcal{N} = \bigcup \mathcal{N}(x_k)$; per definizione di (EVI), $\mathcal{N}(y)$ ha misura di Lebesgue nulla per ogni $y \in H$ e dunque anche \mathcal{N} , unione numerabile di insiemi di misura nulla, è Lebesgue-trascurabile.

Facciamo vedere ora che per ogni $t \in \mathcal{N}$, che sia anche di differenziabilità forte per u (quindi per quasi ogni $t > 0$), vale la disuguaglianza del sottodifferenziale con u' , per ogni $x \in H$ e dunque u è effettivamente un flusso gradiente.

Preso un qualunque $x \in H$, esiste, per densità, una sottosuccessione $x_{n(k)}$ tale che $\pi(x_{n(k)}) \rightarrow \pi(x)$ e $F(x_{n(k)}) \rightarrow F(x)$. Ora, per ogni $t \in \mathcal{N}$ si ha che

$$-\langle x_{n(k)} - u(t), u'(t) \rangle + F(u(t)) \leq F(x_{n(k)})$$

e passando al limite in k

$$-\langle x - u(t), u'(t) \rangle + F(u(t)) \leq F(x)$$

dove ho usato il fatto che $\langle x_{n(k)}, u'(t) \rangle = \langle \pi(x_{n(k)}), u'(t) \rangle$.

Mostriamo ora l'equivalente discreto della (EVI):

$$\frac{1}{2} \frac{\|y_{i+1} - x\|^2 - \|y_i - x\|^2}{\tau} + F(y_{i+1}) \leq F(x). \quad (\text{EVI})_\tau$$

Consideriamo ora la successione costruita prima: vorremmo mostrare che essa soddisfa anche $(\text{EVI})_\tau$. Sappiamo già che $y_{i+1} - y_i \in -\tau \partial F(y_{i+1})$ e dunque

$$\frac{\langle y_{i+1} - y_i, y_{i+1} - x \rangle}{\tau} + F(y_{i+1}) \leq F(x)$$

Quindi mi basterebbe che la differenza tra i due membri di destra delle ultime due disequazioni sia negativa. Ma

$$\|y_{i+1} - x\|^2 - \|y_i - x\|^2 - 2\langle y_{i+1} - y_i, y_{i+1} - x \rangle = -\|y_i - y_{i+1}\|^2 \leq 0$$

e dunque la successione $\{y_i\}$ soddisfa $(\text{EVI})_\tau$.

2.2 Proprietà della soluzione discreta

Discutiamo qui le proprietà di y_k : gli ingredienti essenziali saranno la $(\text{EVI})_\tau$

$$\frac{1}{2} \frac{\|y_{i+1} - x\|^2 - \|y_i - x\|^2}{\tau} + F(y_{i+1}) \leq F(x)$$

e la disuguaglianza

$$\|y_{i+1} - y_i\|^2 + 2\tau F(y_{i+1}) = \min\{\|x - y_i\|^2 + 2\tau F(x)\} \leq 2\tau F(y_i)$$

che chiamiamo stima dell'energia.

Definiamo poi le funzioni S_t^τ e \widehat{S}_t^τ : entrambe, valutate in y_0 ad un tempo $i\tau$, valgono y_i , mentre negli altri tempi \widehat{S}_t^τ è definita interpolando linearmente tra due multipli successivi di τ mentre S_t^τ è definita a gradini, cioè $S_t^\tau y_0 = S_{i\tau}^\tau y_0$ se $i\tau \leq t < (i+1)\tau$.

Con la stima dell'energia si trova che $\widehat{S}_t^\tau y \in \mathcal{AC}_2((0, T); H)$. Infatti $\frac{d}{dt} \widehat{S}_t^\tau y = \frac{1}{\tau}(y_{i+1} - y_i)$ se $i\tau < t < (i+1)\tau$, da cui

$$\int_0^{n\tau} \left\| \frac{d}{dt} \widehat{S}_s^\tau y \right\|^2 ds = \sum_{i=0}^{n-1} \tau \left\| \frac{y_{i+1} - y_i}{\tau} \right\|^2 \leq 2(F(y_0) - F(y_n))$$

Per stime uniformi è utile avere informazioni dal basso di $F(y_i)$ solo in funzione di y_0 . Queste sono automatiche se ad esempio esiste un minimo che, senza perdita di generalità possiamo assumere sia 0. Altrimenti usiamo le stime seguenti quando $\partial F(y_0) \neq \emptyset$:

Lemma 2.10. *Sia $p_0 \in \partial F(y_0)$ e sia $-\frac{1}{\tau}(y_{i+1} - y_i) = p_{i+1} \in \partial F(y_{i+1})$ per $i \geq 0$. Allora valgono:*

- (i) $\|p_i\| \geq \|p_{i+1}\|$, per ogni $i \geq 0$;
- (ii) $0 \geq F(y_i) - F(y_j) \geq -(i-j)\tau\|p_0\|^2$, per ogni $i > j \geq 0$;
- (iii) $\|y_i - y_j\| \leq \sqrt{(i-j)\tau} \sqrt{F(y_j) - F(y_i)} \leq |i-j|\tau\|p_0\|$.

Dimostrazione. Per il punto (i) mi basta dimostrarlo per $i = 0$ e poi riapplicare lo stesso ragionamento per ogni i . Per quanto detto sui sottodifferenziali sappiamo già che

$$\langle p_1 - p_0, y_1 - y_0 \rangle \geq 0.$$

Ora sostituiamo $y_1 - y_0 = -\tau p_1$ e ottengo

$$\begin{aligned} \langle p_1, p_0 \rangle &\geq \|p_1\|^2 \\ \|p_1\| \cdot \|p_0\| &\geq \|p_1\|^2 \\ \|p_0\| &\geq \|p_1\|. \end{aligned}$$

Per il punto (ii), sfruttando la definizione di sottodifferenziale

$$F(y_i) - F(y_j) \geq \langle y_j - y_i, p_i \rangle \geq -\|y_j - y_i\| \cdot \|p_i\|.$$

Usiamo ora la disuguaglianza triangolare e poi il punto (i):

$$\|y_j - y_i\| \leq \sum_{k=i}^{j-1} \|y_{k+1} - y_k\| = \tau \sum_{k=i}^{j-1} \|p_{k+1}\| \leq \tau(i-j)\|p_{i+1}\|.$$

Rimettendo nella prima disuguaglianza, e ricordando di nuovo che $\|p_i\| \leq \|p_0\|$

$$F(y_i) - F(y_j) \geq -\|y_j - y_i\| \cdot \|p_i\| \geq -\tau(i-j)\|p_0\|^2$$

(iii) Dalla stima dell'energia applicata più volte otteniamo:

$$\sum_{k=i}^{j-1} \|y_{k+1} - y_k\|^2 \leq \tau(F(y_i) - F(y_j))$$

e per Cauchy-Schwarz e disuguaglianza triangolare:

$$\|y_i - y_j\| \leq \sum_{k=i}^{j-1} \|y_{k+1} - y_k\| \leq \sqrt{i-j} \sqrt{\sum_{k=i}^{j-1} \|y_{k+1} - y_k\|^2}$$

da cui la tesi mettendo insieme queste due disuguaglianze e anche quella dimostrata al punto (ii). \square

La disuguaglianza del punto (ii) si può leggere in termini di S_t^τ come:

$$F(S_t^\tau y_0) - F(S_s^\tau y_0) \geq -(s-t)\|p_0\|^2$$

quando t, s sono multipli interi di τ .

Teorema 2.11. *Data F convessa s.c.i. e $\bar{y} \in D(F)$ tale che $p \in \partial F(\bar{y}) \neq \emptyset$, esiste una (unica) soluzione $S_t \bar{y}$ alla EVI che parte da \bar{y} . In più essa soddisfa le proprietà:*

$$(i) \quad \|S_t \bar{y} - S_s \bar{y}\| \leq \|p\| \cdot |t - s|$$

$$(ii) \quad \|S_t^\tau \bar{y} - S_t \bar{y}\| \leq C \|p\| \sqrt{\tau t}$$

$$(iii) \quad F(S_t \bar{y}) \leq \inf_{x \in D(F)} \left\{ \frac{1}{2t} \|\bar{y} - x\|^2 + F(x) \right\}$$

$$(iv) \quad \int_0^t \left\| \frac{d}{ds} S_s^\tau \bar{y} \right\|^2 ds \leq 2(F(\bar{y}) - F(S_t \bar{y}))$$

Dimostrazione. Consideriamo le soluzioni discrete con passi $\tau_m = 2^{-m} \tau$ e facciamo vedere che esse tendono uniformemente sui compatti $(0, T)$ ad una curva in \mathcal{AC}_2 .

Procediamo alla stima principale che permette di stimare la differenza tra S_t^τ e $S_t^{\frac{\tau}{2}}$. Consideriamo innanzitutto due passi di lunghezza $\frac{\tau}{2}$ per poterlo paragonare a un passo di lunghezza τ :

$$\begin{aligned} F(\theta) - F(S_\tau^{\frac{\tau}{2}} x) &\geq \frac{1}{2} \left(F(\theta) - F(S_\tau^{\frac{\tau}{2}} x) \right) + \frac{1}{2} \left(F(\theta) - F(S_\tau^{\frac{\tau}{2}} x) \right) \\ &\geq \frac{1}{2} \frac{\|S_\tau^{\frac{\tau}{2}} x - \theta\|^2 - \|S_\tau^{\frac{\tau}{2}} x - \theta\|^2 + \|S_\tau^{\frac{\tau}{2}} x - \theta\|^2 - \|x - \theta\|^2}{\tau} \\ &= \frac{\|S_\tau^{\frac{\tau}{2}} x - \theta\|^2 - \|x - \theta\|^2}{2\tau}. \end{aligned}$$

Quindi sono soddisfatte le due seguenti disuguaglianze:

$$F(\theta) - F(S_\tau^{\frac{\tau}{2}} x) \geq \frac{\|S_\tau^{\frac{\tau}{2}} x - \theta\|^2 - \|x - \theta\|^2}{2\tau}$$

$$F(\eta) - F(S_\tau^\tau y) \geq \frac{\|S_\tau^\tau y - \eta\|^2 - \|y - \eta\|^2}{2\tau}$$

e, sostituendo $\eta = x$ e $\theta = S_\tau^\tau y$ e sommando, ottengo

$$F(x) - F(S_\tau^{\frac{\tau}{2}} x) \geq \frac{\|S_\tau^{\frac{\tau}{2}} x - S_\tau^\tau y\|^2 - \|x - y\|^2}{2\tau}.$$

Iterando questa stima, dove al posto di x e y si sostituiscono rispettivamente $S_{n\tau}^{\frac{\tau}{2}} x$ e $S_{n\tau}^\tau x$, ottengo sommando:

$$2\tau \left(F(x) - F(S_t^{\frac{\tau}{2}} x) \right) \geq \|S_t^{\frac{\tau}{2}} x - S_t^\tau x\|^2.$$

Ora, partendo con $x = \bar{y}$ ottengo la stima

$$\|S_t^{\frac{\tau}{2}} \bar{y} - S_t^\tau \bar{y}\| \leq \sqrt{2\tau} \cdot \sqrt{F(\bar{y}) - F(S_t^{\frac{\tau}{2}} \bar{y})} \leq \|p\| \sqrt{2t\tau}$$

che, iterata con passi τ_m ci dà, se $n < m$:

$$\|S_t^{\tau_m} \bar{y} - S_t^{\tau_n} \bar{y}\| \leq \|p\| \sqrt{2t} \sum_{i=n}^{m-1} \sqrt{\tau_i} \leq \frac{2\|p\| \sqrt{t\tau}}{(\sqrt{2})^n}$$

e dunque, se $2^k t$ è multiplo intero di τ per qualche k , la successione $S_t^{\tau_n} \bar{y}$ converge puntualmente ad un vettore che chiameremo $S_t \bar{y}$. Prendendo $n = 0$ e mandando m ad infinito ottengo la (ii). Prendendo la (iii) del lemma 2.10 e prendendo il limite per $\tau \rightarrow 0$ ottengo la (i). Inoltre è facile vedere interpolando che $\widehat{S_t^{\tau_n} \bar{y}}$ converge per $n \rightarrow \infty$ proprio a $S_t \bar{y}$, uniformemente in t , nei punti in cui è definita. Ma allora converge localmente uniformemente ad una funzione che risulterà continua che estenderà $S_t \bar{y}$; in particolare avrò $S_0 \bar{y} = \bar{y}$. Ma anche le funzioni approssimate a gradini convergono a $S_t \bar{y}$ uniformemente e dunque ci permettono di stimare la derivata distribuzionale: data una funzione $\varphi \in C_c^\infty(0, +\infty)$ vale

$$\begin{aligned} - \int_0^{+\infty} \|S_t \bar{y} - \theta\|^2 \varphi'(t) dt &= \lim_{m \rightarrow +\infty} - \int_0^{+\infty} \|S_t^{\tau_m} \bar{y} - \theta\|^2 \varphi'(t) dt \\ &= \lim_{m \rightarrow +\infty} \sum_{i=1}^{+\infty} (\|S_{i\tau_m}^{\tau_m} \bar{y} - \theta\|^2 - \|S_{(i-1)\tau_m}^{\tau_m} \bar{y} - \theta\|^2) \varphi(i\tau) \\ &\leq \liminf_{m \rightarrow +\infty} \sum_{i=1}^{+\infty} 2\tau (F(\theta) - F(\widehat{S_{i\tau_m}^{\tau_m} \bar{y}})) \varphi(i\tau) \\ &= 2 \int_0^{+\infty} (F(\theta) - F(S_t \bar{y})) \varphi(t) dt. \end{aligned}$$

Nell'ultima uguaglianza si usa il fatto che $\widehat{S_t^{\tau_m} \bar{y}} \rightarrow S_t \bar{y}$ uniformemente sui compatti (come funzioni continue) e $\sum_{i=1}^{+\infty} \delta_{i\tau} \varphi(i\tau) \tau \rightarrow \varphi \mathcal{L}^1$ come misure a supporto compatto (che è sempre *spt* φ). Dunque abbiamo dedotto che la derivata distribuzionale di $\|S_t \bar{y} - \theta\|^2$ è minore o uguale a $2(F(\theta) - F(S_t \bar{y}))$.

Ma dal teorema 2.3 si ha che $S_t \bar{y} \in \mathcal{AC}_{loc}^2$ grazie alle stime uniformi sui compatti che riusciamo ad avere sull'integrale del quadrato della derivata $\widehat{S_t^{\tau_m} \bar{y}}$ mediante il punto (ii) del lemma 2.10. Queste stime della derivata, portate al limite continuo danno proprio la (iv) e inoltre ho che $S_t \bar{y}$ è derivabile quasi ovunque e dunque in particolare sarà vero che, fissato θ , quasi ovunque valga

$$\frac{1}{2} \frac{d}{dt} \|S_t \bar{y} - \theta\|^2 \leq F(\theta) - F(S_t \bar{y}).$$

Dunque $S_t \bar{y}$ è effettivamente una soluzione EVI. □

2.3 Regolarità ed estensione di S_t a $\overline{D(F)}$

In questa sezione ci occupiamo della regolarità del flusso gradiente, in particolare dimostreremo che esiste la derivata temporale destra per tutti i tempi positivi e sarà uguale a $\partial^\circ F$ in tutti i punti (principio di selezione minima). Inoltre c'è anche derivabilità a sinistra ovunque

tranne che in un insieme al più numerabile di punti.

Tutto questo sarà vero anche partendo da un qualsiasi punto in $\overline{D(F)}$ e la soluzione (unica), si troverà in $D(\partial F)$ per tutti i tempi positivi.

Iniziamo col mostrare la regolarità nel caso in cui il dato iniziale sia in $D(\partial F)$.

Teorema 2.12. *Sia y un punto di $D(\partial F)$ e sia $S_t y$ il suo flusso gradiente. Allora*

- (i) $S_t y \in D(\partial F)$, per ogni $t > 0$;
- (ii.a) la funzione $g(t) = |\partial F|(S_t y)$ è decrescente e continua a destra;
- (ii.b) la funzione $G(t) = \partial^\circ F(S_t y)$ è continua a destra;
- (iii) $S_t y$ è derivabile a destra per ogni $t \geq 0$ e la derivata coincide con quella sinistra nei punti di continuità di $g(t)$. Vale inoltre:

$$\frac{d^+}{dt} S_t y = \partial^\circ F(S_t y) \quad \text{per ogni } t \geq 0$$

Dimostrazione. Nella dimostrazione del teorema 2.11 abbiamo visto che $\widehat{S}_t^\tau y \rightarrow S_t y$. Sappiamo che $\widehat{S}_t^\tau y$ è sicuramente in $D(\partial F)$ e in particolare, supponendo $p_0 = \partial^\circ F(y)$ nelle ipotesi del teorema 2.11, si ha

$$|\partial F|(\widehat{S}_t^\tau y) \leq \|p_0\| = |\partial F|(y).$$

Dunque per il corollario 2.6 si ha $S_t y \in D(\partial F)$ e

$$|\partial F|(S_t y) \leq |\partial F|(y).$$

Consideriamo ora il flusso gradiente che parte da $S_s y$, che sappiamo esistere ed essere unico in quanto il suo sottodifferenziale è non vuoto. Per l'unicità questo flusso $S_t(S_s y)$ coincide puntualmente con $S_{s+t} y$, infatti anch'esso soddisfa (GF). Dunque in particolare, usando l'ultima disuguaglianza con flusso gradiente di origine $S_s y$, otteniamo che

$$g(t+s) = |\partial F|(S_{t+s} y) = |\partial F|(S_t(S_s y)) \leq |\partial F|(S_s y) = g(s) \quad \forall s, t \geq 0$$

e dunque g è una funzione decrescente. Per dimostrare la continuità a destra consideriamo $t_k \downarrow t$; quindi ho $g(t_k) \leq g(t)$ e, di nuovo per la semicontinuità di $|\partial F|$, avrò che

$$g(t) = |\partial F|(S_t y) \leq \liminf_{t_k \rightarrow t} |\partial F|(S_{t_k} y) = \limsup_{t_k \rightarrow t} g(t_k) \leq g(t)$$

da cui tutte le disuguaglianze qui sono uguaglianze e quindi $g(t_k) \uparrow g(t)$ (continuità a destra). Poiché le norme convergono, per il corollario 2.6 si ha la convergenza forte $\partial^\circ F(S_{t_k} y) \rightarrow \partial^\circ F(S_t y)$; allo stesso modo si può vedere che c'è convergenza forte anche da sinistra se $g(t)$ è continua a sinistra in quel punto.

Per il punto (iii) ci avvaliamo del fatto, dimostrato in appendice, che, per una curva assolutamente continua, vale il "teorema fondamentale del calcolo integrale".

Applicando il primo punto del teorema 2.11 (con $\bar{y} = S_s y$ e $p = \partial^\circ F(S_s y)$) e sfruttando la proprietà di semigruppato di S_t dimostrata prima, otteniamo che $S_t y$ è $|\partial F|(S_s y)$ -lipschitziana per $t \geq s$ e in particolare, se esiste la derivata del flusso in s , essa dovrà essere in modulo necessariamente minore o uguale di $\|\partial^\circ F(S_s y)\|$; unendo questo all'enunciato (GF)

$$\frac{d}{dt} S_t y \in -\partial F(S_t y) \quad \text{per quasi ogni } t \geq 0$$

ottengo la caratterizzazione più forte, cioè in tutti i punti in cui il flusso sia derivabile, la derivata è forzata ad essere l'elemento di norma minima del sottodifferenziale:

$$\frac{d}{dt} S_t y = -\partial^\circ F(S_t y) \quad \text{per quasi ogni } t \geq 0.$$

Applicando ora il “teorema fondamentale del calcolo integrale” per tempo da t ad s si ha

$$\frac{S_s y - S_t y}{s - t} = -\frac{\int_s^t \partial^\circ F(S_r y) dr}{s - t}.$$

Facendo ora tendere $s \downarrow t$, per la continuità a destra di $\partial^\circ F(S_r y)$, si ha che il limite esiste e dunque vale

$$\frac{d^+}{dt} S_t y = -\partial^\circ F(S_t y) \quad \text{per ogni } t \geq 0.$$

Inoltre nei punti di continuità di $g(t)$ c'è anche continuità a sinistra di $\partial^\circ F(S_r y)$ e quindi

$$\frac{d}{dt} S_t y = -\partial^\circ F(S_t y) \quad \text{per ogni } t \in [0, +\infty)/\Gamma$$

e qui $\Gamma = \{g(t^+) < g(t^-)\}$, l'insieme dei punti di discontinuità di g , è un insieme al più numerabile per la monotonia della funzione stessa. \square

Vogliamo ora estendere la definizione di S_t a tutto $D(F)$; per fare questo, prendiamo x_n una successione che soddisfi le proprietà del lemma 2.8 e vediamo come si comporta $S_t x_n$. Un'analisi preliminare ci porta a:

- $|\partial F|(S_t x_n) \leq |\partial F|(x_n)$, dal punto (ii) del teorema 2.12;
- $S_t x_n$ è di Cauchy, poichè S_t è una contrazione su $D(\partial F)$; chiamando $S_t x := \lim S_t x_n$ si ha in particolare che

$$\|S_t x_n - S_t x\| \leq \|x_n - x\|;$$

e quindi S_t , esteso a $D(F)$, rimane ancora una contrazione;

- anche le energie convergono: $F(S_t x_n) \rightarrow F(S_t x)$ e quindi in particolare $F(S_t x) \leq F(x)$ poichè

$$F(S_t x_n) \leq F(x_n) \leq F(x)$$

e dunque $S_t x \in D(F)$.

Per mostrare l'ultimo punto innanzitutto ricordiamo che per la semicontinuità inferiore vale $\liminf F(S_t x_n) \geq F(S_t x)$; inoltre per la disuguaglianza di sottodifferenziale vale:

$$F(S_t x) \geq \limsup_{n \rightarrow \infty} \{F(S_t x_n) - \langle S_t x - S_t, \partial^\circ F(S_t x_n) \rangle\}.$$

Usando ora Cauchy-Schwarz e le stima prima citate:

$$\begin{aligned} \langle S_t x - S_t x_n, \partial^\circ F(S_t x_n) \rangle &\leq \|S_t x - S_t x_n\| \cdot \|\partial^\circ F(S_t x_n)\| \\ &\leq \|x - x_n\| \cdot \|\partial^\circ F(x_n)\| \rightarrow 0 \end{aligned}$$

e quindi possiamo ignorare l'addendo corrispondente ottenendo

$$\limsup_{n \rightarrow \infty} F(S_t x_n) \leq F(S_t x) \leq \liminf_{n \rightarrow \infty} F(S_t x_n).$$

Possiamo ora enunciare il teorema generale:

Teorema 2.13. *Sia $x \in \overline{D(F)}$. Allora esiste (ed è unico) il flusso gradiente che parte da x . Chiamato $S_t x$ questo flusso si hanno le seguenti proprietà:*

- (i) $S_t \circ S_s = S_{t+s}$, quindi S è un semigruppato;
- (ii) $S_t x \in D(\partial F)$, per ogni tempo t positivo;
- (iii) $S_t x$ risulta derivabile a destra in senso forte per ogni $t > 0$ e vale

$$\frac{d^+}{dt} S_t x = \partial^\circ(S_t x);$$

- (iv) $S_t x$ risulta derivabile in senso forte per ogni $t \notin \Gamma$, dove Γ è l'insieme, al più numerabile, di salti di discontinuità della funzione decrescente $g(t) = |\partial F|(S_t x)$.

Dimostrazione. Affrontiamo prima il caso $x \in D(F)$: definendo $S_t x$ come limite di $S_t x_n$, si ha che la curva così definita appartiene a $AC_2^{loc}([0, +\infty); D(F))$. Infatti, preso un qualunque intervallo chiuso $[0, T]$, le energie in $AC_2([0, T])$ delle curve $S_t x_n$, usando la stima nel punto (iv) del teorema 2.11 e la convergenza puntuale dei valori della funzione, convergono a $F(x) - F(S_t x)$ e dunque sono limitate e allora, per il teorema 2.3 ho che anche $S_t x$ è una curva assolutamente continua e quindi derivabile quasi ovunque.

Inoltre, passando al limite la (EVI) debole per $S_t x_n$:

$$- \int_0^{+\infty} \|S_t x_n - y\| \varphi'(t) dt \leq 2 \int_0^{+\infty} (F(\theta) - F(S_t x_n)) \varphi(t) dt$$

si ottiene, per convergenza di $S_t x_n$ a $S_t x$ e di $F(S_t x_n)$ a $F(S_t x)$

$$- \int_0^{+\infty} \|S_t x - y\| \varphi'(t) dt \leq 2 \int_0^{+\infty} (F(\theta) - F(S_t x)) \varphi(t) dt$$

cioè la (EVI) debole per $S_t x$ che, unita alla derivabilità quasi ovunque del flusso permette di dimostrare che $S_t x$ risolve (EVI), e dunque (GF). Abbiamo dunque ottenuto l'esistenza del

(GF) per ogni $x \in D(F)$.

La proprietà (i), come è già stato fatto vedere, è conseguenza dell'unicità, dimostrata all'inizio del capitolo. Detto questo sappiamo che per ogni $t_0 > 0$ esiste un tempo $0 < s < t_0$ in cui è verificata l'inclusione differenziale del (GF), in quanto è verificata quasi ovunque. Quindi $S_s x \in D(\partial F)$ e, usando la proprietà di semigruppato e il teorema 2.12 otteniamo le proprietà (ii), (iii), (iv) per ogni tempo $t \geq t_0$; data l'arbitrarietà di t_0 , otteniamo le tesi per ogni $t > 0$. Consideriamo ora $x \in \overline{D(F)}$ e una successione $x_k \in D(\partial F)$ che converga a x . Come prima poiché S_t risulta essere una contrazione, si ha che $S_t x_k$ è una successione di Cauchy, che converge ad un punto che chiamiamo $S_t x$. Ora usando il punto (iii) del teorema 2.11 e la semicontinuità, otteniamo:

$$\begin{aligned} \min_{y \in D(F)} \left\{ \frac{\|x - y\|^2}{2s} + F(y) \right\} &= \frac{\|y_0 - x\|^2}{2s} + F(y_0) \\ &= \lim_{k \rightarrow +\infty} \frac{\|y_0 - x_k\|^2}{2s} + F(y_0) \\ &\geq \liminf_{k \rightarrow +\infty} \min_{y \in D(F)} \left\{ \frac{\|y - x_k\|^2}{2s} + F(y) \right\} \\ &\geq \liminf_{k \rightarrow +\infty} F(S_s x_k) \geq F(S_s x) \end{aligned}$$

il che dimostra $S_s x \in D(F)$ per ogni tempo $s > 0$, e questo ci permette ancora una volta di dimostrare la proprietà di semigruppato: infatti $S_s x_k \rightarrow S_s x$, allora per la contrattività $S_t(S_s x_k) \rightarrow S_t(S_s x)$.

Per la proprietà di semigruppato in $D(F)$ so già che $S_t(S_s x_k) = S_{t+s} x_k$, che, per definizione so che converge a $S_{t+s} x$ che quindi coinciderà con $S_t(S_s x)$. In particolare per $t \geq 0$ ho che $S_{t+s} x$ è un flusso gradiente, ma allora, per l'arbitrarietà di s lo è per ogni tempo positivo e in più eredita tutte le proprietà (ii), (iii), (iv) per la proprietà di semigruppato, come fatto prima. \square

2.3.1 Proprietà di stabilità

Finora abbiamo descritto come funzionano i flussi gradiente fissato un funzionale e il punto iniziale. La proprietà di contrazione del flusso gradiente ci dà la stabilità al variare del punto iniziale; analizziamo ora la situazione in cui variamo il funzionale, in particolare con una perturbazione lineare.

Data $F : H \rightarrow \mathbb{R} \cup \{+\infty\}$ una funzione convessa e s.c.i., consideriamo, per ogni $v \in H$, la funzione:

$$F_v(x) := \begin{cases} F(x) + \langle v, x \rangle & \text{se } x \in D(F) \\ +\infty & \text{altrimenti} \end{cases}$$

Come prima, chiamiamo S_t e U_t i flussi gradiente rispettivamente secondo le funzioni F e F_v . Con la stessa idea di prima, facciamo un procedimento di discretizzazione e vediamo come

sono legate le funzioni S_τ^τ e U_τ^τ e notiamo che:

$$\begin{aligned} U_\tau^\tau y &= \arg \min_{x \in D(F)} \left\{ \frac{1}{2\tau} \|y - x\|^2 + F(x) + \langle v, x \rangle \right\} \\ &= \arg \min_{x \in D(F)} \left\{ \frac{1}{2\tau} \|y - \tau v - x\|^2 + F(x) \right\} \\ &= S_\tau^\tau(y - \tau v) \end{aligned}$$

Dunque, usando la proprietà di contrazione di S_τ^τ otteniamo (per semplicità di notazione d'ora in poi $S_\tau^\tau = \mathcal{S}$ e $U_\tau^\tau = \mathcal{U}$)

$$\|\mathcal{S}x - \mathcal{U}x\| = \|\mathcal{S}x - \mathcal{S}(x - \tau v)\| \leq \|x - (x - \tau v)\| = \tau \|v\| \quad \forall x \in D(F)$$

Iterando il procedimento usando la disuguaglianza triangolare, otteniamo che:

$$\begin{aligned} \|\mathcal{S}^{n+1}y - \mathcal{U}^{n+1}y\| &\leq \|\mathcal{S}(\mathcal{S}^n y) - \mathcal{S}(\mathcal{U}^n y)\| + \|\mathcal{S}(\mathcal{U}^n y) - \mathcal{U}(\mathcal{U}^n y)\| \\ &\leq \|\mathcal{S}^n y - \mathcal{U}^n y\| + \tau \|v\| \end{aligned}$$

e quindi, per induzione, otteniamo $\|\mathcal{S}_t^\tau y - \mathcal{U}_t^\tau y\| \leq t\|v\|$, per ogni tempo t multiplo intero di τ . Prendendo il limite per $\tau \rightarrow 0$ e un'ulteriore disuguaglianza triangolare, otteniamo:

$$\|\mathcal{S}_t y - \mathcal{U}_t x\| \leq \|y - x\| + t\|v\|.$$

In particolare, otteniamo il seguente teorema:

Teorema 2.14. *Sia F una funzione s.c.i. e convessa. Siano poi dati $x_n \rightarrow x$ dei punti di $\overline{D(F)}$, e $v_n \rightarrow v$ dei punti di H . Denotiamo poi con X_t^n il flusso gradiente secondo la funzione F_{v_n} , che parte da x_n . Allora vale che*

$$\|X_t^n - X_t\| \leq \|x_n - x\| + t\|v_n - v\| \rightarrow 0$$

dove X_t è il flusso gradiente secondo F_v che parte da x . In particolare c'è convergenza puntuale per ogni tempo delle traiettorie dei flussi gradiente.

In realtà vale un teorema molto più generale, che però non fornisce informazioni quantitative sulla convergenza dei flussi, ed è strettamente legato alla nozione di Γ -convergenza. Infatti in sostanza dice che “il limite dei flussi gradiente è un flusso gradiente del Γ -limite dei funzionali”. Innanzitutto richiamiamo cos'è Γ -convergenza, una nozione introdotta da De Giorgi che ha avuto grande risonanza e applicazioni nel campo del Calcolo delle Variazioni:

Definizione 2.15. (Γ -convergenza) *Siano $F_n : X \rightarrow \mathbb{R}$ delle funzioni su uno spazio metrico X . Si dice che esse Γ -convergono alla funzione F se sono verificate le due seguenti proprietà:*

- data una qualunque successione x_n che converge a x ,

$$F(x) \leq \liminf_{n \rightarrow \infty} F_n(x_n)$$

- dato $z \in X$, esiste una successione z_n che converge a z tale che

$$F(z) \geq \limsup_{n \rightarrow \infty} F_n(z_n)$$

Questa nozione di convergenza è alquanto particolare, in quanto la successione costante $F_n = G$ non è detto converga a G , in quanto ci converge se e solo se G è semicontinua inferiormente. Vedremo ora che questa nozione è esattamente quello che ci serve appunto per la stabilità generale dei flussi gradiente:

Teorema. Siano u_n delle curve in $\mathcal{AC}_{loc}^p((0, +\infty); H)$ (con $p > 1$), che convergono, puntualmente per ogni $t \geq 0$, ad una curva u . Siano poi F_n delle funzioni convesse ed s.c.i., che Γ -convergono ad F , funzione s.c.i. e convessa anch'essa. Allora se u_n è flusso gradiente di F_n , cioè

$$u_n'(t) \in -\partial F_n(u_n(t)) \quad \text{per quasi ogni } t > 0.$$

allora u è flusso gradiente di F :

$$u'(t) \in -\partial F(u(t)) \quad \text{per quasi ogni } t > 0.$$

Dimostrazione. La dimostrazione è essenzialmente basata sul fatto che una curva assolutamente continua è flusso gradiente se e solo se soddisfa (EVI) in senso distribuzionale. Innanzitutto sappiamo dal teorema 2.3 che u è una curva assolutamente continua. Bisogna dunque verificare solo la disuguaglianza (EVI) in senso distribuzionale. Per fare questo portiamo al limite le disuguaglianze (EVI) per le u_n : sappiamo che, per ogni $\psi \geq 0 \in C_c^\infty(0, +\infty)$, vale:

$$\int_0^{+\infty} \frac{1}{2} \|u_n(t) - y\|^2 \psi'(t) dt \geq \int_0^{+\infty} (F_n(u_n(t)) - F_n(y)) \psi(t) dt \quad \forall y \in H$$

Per ogni $z \in H$ si consideri z_n che soddisfi la seconda condizione della Γ -convergenza e mandiamo al limite la disuguaglianza appena scritta sostituendo $y = z_n$:

$$\begin{aligned} \int_0^{+\infty} \frac{1}{2} \|u(t) - z\|^2 \psi'(t) dt &= \lim_{n \rightarrow \infty} \int_0^{+\infty} \frac{1}{2} \|u_n(t) - z_n\|^2 \psi'(t) dt \\ &\geq \liminf_{n \rightarrow +\infty} \int_0^{+\infty} (F_n(u_n(t)) - F_n(z_n)) \psi(t) dt \\ &\geq \int_0^{+\infty} \liminf_{n \rightarrow +\infty} \{(F_n(u_n(t)) - F_n(z_n)) \psi(t)\} dt \\ &\geq \int_0^{+\infty} (\underline{\lim} F_n(u_n(t)) - \overline{\lim} F_n(z_n)) \psi(t) dt \\ &\geq \int_0^{+\infty} (F(u(t)) - F(z)) \psi(t) dt. \end{aligned}$$

Dunque u è una soluzione (EVI) per F ed è quindi anche flusso gradiente. \square

Capitolo 3

Sticky Particle System

In questo capitolo studieremo il sistema di equazioni differenziali detto Sticky Particle System o, in breve, SPS:

$$\begin{cases} \partial_t \rho_t + \partial_x(v_t \rho_t) = 0 \\ \partial_t(v_t \rho_t) + \partial_x(v_t^2 \rho_t) = 0. \end{cases}$$

Qui ρ_t sono delle misure dipendenti dal tempo e v_t sono delle funzioni definite a meno di insiemi ρ_t trascurabili: incontriamo già ora la prima difficoltà, che risiede nel fatto che uno dei dati appartiene ad uno spazio che cambia nel tempo. Inoltre entrambe le equazioni hanno la forma di legge di conservazione, quindi il problema rientra nella classe dei sistemi di leggi di conservazione, di cui si sa poco o nulla, in particolare non c'è unicità e in generale l'esistenza di soluzioni lisce c'è solo per tempi piccoli, anche se si parte da condizioni iniziali lisce.

Per risolvere questi problemi abbiamo quindi bisogno di :

- selezionare una soluzione particolare, chiamata *soluzione entropica*;
- avere anche una nozione di soluzione debole, che permetta l'esistenza per tutti i tempi.

Iniziamo dal secondo punto: restringiamo innanzitutto la velocità ad essere un elemento di $L^2(\rho_t)$ e supponiamo che ρ_t abbia sempre massa finita che, per la prima equazione, sarà costante e porremo uguale a 1; inoltre assumiamo anche che ρ_t abbia sempre momento secondo finito. Per il primo punto invece consideriamo la derivata spaziale in senso debole e quella temporale in senso forte (sarà chiaro tra poco cosa vuol dire). Prendiamo dunque come definizione di soluzione questa:

Definizione 3.1. Sia (ρ_t, v_t) una curva, dove ρ_t è una misura in $\mathcal{P}_2(\mathbb{R})$ e v_t appartiene a $L^2(\rho_t)$; essa sarà chiamata soluzione di (SPS) se:

- sia ρ_t che $\rho_t v_t$ sono debolmente continue in dualità con le funzioni continue e limitate;
- la funzione $\|v_t\|_{L^2(\rho_t)}^2$ sia in $L^1_{loc}(0, +\infty)$;

- siano soddisfatte, per ogni $\zeta \in C_c^\infty(\mathbb{R})$, le seguenti equazioni in senso distribuzionale:

$$\frac{d}{dt} \left(\int_{\mathbb{R}} \zeta(x) d\rho_t(x) \right) = \int_{\mathbb{R}} \zeta'(x) v_t(x) d\rho_t(x)$$

$$\frac{d}{dt} \left(\int_{\mathbb{R}} \zeta(x) v_t(x) d\rho_t(x) \right) = \int_{\mathbb{R}} \zeta'(x) v_t^2(x) d\rho_t(x).$$

Abbiamo dunque chiarito intendiamo per soluzione debole, mancherebbe ancora di definire cos'è una soluzione entropica.

Per fare questo analizziamo prima casi semplici per cercare di capire come funziona questo sistema e quali possono essere le soluzioni fisicamente accettabili: consideriamo come misura iniziale una somma finita di delte di Dirac, viste come N particelle di massa finita. La funzione v_t ha bisogno di essere definita solo nei punti dove c'è massa e rappresenta la velocità della particella corrispondente. La prima equazione non è altro che la conservazione della massa, mentre la seconda è la conservazione della quantità di moto.

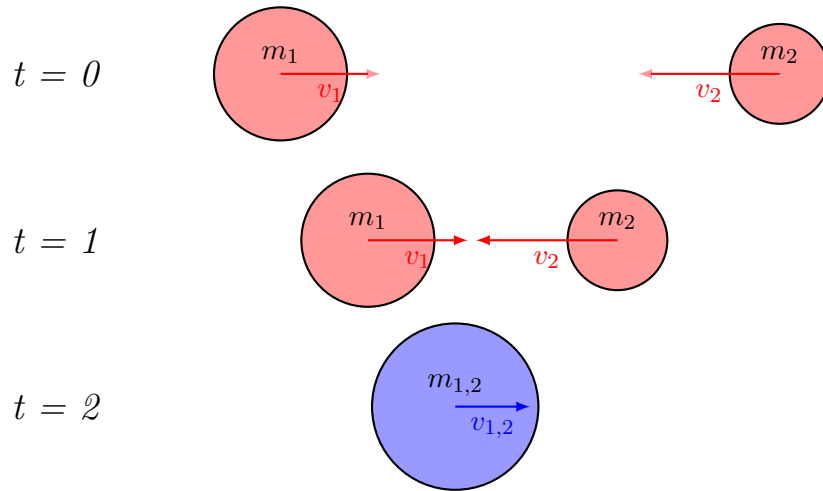
L'evoluzione ovvia è quella di muovere le particelle di moto rettilineo uniforme fino a che non c'è il primo urto tra due particelle: la non unicità della soluzione è data dal fatto che non sappiamo come si comportano gli urti, se in modo elastico o anelastico; come suggerisce il nome *Sticky Particle System*, noi considereremo solo l'evoluzione temporale in cui le particelle, quando avviene un urto, si attaccano conservando la quantità di moto, e quindi gli urti sono completamente anelastici, minimizzando dunque di volta in volta l'energia cinetica. Prendendo questa evoluzione descrittiva, trattata per la prima volta in un modello di Zeldovich [Zel70], abbiamo sicuramente unicità della soluzione nel caso della misura data da un numero finito di particelle; vedremo che questa soluzione è effettivamente la soluzione entropica¹, che, in un certo senso, è sempre quella che minimizza l'energia. Una descrizione più dettagliata di questa evoluzione è trattata nella seconda sezione di questo capitolo.

A motivo di esempio, riportiamo in figura l'evoluzione temporale nel caso di due particelle, di massa m_1 ed m_2 ; quando urtano, esse si fondono in un'unica particella di massa $m_{1,2}$. Le relazioni tra masse iniziali e finali e velocità iniziali e finali sono:

$$m_{1,2} = m_1 + m_2 \quad v_{1,2} = \frac{m_1 v_1 + m_2 v_2}{m_1 + m_2}$$

Cioè le particelle si attaccano (da qui *sticky particle system*) e la quantità di moto totale rimane invariata.

¹nozione, ricordiamo, tipica delle leggi di conservazione



L'approccio usato nell'affrontare questo problema è stato solitamente quello di approssimare dei dati iniziali generali (ρ, v) mediante dati iniziali discreti (ρ^N, v^N) , per i quali si conosce l'evoluzione, per poi cercare di dimostrare che effettivamente le orbite (ρ_t^N, v_t^N) convergono per $N \rightarrow +\infty$ ad una curva (ρ_t, v_t) che soddisfa ancora SPS. Questo metodo è usato per dimostrare che esistono soluzioni che soddisfano opportune condizioni di entropia [Bou94, Bouchut] da Grenier [Gre95] e Weinan, Rykov e Sinai [WRS96] (interessante anche il contributo di Martin e Piasecki [MP94]). Ricordiamo anche gli approcci diversi usati da Bouchut e James [BJ95], Poupaud e Rascle [PR97], e da Sever [Sev01] nel caso multidimensionale; regolarizzazioni viscosi di SPS sono state studiate da Sobolevski [Sob97] e Boudin [Bou00], e un modello differente, basato su particelle di dimensione finita, è stato proposto da Wolansky [Wol07].

I risultati di convergenza sono poi stati estesi e migliorati in Brenier e Grenier [BG98], Huang e Wang [HW01], e Nguyen e Tudorascu [NT08] (con un approccio probabilistico Moutsinga [Mou08] è stato capace di considerare anche velocità con salti negativi in qualsiasi punto del supporto di ρ_0): l'assunzione chiave è che la velocità iniziale v sia limitata, oppure continua con una crescita al più quadratica, e che la successione ρ^N (con la massa totale di ogni misura uguale ad 1) converga a ρ rispetto alla distanza di Wasserstein W_2 nello spazio $\mathcal{P}_2(\mathbb{R})$. Questo include anche il caso (considerato in [9]) di una successione ρ^N con supporto uniformemente limitato e debolmente convergenti a ρ nella dualità con le funzioni continue.

Il cambio di punto di vista adottato in [NS09] rispetto ai precedenti lavori risiede nel fatto di osservare che nel caso di misure discrete l'evoluzione temporale prima del primo urto è una geodetica di velocità costante in $\mathcal{P}_2(\mathbb{R})$, e quindi, come in ogni varietà Riemanniana, è un riscaldamento del flusso gradiente rispetto alla funzione $\phi(\rho) = -\frac{1}{2}W_2^2(\rho, \rho_0)$, che può essere tradotto, come abbiamo già accennato, in un flusso gradiente rispetto ad una funzione convessa nello spazio di Hilbert $H = L^2(0, 1)$.

Sorprendentemente questo flusso gradiente nello spazio di Wasserstein, nel caso di dato iniziale discreto, dopo il primo urto, evolve esattamente come lo Sticky Particle System, come mostrato in [AGS08, Teorema 10.4.12]: l'approccio sarà dunque far vedere, sfruttando le ottime proprietà di stabilità dei flussi gradiente, che anche partendo da dati qualunque, la

soluzione mutuata dal flusso gradiente è effettivamente soluzione di SPS.

La prima sezione sarà dedicata all'analisi accurata di questo flusso gradiente, trovandone anche una descrizione esplicita in termini della proiezione sul cono \mathcal{K} delle funzioni quadrato integrabili essenzialmente crescenti:

$$X(t) = P_{\mathcal{K}}(X_0 + tV_0).$$

Nella seconda parte invece sarà presentato il procedimento di limite per far vedere che in effetti partendo da qualunque dato iniziale la curva data dal flusso gradiente è soluzione debole entropica del SPS.

3.1 Proprietà di \mathcal{K}

Abbiamo chiamato \mathcal{K} l'insieme delle classi funzioni in $L^2(0, 1)$ essenzialmente crescenti, cioè che hanno un rappresentante crescente, che possiamo identificare con quello continuo a sinistra. In questa sezione indaghiamo le proprietà di questo sottoinsieme di $L^2(0, 1)$ e della sua funzione indicatrice $I_{\mathcal{K}}$ definita come:

$$I_{\mathcal{K}}(f) = \begin{cases} 1 & \text{se } f \in \mathcal{K} \\ +\infty & \text{altrimenti.} \end{cases}$$

L'insieme \mathcal{K} risulta essere un convesso e per questo possiamo definire in modo univoco la proiezione su \mathcal{K} :

$$g = P_{\mathcal{K}}f \iff \langle f - g, h - g \rangle \leq 0 \quad \forall h \in \mathcal{K}$$

In questa usuale definizione per la proiezione su un convesso possiamo prendere $h = \lambda g$ con $\lambda > 0$, poiché \mathcal{K} è anche un cono e troviamo che $\langle f - g, g \rangle = 0$, quindi una definizione alternativa risulta essere:

$$g = P_{\mathcal{K}}f \iff \begin{aligned} \langle f - g, h \rangle &\leq 0 \quad \forall h \in \mathcal{K} \\ \langle f - g, g \rangle &= 0. \end{aligned}$$

Sarà utile in seguito introdurre la nozione di cono polare

$$\mathcal{K}^o = \{g \in L^2(0, 1) \mid \langle g, f \rangle \leq 0 \quad \forall f \in \mathcal{K}\},$$

grazie al quale è facile riscrivere in forma compatta

$$g = P_{\mathcal{K}}f \iff f - g \in \langle g \rangle^{\perp} \cap \mathcal{K}^o.$$

Forniamo ora la formula chiusa per la proiezione:

Teorema 3.2. *Data $f(x) \in L^2(0, 1)$, si prenda $F(x)$ una sua primitiva. Denotando con $F^{**}(x)$ l'inviluppo convesso di F , cioè la più grande funzione convessa su $[0, 1]$ minore o uguale ad F , si ha:*

$$P_{\mathcal{K}}f = \frac{d}{dx} F^{**}(x).$$

Dimostrazione. Per dimostrare questo teorema saranno essenziali due fatti sugli involucri convessi di funzioni continue in $[0, 1]$, dimostrati in appendice:

- $F(0) = F^{**}(0)$ e $F(1) = F^{**}(1)$;
- L'insieme $A = \{F(x) > F^{**}(x)\}$ è un aperto e se $(a, b) \subseteq A$ allora $F^{**}|_{(a,b)}$ è una funzione lineare.

Consideriamo la derivata dell'involuppo convesso: $g(x) = \frac{d}{dx} F^{**}(x)$ che risulta una funzione crescente in quanto derivata di una funzione convessa e inoltre $g \in L^2(0, 1)$ da cui $g \in \mathcal{K}$. Consideriamo ora una funzione limitata $h \in \mathcal{K}$ e sia μ_h la sua derivata distribuzionale. Il fatto che h sia essenzialmente crescente si traduce nel fatto che μ_h è una misura positiva. Ora abbiamo che

$$\int_0^1 (f - g)h dx = [(F - F^{**})h]_0^1 - \int_0^1 (F - F^{**})d\mu_h$$

Ma $F(0) = F^{**}(0)$ e $F(1) = F^{**}(1)$ e inoltre $F \geq F^{**}$ da cui $\langle f - g, h \rangle \leq 0$; poiché questo è vero per un denso in \mathcal{K} , è vero anche per tutto \mathcal{K} . Rimane solo da vedere che $\langle f - g, g \rangle \geq 0$. Ma sappiamo che $\mu_g(A) = 0$ poiché g risulta costante nelle componenti connesse di A e per definizione $F = F^{**}$ in A^c da cui la tesi. \square

Con poca fatica in più si può dimostrare il seguente teorema:

Teorema 3.3. *Siano $f, g \in L^2(0, 1)$ e sia $\psi : \mathbb{R} \rightarrow \mathbb{R}$ una funzione s.c.i. e convessa. Allora vale:*

$$\int_0^1 \psi(P_{\mathcal{K}}f - P_{\mathcal{K}}g) dx \leq \int_0^1 \psi(f - g) dx.$$

Dimostrazione. Consideriamo preliminarmente $\psi \in C^1$ e $z \in \mathcal{K}$. Consideriamo poi il seguente integrale:

$$\int_0^1 (f - P_{\mathcal{K}}f)\psi'(P_{\mathcal{K}}f - z).$$

Spezziamo ora l'integrale nella somma degli integrali fatti nelle componenti connesse di $A = \{F(x) > F^{**}(x)\}$, che sono intervalli aperti; infatti il contributo in A^c è nullo, in quanto qui $f = P_{\mathcal{K}}f$ quasi ovunque. Consideriamo una di queste componenti connesse, diciamo che sia (a, b) ; senz'altro abbiamo che $F(a) - F^{**}(a) = F(b) - F^{**}(b) = 0$, e inoltre, per il teorema precedente, sappiamo che $P_{\mathcal{K}}f|_{(a,b)}$ è costante e dunque $\psi'(P_{\mathcal{K}}f - z)$ è una funzione decrescente e dunque la sua derivata distribuzionale, che chiamiamo $\gamma_{a,b}$, sarà negativa o nulla.

Applicando ora l'integrazione per parti otteniamo. per quanto detto:

$$\int_a^b (f - P_{\mathcal{K}}f)\psi'(P_{\mathcal{K}}f - z) = - \int_a^b F - F^{**} d\gamma_{a,b} \geq 0$$

da cui l'integrale considerato in partenza risulta essere positivo o nullo. Applicando ora la disuguaglianza di sottodifferenziale si ha:

$$\begin{aligned}\psi(f - g) &\geq \psi(\mathcal{P}_K f - \mathcal{P}_K g) + (f - \mathcal{P}_K f - (g - \mathcal{P}_K g))\psi'(\mathcal{P}_K f - \mathcal{P}_K g) \\ &= \psi(\mathcal{P}_K f - \mathcal{P}_K g) + (f - \mathcal{P}_K f)\psi'(\mathcal{P}_K f - \mathcal{P}_K g) - (g - \mathcal{P}_K g)\psi'(\mathcal{P}_K f - \mathcal{P}_K g)\end{aligned}$$

ora integriamo questa disuguaglianza in $(0, 1)$ e utilizziamo il lemma appena dimostrato prima con $\psi(x)$ ed f e poi con $\psi(-x)$ e g , per ottenere la tesi. Per una generica ψ , consideriamo le disuguaglianze con le approssimazioni di Moreau-Yosida $\psi_{2n}(x) = \min \psi(y) + n|x - y|^2$, quindi mandiamo al limite, arrivando anche qui al risultato voluto. \square

Questo risultato ci mostra, scegliendo $\psi(x) = |x|^p$, che \mathcal{P}_K è in realtà una contrazione in ogni L^p con $1 \leq p < +\infty$.

Ora quello che ci interessa è indagare ∂I_K , ai fini di dimostrare che la soluzione della SPS con dati discreti è effettivamente un flusso gradiente. Da definizione di sottodifferenziale:

$$\partial I_K(f) = \{h \in L^2(0, 1) \mid I_K(g) \geq I_K(f) + \langle g - f, h \rangle \quad \forall g \in L^2(0, 1)\}.$$

Per identificare meglio cosa sia questo insieme, notiamo che se $g \notin \mathcal{K}$ la disuguaglianza è sempre verificata, quindi nella definizione basta considerare $g \in \mathcal{K}$. Ricordando inoltre che \mathcal{K} è un cono e quindi prendendo $g = \lambda f$ si ottiene che $\langle h, f \rangle = 0$, da cui $\langle g, h \rangle \leq 0$ per ogni $g \in \mathcal{K}$; ricordando la definizione di cono polare, il sottodifferenziale ha una forma piuttosto semplice, che ricorda un po' la formula di proiezione:

$$\partial I_K(f) = \langle f \rangle^\perp \cap \mathcal{K}^\circ.$$

Infatti, considerando una $g \in \mathcal{K}$ e $f, p \in L^2(0, 1)$, tali che $f = g + p$, si ha che:

$$p \in \partial \mathcal{K}(g) \iff P_{\mathcal{K}}(g + p) = g$$

$$f - g \in \partial \mathcal{K}(g) \iff P_{\mathcal{K}}(f) = g$$

Cercheremo ora una ulteriore caratterizzazione del sottodifferenziale, più utile per le nostre applicazioni; per comodità, introduciamo la seguente convenzione: data $f \in L^2(0, 1)$, la lettera maiuscola F indicherà la primitiva di f che si annulla in 0, dunque

$$F(x) = \int_0^x f(t) dt.$$

Allo stesso modo saranno definite G come la primitiva di g , eccetera.

Consideriamo ora i seguenti sottoinsiemi:

$$\mathcal{N}_\Omega = \{g \in L^2(0, 1) : G \geq 0 \text{ in } [0, 1], \quad G = 0 \text{ in } [0, 1] \setminus \Omega\}$$

chiamando $\mathcal{N} = \mathcal{N}_{(0,1)}$ è facile mostrare che:

Lemma 3.4. $\mathcal{K}^\circ = \mathcal{N}$.

Dimostrazione. Prendiamo $f \in \mathcal{K}^o$; poiché tutte le costanti sono in \mathcal{K} allora sicuramente avrò $F(1) = 0$. Considerando poi la funzione crescente $g_t(x) = -\mathbb{1}_{[0,t)}(x)$, la relazione $\langle f, g_t \rangle \leq 0$ diventa proprio $F(t) \geq 0$ e dunque f appartiene a \mathcal{N} .

Sia ora $f \in \mathcal{N}$; voglio verificare che appartiene al cono polare: mi basta provare la disuguaglianza per una qualunque g in $C^1(0,1) \cap \mathcal{K}$, che è un sottoinsieme denso di \mathcal{K} . Si ha:

$$\int_0^1 g(t)f(t)dt = \int_0^1 g'(t)F(t)dt \geq 0 \quad (3.1)$$

dove abbiamo usato che $F(0) = F(1) = 0$ e che sia g' che F sono funzioni positive o nulle per ipotesi. \square

Dall'equazione (3.1) deduciamo inoltre che se $g \in C^1(0,1) \cap \mathcal{K}$, allora il sottodifferenziale nel punto g consta di funzioni f la cui primitiva può essere diversa da zero solo dove $g' = 0$ cioè dove g è localmente costante. Questo ci spinge a definire, per ogni $h \in L^2(0,1)$, gli insiemi

$$\Omega_h = \{\omega \in (0,1) : h \text{ è essenzialmente costante in un intorno di } \omega\},$$

$$\mathcal{H}_h = \{f \text{ è essenzialmente costante nelle componenti connesse di } \Omega_h\}.$$

Chiamando $\mathcal{P}_{\mathcal{H}_h}$ la proiezione sul sottospazio $\mathcal{H}_h \subseteq L^2(0,1)$ è facile verificare che:

$$\mathcal{P}_{\mathcal{H}_h}g(x) = \begin{cases} g(x) & \text{se } x \text{ non appartiene a } \Omega_g \\ \int_I g(y)dy & \text{se } x \in I, \text{ componente connessa di } \Omega_g. \end{cases}$$

Teorema 3.5. $\partial I_{\mathcal{K}}(g) = \mathcal{N}_{\Omega_g} = \mathcal{H}_g^\perp \cap \mathcal{N}$.

Dimostrazione. Mostriamo le varie inclusioni:

1) Notiamo innanzitutto che $\mathcal{N}_{\Omega_g} \subseteq \mathcal{H}_g^\perp \cap \mathcal{N}$.

Per fare questo consideriamo una funzione $h \in \mathcal{N}_{\Omega_g}$, e la sua primitiva H ; consideriamo poi $\mathcal{P}_{\mathcal{H}_g}h$, e la sua primitiva \widehat{H} . Mostriamo che $\widehat{H} \equiv 0$.

Sia $x \in \Omega_g^c$; la funzione indicatrice dell'intervallo $[0, x)$ appartiene allora al sottospazio \mathcal{H}_g , da cui:

$$\widehat{H}(x) = \langle \mathbb{1}_{[0,x)}, \mathcal{P}_{\mathcal{H}_g}h \rangle = \langle \mathbb{1}_{[0,x)}, h \rangle = H(x) = 0$$

Inoltre per la proprietà della proiezione citata prima si ha che, se x è nell'intervallo (a, b) , componente connessa di Ω_g :

$$\mathcal{P}_{\mathcal{H}_g}h(x) = \int_a^b h(y)dy = H(b) - H(a) = 0$$

poiché $a, b \notin \Omega_g$. Questo permette di dedurre che $\widehat{H} \equiv 0$ e dunque $\mathcal{P}_{\mathcal{H}_g}h = 0$, cioè h è ortogonale al sottospazio \mathcal{H}_g .

2) Ovviamente poi vale $\mathcal{H}_g^\perp \cap \mathcal{N} \subseteq \langle g \rangle^\perp \cap \mathcal{N} = \partial I_{\mathcal{K}}(g)$, poiché $g \in \mathcal{H}_g$.

3) Ora resta solo da mostrare che $\partial I_{\mathcal{K}}(g) = \langle g \rangle^{\perp} \cap \mathcal{N} \subseteq \mathcal{N}_{\Omega_g}$.

Applicando il lemma 3.7, infatti possiamo concludere che se $f \in \mathcal{N}$ è ortogonale a g allora vale:

$$\int_0^1 F(x) d\gamma(x) = 0$$

dove, al solito, $F(x) = \int_0^x f(\omega) d\omega$ e invece γ è la derivata distribuzionale di g , ed è quindi una misura di Radon su $(0, 1)$ nonnegativa, il cui supporto è proprio Ω_g . Poiché $F \geq 0$ è una funzione continua, essa dovrà essere nulla sul supporto di γ affinché il suo integrale sia nullo. Quindi ottengo l'ultima inclusione, cioè $\langle g \rangle^{\perp} \cap \mathcal{N} \subseteq \mathcal{N}_{\Omega_g}$.

□

Una conseguenza di questo teorema è il seguente corollario:

Corollario 3.6. *Sia $f \in L^2(0, 1)$. Se $g = \mathcal{P}_{\mathcal{K}}f$ allora vale*

$$g = \mathcal{P}_{\mathcal{H}_g}f \quad \mathcal{H}_g \subseteq \mathcal{H}_f$$

Dimostrazione. Per l'osservazione fatta prima sul legame tra la proiezione su \mathcal{K} e il sotto-differenziale di $\mathcal{I}_{\mathcal{K}}$, sappiamo che $f - g \in \partial \mathcal{I}_{\mathcal{K}}$, ma allora $f - g \in \mathcal{H}_g^{\perp}$ e da questo segue $g = \mathcal{P}_{\mathcal{H}_g}f$.

Per il secondo punto ci dobbiamo ricordare la definizione di $\mathcal{P}_{\mathcal{K}}$: basta dimostrare che se F , la primitiva di f , è lineare in un intorno di ω , lo è anche F^{**} . Se $\omega \in \{F^{**} < F\}$, allora sappiamo già dal teorema 3.2 che F^{**} è lineare in un intorno di ω . Sia invece ω tale che $F^{**}(\omega) = F(\omega)$ e tale che

$$F(x) = m(x - \omega) + F(\omega) \quad \forall x \in \overline{B}(\omega, \epsilon).$$

Consideriamo prima $x \notin \{0, 1\}$. Per definizione di F^{**} come sup delle rette che restano sotto il grafico di F , sappiamo che esistono a_n, b_n tali che

$$F(x) \geq a_n(x - \omega) + b_n \quad \forall x \in [0, 1],$$

tali che $b_n \geq F(\omega) - \frac{1}{n}$; considerando $x = \omega \pm \epsilon$, otteniamo

$$|a_n - m| \leq \frac{1}{n\epsilon},$$

da cui, considerando $n \rightarrow +\infty$, abbiamo

$$F(x) \geq m(x - \omega) + F(\omega) \quad \forall x \in [0, 1].$$

Questo ci permette di dire che $F^{**}|_{[\omega-\epsilon, \omega+\epsilon]} = F$, ed è quindi lineare in un intorno di ω . Un ragionamento simile si usa per mostrare che vale la stessa cosa nel caso in cui ω sia uno degli estremi: ad esempio se $\omega = 0$ si considera la retta di pendenza massima che passi per $(0, F(0))$ e stia sempre sotto il grafico di F , e si dimostra che essa sarà il grafico di F^{**} in un intorno di 0. □

Lemma 3.7. *Sia $g \in \mathcal{K}$ e una funzione $f \in \mathcal{N}$. Indicando con γ la misura di Radon su $(0, 1)$ che identifica la derivata distribuzionale di g , e con F la primitiva di f tale che $F(0) = 0$, si ha che $F \in L^1(\gamma)$ e inoltre vale:*

$$\int_0^1 f(t)g(t)dt = - \int_0^1 F(t)d\gamma(t)$$

Dimostrazione. Abbiamo già dimostrato questo fatto per $g \in C^1([0, 1])$; per farlo in generalità dobbiamo effettuare un'analisi più accurata, poiché γ potrebbe non essere una misura finita. Iniziamo col definire la nostra misura in questo modo, prendendo g come il rappresentante continuo a sinistra:

$$\gamma((a, b]) = g(b) - g(a) \quad \text{per ogni } 0 < a < b < 1$$

Questa misura è sicuramente di Radon su $(0, 1)$ e inoltre verifica

$$\int_0^1 fg = - \int_0^1 Fd\gamma$$

quando f è costante a tratti, a media nulla e a supporto compatto in $(0, 1)$; per approssimazione, la formula vale anche per f continua a supporto compatto e media nulla e quindi anche per $f \in L^2(0, 1)$ a supporto compatto e media nulla. Dimostriamo ora che la formula è vera anche quando $f \in \mathcal{N}$, ma il supporto non è necessariamente compatto: consideriamo delle funzioni nonnegative $\phi_n \in C_c^1(0, 1)$, costantemente uguali a 1 in $[\frac{1}{n}, 1 - \frac{1}{n}]$ e tali che $|\phi'_n| \leq 2n$. Allora vale:

$$\int_0^1 F\phi_n d\gamma = - \int_0^1 f\phi_n g dx - \int_0^1 F\phi'_n g dx. \quad (3.2)$$

Dimostriamo ora che il secondo addendo del secondo membro tende a 0. Infatti

$$\int_0^1 F\phi'_n g dx = \int_0^{\frac{1}{n}} F\phi'_n g dx + \int_{1-\frac{1}{n}}^1 F\phi'_n g dx;$$

lavorando sul primo pezzo (il secondo si fa allo stesso modo ricordando che $F(1) = 0$), si applicano Cauchy-Schwarz e la disuguaglianza di Hardy:

$$\begin{aligned} \left| \int_0^{\frac{1}{n}} F\phi'_n g dx \right| &\leq 2n \left(\int_0^{\frac{1}{n}} \frac{F(x)^2}{x^2} dx \right)^{\frac{1}{2}} \left(\int_0^{\frac{1}{n}} x^2 g(x)^2 dx \right)^{\frac{1}{2}} \\ &\leq 2n \cdot 2 \left(\int_0^{\frac{1}{n}} f(x)^2 dx \right)^{\frac{1}{2}} \frac{1}{n} \|g\|_{L^2(0, \frac{1}{n})} \\ &\leq 4 \|f\|_{L^2(0, \frac{1}{n})} \|g\|_{L^2(0, \frac{1}{n})} \end{aligned}$$

e dunque questo contributo va a 0 quando n tende ad infinito.

Possiamo ora passare al limite la (3.2) e, usando la convergenza dominata al secondo membro e il teorema di Lebesgue al primo membro (sfruttando che $F \geq 0$), si ottiene proprio la tesi:

$$\int_0^1 F(t) d\gamma(t) = - \int_0^1 f(t)g(t)dt.$$

□

Nella prossima sezione avremo più volte bisogno della seguente proposizione:

Lemma 3.8. *Siano $f, g \in \mathcal{K}$, due funzioni essenzialmente crescenti; vale allora che*

$$f_g = f - \mathcal{P}_{\mathcal{H}_g} f \in -\partial\mathcal{L}_{\mathcal{K}}(g)$$

Dimostrazione. Per come è costruita, f_g è sicuramente ortogonale a g ; per la caratterizzazione del sottodifferenziale enunciata prima, basta dimostrare che $f_g \in -\mathcal{N}$. Sia (a, b) una componente connessa di Ω_g ; in questo tratto f_g sarà uguale a $f - c$, dove c è scelto in modo tale che l'integrale di f_g in (a, b) sia nullo. Essendo f una funzione crescente, l'integrale di f_g , che chiameremo F_g , è convesso in questo intervallo e $F_g(a) = F_g(b)$. Da questo deduciamo che

$$F_g(x) = 0 \quad \text{se } x \text{ è in } [0, 1] \setminus \Omega_g$$

E invece è nonpositivo altrove per la convessità di F_g nelle componenti connesse di Ω_g . Effettivamente dunque per definizione vale $f_g \in -\mathcal{N}$, da cui la tesi. □

3.2 Il flusso gradiente come soluzione del SPS

Lo scopo di questa sezione è mostrare che, nel caso in cui la misura di partenza μ_0 sia discreta, la soluzione μ_t del problema delle Sticky Particle System è equivalente ad un flusso gradiente nello spazio $L^2(0, 1)$ delle funzioni X_{μ_t} .

Raggiungeremo questo risultato mediante una descrizione esauriente dell'evoluzione temporale, possibile nel caso semplice in cui μ_0 è discreta. Dopo mostreremo che la soluzione del flusso gradiente con dato iniziale generale risulterà anch'esso soluzione SPS; qui useremo le proprietà di stabilità e convergenza del flusso gradiente.

3.2.1 Descrizione nel caso di misura iniziale discreta

Supponiamo di avere all'inizio N particelle puntiformi, di massa m_i (e massa totale normalizzata ad 1), situate nelle posizioni x_i^0 con velocità iniziale v_i^0 ; le particelle le consideriamo in

ordine da sinistra a destra e quindi si ha $x_i^0 < x_{i+1}^0$. Avremo dunque come condizioni iniziali:

$$\begin{aligned}\rho_0 &= \sum_{i=1}^N m_i \delta(x - x_i^0) \\ v_0 \rho_0 &= \sum_{i=1}^N v_i^0 m_i \delta(x - x_i^0).\end{aligned}$$

Per descrivere l'evoluzione temporale bisogna considerare i tempi in cui avvengono gli urti: chiamiamo questo insieme Υ , che consisterà di un numero finito di tempi, ordinati in modo crescente $0 = t_0 < t_1 < t_2 < \dots < t_M$, dove includiamo anche $t_0 = 0$ per comodità. In particolare, se $t < t_1$ si ha che le particelle si muovono liberamente con velocità costante e quindi:

$$\begin{aligned}\rho_t &= \sum_{i=1}^N m_i \delta(x - x_i^0 - tv_i^0) && \text{per } t < t_1, \\ v_t \rho_t &= \sum_{i=1}^N v_i^0 m_i \delta(x - x_i^0 - tv_i^0) && \text{per } t < t_1.\end{aligned}$$

Consideriamo ora $X_t = X_{\rho_t}$, la pseudo-inversa della funzione di ripartizione di ρ_t , definita al capitolo 1; ricordiamo la sua proprietà fondamentale, cioè che $(X_t)_{\#} \mathcal{L}^1 = \rho_t$. Poi definiamo $V_t = v_t \circ X_t$, di cui dobbiamo verificare la buona definizione; infatti v_t è definita a meno di insiemi ρ_t trascurabili. Ma se avessi v_t e \tilde{v}_t due rappresentanti diversi:

$$\begin{aligned}\int_0^1 |V_t(\omega) - \tilde{V}_t(\omega)|^2 d\omega &= \int_0^1 |v_t(X_t(\omega)) - \tilde{v}_t(X_t(\omega))|^2 dx \\ &= \int_0^{+\infty} |v_t(x) - \tilde{v}_t(x)|^2 d\rho_t(x) = 0\end{aligned}$$

e dunque $V_t = \tilde{V}_t$ Lebesgue-quasi ovunque in $(0, 1)$. Dunque la velocità Lagrangiana V_t è ben definita e in $L^2(0, 1)$. Questa velocità è molto comoda poiché risulta essere proprio la velocità della curva X_t in $L^2(0, 1)$. Difatti è facile vedere che

$$X_t(\omega) = x_{k+1}^0 + tv_{k+1}^0 \quad \text{se } \omega_k < \omega \leq \omega_{k+1}$$

dove ω_k è la massa delle prime k particelle e $\omega_0 = 0$. Da cui ottengo che

$$\frac{d}{dt} X_t(\omega) = V_t(\omega) \quad \text{se } t < t_1.$$

Ovviamente queste osservazioni valgono ogniqualvolta t non è un tempo di collisione.

Per capire cosa succede nei tempi di collisione bisogna definire ancora qualche oggetto: innanzitutto immaginiamo che se al tempo t_1 si urtano le particelle $i, i+1, \dots, j$, esse andranno

a formare una unica particella etichettata con j , il massimo indice tra quelli coinvolti (ovviamente ci possono essere più urti contemporaneamente e facciamo la stessa cosa in tutti); lo stesso per i tempi di collisione successivi, cioè lasciamo l'indice maggiore.

Ad ogni tempo $t_k \in \Upsilon$ è dunque associato un insieme di indici ‘residui’ I_k , tali che $I_{k+1} \subseteq I_k$, e un insieme di “masse cumulate residue”, che chiamiamo $J_k = \{\omega_i \text{ t.c. } i \in I_k\}$. Consideriamo ora i sottospazi $\mathcal{H}_t := \mathcal{H}_{X_t}$. Si ha che

$$\mathcal{H}_{t_k} = \{f \in L^2(0, 1) : f \text{ è localmente costante in } (0, 1) \setminus J_k\}$$

$$\mathcal{H}_t = \mathcal{H}_{t_k} \quad \text{se } t_k \leq t < t_{k+1}.$$

Ovviamente questi spazi sono decrescenti, cioè $\mathcal{H}_s \subseteq \mathcal{H}_t$ quando $s \geq t$.

Questi spazi sono utili per capire come cambia la velocità: consideriamo il primo urto al tempo t_1 e supponiamo collidano insieme le particelle i, \dots, j . Per la conservazione della quantità di moto, caratteristica di questo sistema, si ha che la velocità della nuova particella è:

$$v_j(t_k^+) = \frac{\sum_{k=i}^j m_k v_k(t_k^-)}{\sum_{k=i}^j m_k}$$

ma è facile vedere che questa formula è equivalente a $V_{t_1} = \mathcal{P}_{\mathcal{H}(t_1)}(V_0)$, infatti all'infuori di (ω_i, ω_j) rimane tutto invariato, mentre, in questo intervallo, si sostituisce la funzione con la sua media, che non è altro che la media delle velocità pesate con le masse. Similmente si ha che

$$\begin{aligned} V_{t_k} &= \mathcal{P}_{\mathcal{H}_{t_k}}(V_{t_{k-1}}) = \mathcal{P}_{\mathcal{H}_{t_k}}(V_0) \\ V_t &= \mathcal{P}_{\mathcal{H}_t}(V_0) \end{aligned}$$

Riassumiamo quello che abbiamo osservato finora: nella notazione lagrangiana, se il dato iniziale è una misura discreta, si ha che :

- $\mathcal{H}_t = \mathcal{H}_{X_t}$ sono sottospazi decrescenti di $L^2(0, 1)$;
- $V_t = \mathcal{P}_{\mathcal{H}_t}(V_0)$
- $\frac{d}{dt} X_t(\omega) = V_t(\omega)$, per ogni $t \notin \Upsilon$;
- $\frac{d}{dt}^+ X_t(\omega) = V_t(\omega)$, per qualsiasi t ;

3.2.2 SPS discreto è un flusso gradiente

Dimostriamo ora il seguente teorema:

Teorema 3.9. *Sia (X_t, V_t) la soluzione lagrangiana descritta prima. Indichiamo inoltre con $\tilde{X}_t = X_0 + tV_0$, la soluzione con le particelle libere; sono allora verificate le seguenti:*

- $X_t = \mathcal{P}_{\mathcal{H}_t}(\tilde{X}_t)$;
- $t \frac{d}{dt}^+ X_t \in X_t - X_0 - \partial \mathcal{I}_{\mathcal{K}}(X_t)$;
- $V_t - V_0 \in \partial \mathcal{I}_{\mathcal{K}}(X_t)$;

- $X_t = \mathcal{P}_{\mathcal{K}}(\tilde{X}_t)$;

Dimostrazione. Mostriamo il primo punto, notando prima che X_t è una curva $\|V_0\|_2$ -lipschitziana e dunque vale il teorema fondamentale del calcolo integrale:

$$\begin{aligned} X_t &= X_0 + \int_0^t V_s ds = \mathcal{P}_{\mathcal{H}_t} \left(X_0 + \int_0^t V_s ds \right) \\ &= \mathcal{P}_{\mathcal{H}_t}(X_0) + \mathcal{P}_{\mathcal{H}_t} \left(\int_0^t \mathcal{P}_{\mathcal{H}_s}(V_0) ds \right) \\ &= \mathcal{P}_{\mathcal{H}_t}(X_0) + \left(\int_0^t \mathcal{P}_{\mathcal{H}_t}(V_0) ds \right) \\ &= \mathcal{P}_{\mathcal{H}_t}(X_0 + tV_0) = \mathcal{P}_{\mathcal{H}_t}(\tilde{X}_t) \end{aligned}$$

Per il secondo punto, vogliamo calcolare tV_t , e per farlo lo esplicitiamo dalla formula del punto 1:

$$\begin{aligned} X_t &= \mathcal{P}_{\mathcal{H}_t}(X_0 + tV_0) \\ X_t &= \mathcal{P}_{\mathcal{H}_t}(X_0) + t\mathcal{P}_{\mathcal{H}_t}(V_0) \\ X_t - X_0 &= (\mathcal{P}_{\mathcal{H}_t}(X_0) - X_0) + tV_t \\ tV_t &= X_t - X_0 + (X_0 - \mathcal{P}_{\mathcal{H}_t}(X_0)) \end{aligned}$$

e quindi, per il lemma 3.8, ho la tesi.

Infine, noto che posso scrivere $V_t - V_0$ come $\sum_{i=0}^k (V_{t_{i+1}} - V_{t_i})$ per qualche k (il più grande indice tale che $t \geq t_k$). Ora so che, sostituendo $V_{t_i}(t_{i+1} - t_i) = X_{t_{i+1}} - X_{t_i}$:

$$\begin{aligned} (t_{i+1} - t_i)(V_{t_{i+1}} - V_{t_i}) &= (t_{i+1} - t_i)(\mathcal{P}_{\mathcal{H}_{t_{i+1}}}(V_i) - V_i) \\ &= \mathcal{P}_{\mathcal{H}_{t_{i+1}}}(X_{t_{i+1}} - X_{t_i}) - X_{t_{i+1}} + X_{t_i} \\ &= X_{t_i} - \mathcal{P}_{\mathcal{H}_{t_{i+1}}}(X_{t_i}) \in -\mathcal{N} \end{aligned}$$

per il lemma 3.8 ; dunque, poiché \mathcal{N} è un cono, essendo $V_0 - V_t$ somma di elementi nel cono, appartiene anch'esso al cono. Notando poi che $\mathcal{P}_{\mathcal{H}_t}(V_0 - V_t) = 0$ si ottiene

$$V_0 - V_t \in \mathcal{H}_{X_t}^\perp \cap \mathcal{N} = \partial\mathcal{I}_{\mathcal{K}}(X_t),$$

cioè la tesi del terzo punto. Ora è facile concludere anche l'ultimo punto, osservando che

$$\tilde{X}_t - X_t = \int_0^t (V_0 - V_s) ds \in \mathcal{N}$$

e ricordando la caratterizzazione della proiezione su \mathcal{K} . □

Abbiamo dunque stabilito nel secondo punto del precedente teorema che X_t soddisfa un flusso gradiente riscaldato, in particolare X_{e^t} è un flusso gradiente per $t \in (-\infty, +\infty)$, secondo la funzione $-(X_0 - X)^2 + \mathcal{I}_{\mathcal{K}}(X)$, che però non risulta convessa. Tuttavia per il terzo punto si ha che:

$$V_t - V_0 \in -\partial\mathcal{I}_{\mathcal{K}}(X_t)$$

$$\frac{d}{dt}X_t \in V_0 - \partial\mathcal{I}_{\mathcal{K}}(X_t) \quad \text{per quasi ogni } t,$$

cioè per $t \in [0, +\infty) \setminus \Upsilon$. Questo ci permette di dire che effettivamente X_t è flusso gradiente, per tempi positivi, della funzione convessa

$$F(X) = \mathcal{I}_{\mathcal{K}}(X) - \langle V_0, X \rangle$$

3.3 Approssimazione con dati discreti

Preso un dato iniziale (ρ, v) , consideriamo una successione (ρ^N, v^N) di dati discreti che convergano ad esso: per fare questo è utile prima definire uno spazio ambiente, possibilmente metrico, nel quale avere le coppie misura-velocità. Lo spazio migliore è:

$$\mathcal{V}_2(\mathbb{R}) = \{(\rho, v\rho) \in \mathcal{P}_2(\mathbb{R}) \times \mathcal{M}(\mathbb{R}) : v \in L^2(\rho)\}$$

munito della distanza (da qui in poi $\mu^i = (\rho^i, v^i\rho^i)$)

$$D_2^2(\mu^1, \mu^2) = \int_0^1 (X_{\rho^1} - X_{\rho^2})^2 + (v^1(X_{\rho^1}) - v^2(X_{\rho^2}))^2 dt,$$

dove X_ρ denota, come nel capitolo 1, la pseudo inversa della funzione di ripartizione della misura ρ . Notiamo subito che il primo addendo nel membro di destra non è altro che $W_2^2(\rho^1, \rho^2)$, e la quantità rimanente risulta essere positiva, e la chiameremo $U_2^2(\mu^1, \mu^2)$. In particolare, la convergenza in \mathcal{V}_2 implica la convergenza nello spazio di Wasserstein delle misure.

Inoltre questa topologia è più forte di quella indotta dalla convergenza debole di $v\rho$:

Lemma 3.10. *Sia $\mu_n = (\rho_n, v_n\rho_n)$ una successione di elementi di \mathcal{V}_2 . Essa converge a $\mu = (\rho, v\rho)$ secondo la distanza D_2 se e solo se [AGS08, Prop. 5.4.3]:*

- $W_2(\rho_n, \rho) \rightarrow 0$;
- $v_n\rho_n \rightharpoonup v\rho$, debolmente in $\mathcal{M}(\mathbb{R})$;
- c'è convergenza delle energie: $\|v_n\|_{L^2(\rho_n)} \rightarrow \|v\|_{L^2(\rho)}$.

Il sottoinsieme delle misure discrete

$$\hat{\mathcal{V}}_2 = \left\{ (\rho, v\rho) : \rho = \sum_{i=1}^m m_i \delta_{x_i}(x), m_i > 0, \sum_{i=1}^m m_i = 1 \right\}$$

è denso in \mathcal{V}_2 , ma esso non è completo. Infatti se consideriamo la successione:

$$\mu_t = \left(\frac{1}{2}\delta_{-t} + \frac{1}{2}\delta_{-t}, \frac{1}{2}\delta_t - \frac{1}{2}\delta_{-t} \right),$$

essa è di Cauchy per $t \rightarrow 0$, infatti $D_2(\mu_s, \mu_t) = |t - s|/\sqrt{2}$, tuttavia se convergesse a $\mu = (\rho, v\rho)$, dovrei avere $\rho = \delta_0$, e, per la convergenza debole della seconda componente, $v = 0$. Ma questo va contro il terzo punto della caratterizzazione di convergenza in \mathcal{V}_2 , poiché $\|v_t\| = 1$ mentre $\|v\| = 0$.

Tuttavia questa distanza è comoda poichè, riletta nella notazione lagrangiana, essa risulta essere quella naturale nello spazio prodotto $L^2(0, 1) \times L^2(0, 1)$, a cui appartengono i dati (X_t, V_t) .

Mostriamo ora dei risultati di stabilità, validi al momento per dati iniziali in $\hat{\mathcal{V}}_2$: qui denotiamo con \mathcal{S}_t il semigruppato di evoluzione descritto in precedenza, quello mutuato dal flusso gradiente. Per ora $\mathcal{S}_t : \hat{\mathcal{V}}_2 \rightarrow \hat{\mathcal{V}}_2$, e il nostro scopo sarà estenderlo a tutto \mathcal{V}_2 .

Teorema 3.11. *Siano $\mu^1 = (\rho^1, v^1\rho^1)$ e $\mu^2 = (\rho^2, v^2\rho^2)$ due elementi di $\hat{\mathcal{V}}_2$ e siano $\mu_t^i = (\rho_t^i, v_t^i\rho_t^i) = \mathcal{S}_t(\mu^i)$ le loro evoluzioni temporali. Valgono allora, per ogni costo convesso φ :*

$$\mathcal{C}_\varphi(\rho_t^1, \rho_t^2) \leq \int_{\mathbb{R} \times \mathbb{R}} \varphi(x + tv^1(x) - y - tv^2(y)) d\pi(x, y), \quad \pi = \Gamma_0(\rho^1, \rho^2)$$

$$W_2(\rho_t^1, \rho_t^2) \leq W_2(\rho^1, \rho^2) + tU_2(\mu^1, \mu^2),$$

$$\int_0^t U_2^2(\mu_r^1, \mu_r^2) dr \leq C(1+t)(D_2(\mu^1, \mu_0) + D_2(\mu^2, \mu_0))(W_2(\rho^1, \rho^2) + U_2(\mu^1, \mu^2)),$$

dove C è una costante universale, e $\mu_0 = (\delta_0, 0)$.

Dimostrazione. La prima disuguaglianza discende immediatamente dal teorema 3.3 ricordando applicando la caratterizzazione della soluzione data al quarto punto dal teorema 3.5. La seconda disuguaglianza è una banale applicazione della disuguaglianza triangolare per la norma L^2 , dopo aver usato la prima disuguaglianza nel caso $\varphi(z) = z^2$.

Il terzo punto invece, più tecnico, riguarda in generale una proprietà di stabilità dei flussi gradienti, ed è dimostrata in [Sav93]. \square

Ora, dato che $\mu_0^N \rightarrow \mu_0$ implica

$$X_0^N \rightarrow X_0 \quad \text{in } L^2(0, 1), \quad V_0^N \rightarrow V_0 \quad \text{in } L^2(0, 1)$$

con il teorema 2.14 si vede che c'è convergenza dei flussi gradiente, che inoltre convergono esattamente ad un flusso gradiente. Per la formula della proiezione trovata nella sezione precedente, troviamo che la curva a cui i flussi gradienti convergono è:

$$X(t) = \mathcal{P}_K(X_0 + tV_0)$$

Questo basterebbe a garantire l'esistenza di una soluzione al sistema iniziale SPS se solo valesse $V(t) \in \mathcal{H}_{X(t)}$, dimodoché si possa scrivere $V(t) = v_t(X(t))$ e ritrovare dunque la

descrizione iniziale. Per fare questo, notiamo che, dal punto 2 del teorema 3.9, le soluzioni del SPS con dato iniziale discreto, se riscalate, risolvono anche il flusso gradiente mediante una funzione che abbiamo già notato non essere convessa. Tuttavia è λ -convessa² e questo basta ad avere la medesima stabilità enunciata nel teorema 2.3.1, che ci permette dunque di dire che:

$$tV_t \in X_t - X_0 - \partial\mathcal{I}_{\mathcal{K}}(X_t) \quad \text{per q.o. } t$$

In realtà, ricordando la proprietà di selezione minima enunciata nel capitolo 2, valida anche nel caso λ -convesso, sappiamo che:

$$t \frac{d^+}{dt} X_t = X_t - X_0 - \Xi \quad \forall t > 0$$

Dove Ξ è l'elemento di $\partial\mathcal{I}_{\mathcal{K}}(X_t)$ che rende minima la norma

$$\|X_t - X_0 - \Xi\|.$$

In questo caso siamo fortunati e riusciamo a trovare chi è esattamente Ξ , e dunque V_t . Come già dimostrato in precedenza, sappiamo che $\partial\mathcal{I}_{\mathcal{K}}(X_t) \subseteq \mathcal{H}_{X_t}^\perp$ e dunque possiamo scrivere:

$$X_t - X_0 - \Xi = (X_t - \mathcal{P}_{\mathcal{H}_{X_t}}(X_0)) + (\mathcal{P}_{\mathcal{H}_{X_t}}(X_0) - X_0 - \Xi)$$

Ora, nella prima parentesi ci due elementi di \mathcal{H}_{X_t} , mentre nella seconda parentesi ci sono due vettori di $\mathcal{H}_{X_t}^\perp$ (che sono Ξ , per quanto detto prima, e $\mathcal{P}_{\mathcal{H}_{X_t}}(X_0) - X_0$, per definizione di proiezione su un sottospazio vettoriale): dunque è facile calcolarsi la norma che risulta essere:

$$\|X_t - X_0 - \Xi\|^2 = \|X_t - \mathcal{P}_{\mathcal{H}_{X_t}}(X_0)\|^2 + \|\mathcal{P}_{\mathcal{H}_{X_t}}(X_0) - X_0 - \Xi\|^2.$$

Dal teorema 2.14 sappiamo che $\mathcal{P}_{\mathcal{H}_{X_t}}(X_0) - X_0$ è un elemento del sottodifferenziale di $\mathcal{I}_{\mathcal{K}}$ in X_t e quindi sicuramente per $\Xi = \mathcal{P}_{\mathcal{H}_{X_t}}(X_0) - X_0$ ottengo la norma minima. Quindi, il risultato finale risulta essere:

$$tV_t = X_t - \mathcal{P}_{\mathcal{H}_{X_t}}(X_0) \quad \forall t \in (0, +\infty) \setminus \Theta,$$

dove Θ è l'insieme dei tempi dove c'è discontinuità della norma di V_t , cioè una quantità al più numerabile di tempi. Se inoltre definisco V_t come la derivata destra di X_t ho che la formula precedente vale per tutti i tempi positivi.

Ora ricordiamo che $\mathcal{P}_{\mathcal{K}}(X_0 + tV_0) = X_t$ e, utilizzando il corollario 3.6, otteniamo che

$$X_t = \mathcal{P}_{\mathcal{H}_t}(X_0 + tV_0) = \mathcal{P}_{\mathcal{H}_t}(X_0) + t\mathcal{P}_{\mathcal{H}_t}(V_0),$$

da cui otteniamo, confrontando con la formula di prima: $V_t = \mathcal{H}_t(V_0)$. Inoltre si ha anche che

$$\mathcal{H}_t = \mathcal{H}_{X_t} \subseteq \mathcal{H}_{X_0 + tV_0} \subseteq \mathcal{H}_{X_0} = \mathcal{H}_0$$

²Una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$, si dice λ -convessa, se per ogni $x, y \in \mathbb{R}$, vale:

$$f(tx + (1-t)y) \leq g e q t f(x) + (1-t)f(y) - \frac{1}{2}\lambda t(1-t)|x-y|^2 \quad \forall t \in [0, 1]$$

Inoltre, usando l'ultima disuguaglianza del teorema 3.11 è facile anche vedere che le anche le velocità convergono in L^2 per quasi ogni tempo, e quindi possiamo anche considerare il problema le cui condizioni iniziali sono $(\nu_0^N, u_0^N) = (\rho_t^N, v_t^N)$, e quindi dimostrare che valgono, per ogni $s < t$:

$$\mathcal{H}_t \subseteq \mathcal{H}_s, \quad V_t = \mathcal{P}_{\mathcal{H}_t} V_s, \quad X_t = \mathcal{P}_{\mathcal{K}}(X_s + (t-s)V_s).$$

Riassumendo, possiamo enunciare il seguente risultato di esistenza:

Teorema 3.12. (Esistenza) *Sia ρ_0 una misura in $\mathcal{P}_2(\mathbb{R})$ e sia v_0 una funzione in $L^2(\rho_0)$. Allora per ogni tempo $t > 0$ esistono una misura ρ_t in $\mathcal{P}_2(\mathbb{R})$ e una funzione v_t in $L^2(\rho_t)$ che risolvono SPS con dato iniziale (ρ_0, v_0) .*

Dimostrazione. Bisogna dimostrare che effettivamente valgono le formule di inizio capitolo (definizione 6): tuttavia, le derivate temporali, sono da intendere come derivate destre in quanto si tratta di problemi di evoluzione. Detto ciò, per verificare le due identità, usiamo essenzialmente tre fatti:

- $\frac{d^+}{dt} X_t = V_t = v_t \circ X_t$, per ogni tempo $t \geq 0$;
- $V_t = \mathcal{P}_{\mathcal{H}(t)} V_0$;
- presa $\zeta(x) \in C_b(\mathbb{R})$ e una successione f_n che converge in probabilità ad f , allora $\zeta \circ f_n \xrightarrow{*} \zeta \circ f$, in L^∞ .

I primi due punti sono già stati dimostrati in precedenza, dunque dimostriamo l'ultimo punto: sappiamo che se una successione di variabili aleatorie f_n tende in probabilità ad f , allora $\zeta \circ f_n \rightarrow \zeta \circ f$ in tutti gli L^p , con $p < \infty$. In particolare sarà vero anche in L^2 ; consideriamo ora i funzionali

$$T_n : L^1 \rightarrow \mathbb{R} \quad T_n(g) = \int_X (\zeta \circ f_n - \zeta \circ f)g$$

Ovviamente le mappe T_n sono equilimitate e, ristrette a L^2 , tendono puntualmente a 0. Per densità dunque $T_n \xrightarrow{*} 0$, cioè la tesi.

Sia ora ζ una funzione in $C^2(\mathbb{R})$, con derivata seconda limitata. Sappiamo allora, per la formula di Taylor, per qualche c compreso tra $X_t(\omega)$ e $X_s(\omega)$, vale:

$$\zeta(X_t(\omega)) - \zeta(X_s(\omega)) - \zeta'(X_t(\omega))(X_t(\omega) - X_s(\omega)) = \frac{1}{2} \zeta''(c)(X_t(\omega) - X_s(\omega))^2.$$

Integrando ora in ω e dividendo per $t-s$ otteniamo:

$$\left| \frac{\int_0^1 \zeta \circ X_t d\omega - \int_0^1 \zeta \circ X_s d\omega}{t-s} - \int_0^1 \zeta' \circ X_t \frac{X_t - X_s}{t-s} d\omega \right| \leq \frac{M}{t-s} \|X_t - X_s\|^2.$$

Facendo tendere s a t da destra e usando la lipschitzianità della curva X_t nel tempo otteniamo che il membro di destra tende a zero e quindi, con la derivabilità forte della curva X_t , possiamo dire che:

$$\frac{d^+}{dt} \int_0^1 \zeta \circ X_t d\omega = \int_0^1 \zeta' \circ X_t \cdot V_t d\omega,$$

che, usando $(X_t)_\# \mathcal{L} = \rho_t$ e $V_t = v_t(X_t)$ fa ottenere proprio la prima equazione:

$$\frac{d^+}{dt} \int_{\mathbb{R}} \zeta(x) d\rho_t(x) = \int_{\mathbb{R}} \zeta'(x) \cdot v_t(x) d\rho(x).$$

Per la seconda equazione sfruttiamo un procedimento analogo: innanzitutto però notiamo che

$$\int_0^1 \zeta(X_t) V_t d\omega = \int_0^1 \zeta(X_t) V_0 d\omega$$

in virtù del fatto che $V_t = \mathcal{P}_{\mathcal{H}(t)} V_0$. Stavolta l'assunzione che facciamo su ζ deve essere un po' più forte, cioè $\zeta \in C^1(\mathbb{R})$, con derivata limitata; la scelta è dettata da questioni di integrabilità. Consideriamo ora la funzione $X_{t,s}$ tale che risolva la seguente equazione per ogni ω :

$$\zeta(X_t(\omega)) - \zeta(X_s(\omega)) = \zeta'(X_{t,s}(\omega))(X_t(\omega) - X_s(\omega))$$

in modo tale che $X_{t,s}$ risulti sempre essere compreso tra X_t e X_s . In questo modo si ha che $X_{t,s} \rightarrow X_t$ in L^2 e quindi, anche in probabilità e, per il lemma di prima dunque $\zeta'(X_{t,s}(\omega)) \rightarrow \zeta'(X_t(\omega))$ debolmente-* in $(L^1)^*$. Poiché $V_0 \frac{X_t(\omega) - X_s(\omega)}{t-s} \rightarrow V_0 V_t$ in L^1 , possiamo concludere che:

$$\begin{aligned} \frac{d^+}{dt} \int_{\mathbb{R}} \zeta(x) v_t(x) d\rho_t(x) &= \lim_{s \rightarrow t^+} \frac{1}{t-s} \int_0^1 V_0 \cdot (\zeta \circ X_t - \zeta \circ X_s) d\omega \\ &= \lim_{s \rightarrow t^+} \int_0^1 \zeta'(X_{t,s}) V_0 \frac{X_t - X_s}{t-s} d\omega \\ &= \int_0^1 \zeta'(X_t) V_0 V_t d\omega \\ &= \int_0^1 \zeta'(X_t) V_t^2 d\omega = \int_{\mathbb{R}} v_t^2(x) \zeta'(x) d\rho(x) \end{aligned}$$

e quindi anche la seconda equazione è verificata. \square

Con gli stessi accorgimenti di quest'ultima dimostrazione, si può facilmente vedere che, se $t \rightarrow 0$, allora:

$$\rho_t \rightarrow \rho \quad v_t \rho_t \rightarrow v_0 \rho_0 \quad v_t^2 \rho_t \rightarrow v_0^2 \rho_0.$$

Inoltre è facile verificare dall'identità $tV_t = X_t - \mathcal{P}_{\mathcal{H}_t}(X_0)$ che vale la condizione di Oleinik sulla velocità:

$$\frac{v_t(x_1) - v_t(x_2)}{x_1 - x_2} \leq \frac{1}{t} \quad \text{per } \rho_t\text{-quasi ogni } x_1 < x_2.$$

Questa condizione, aggiunta a $v_t^2 \rho_t \rightarrow v_0^2 \rho_0$, permette di far vedere, per [HW01] che questa soluzione è effettivamente l'unica soluzione entropica, almeno nel caso in cui la velocità iniziale v_0 è limitata. Un prossimo passo da fare è quello di dimostrare l'unicità anche nel caso delle nostre condizioni iniziali.

Richiamiamo ora il concetto di soluzione entropica.

3.4 Condizioni di entropia

Descriveremo di seguito le varie definizioni di soluzione entropica data negli anni per il sistema (SPS). Storicamente, la condizione di entropia considerata inizialmente è quella data dalla vanishing viscosity; la introduciamo nel contesto delle leggi di conservazione scalari in 1-D.

3.4.1 Vanishing viscosity

Supponiamo di voler risolvere la legge di conservazione scalare

$$\partial_t u + \partial_x(F(u)) = 0 \quad (3.3)$$

alla quale affianchiamo la condizione iniziale $u(0, x) = u_0(x)$. È fatto noto che anche se sia la condizione iniziale che il flusso sono funzioni lisce, non è detto che esista una soluzione liscia per tutti i tempi. In questo caso la soluzione generalizzata è da cercare nella classe BV , delle funzioni a variazione limitata. Allargando tuttavia il dominio del problema e ammettendo soluzioni in tempo debole, viene meno l'unicità (si veda [Eva10] per maggiori approfondimenti); tuttavia le leggi di conservazione scalare governano solitamente fenomeni fisici, dove un'unica evoluzione è ammessa.

Per determinare qual è la soluzione significativa, vengono chiamate in gioco alcune disuguaglianze legate alle quantità fisiche, che svolgono il ruolo del secondo principio della termodinamica, e per questo le soluzioni che vengono trovate vengono chiamate entropiche. Matematicamente, un modo per spiegare queste disuguaglianze viene dalla vanishing viscosity. Immaginiamo che il nostro sistema sia in realtà idealizzato e in realtà dovrebbe essere presente nell'equazione anche un termine *viscoso* appunto, proporzionale a $\partial_{xx}u$. Consideriamo dunque la soluzione u_ε , unica poiché l'equazione diviene parabolica, del problema perturbato

$$\partial_t u + \partial_x(F(u)) = \varepsilon \partial_{xx}u$$

Se poi $u_\varepsilon \rightarrow \bar{u}$ uniformemente, allora, con opportune ipotesi su F , \bar{u} risulta essere proprio la nostra soluzione viscosa. Sia poi (S, R) una coppia entropia-flusso per 3.3, cioè tale che S è convessa regolare e $R'(x) = F'(x)S'(x)$. Moltiplicando l'equazione perturbata per $S'(u)$, dove S è una funzione convessa regolare, svolgendo i calcoli, otteniamo

$$\partial_t(S(u_\varepsilon)) + \partial_x(R(u_\varepsilon)) = \varepsilon \partial_{xx}(S(u_\varepsilon)) - \varepsilon S''(u_\varepsilon) |\partial_x u_\varepsilon|^2 \leq \varepsilon \partial_{xx}(S(u_\varepsilon))$$

Portando al limite questa disuguaglianza distribuzionalmente si trova che

$$\partial_t(S(u)) + \partial_x(R(u)) \leq 0,$$

che è la condizione di entropia classica richiesta ad u per essere soluzione entropica.

3.4.2 Soluzione entropica di SPS

Nel caso delle equazioni di Eulero senza pressione, con un procedimento simile si arriva, come in [Bou94], a una condizione di entropia del tipo

$$\partial_t(\rho S(v)) + \partial_x(\rho v S(v)) \leq 0 \quad \forall S \text{ funzione convessa } C^1$$

tuttavia è stato verificato che questa condizione da sola non garantisce l'unicità mentre la condizione giusta dovrebbe essere una condizione di tipo Oleinik, in particolare in [WRS96] viene suggerito che la condizione debba essere:

$$\partial_x v_t < \frac{1}{t}$$

In effetti, in [WD97] e [BJ95] viene effettivamente dimostrata l'unicità di soluzioni che soddisfano

$$\frac{v_t(x_1) - v_t(x_2)}{x_1 - x_2} \leq \frac{1}{t}, \quad (3.4)$$

tuttavia la loro dimostrazione mostra l'unicità della soluzione entropica quando il dato iniziale ρ_0 è una funzione misurabile limitata; questo risultato sembra restrittivo, in quanto la soluzione è in generale una misura.

Il significato fisico di questa condizione lo possiamo intuire considerando il sistema con le particelle discrete. Ad un certo istante t_0 consideriamo due particelle successive che sono nelle posizioni, rispettivamente x_1 e x_2 : se non fosse verificata la condizione (3.4) allora, ripercorrendo le traiettorie delle due particelle andando indietro nel tempo, esse si incontrerebbero ad un tempo $t_1 > 0$, per cui, tornando all'evoluzione andando avanti nel tempo, significherebbe che una particella al tempo t_1 si è divisa in due particelle più piccole, cosa che violerebbe un principio di "stabilità in avanti" del sistema.

Tuttavia lo stesso ragionamento può essere fatto nel caso di uguaglianza nella (3.4), nel quale caso avremmo un identico comportamento, ma con $t_1 = 0$, che comunque non va bene. Infatti nel caso generale la condizione (1) non è sufficiente: considerando $(\rho_0, v_0) = (\delta_0, 0)$, le curve parametrizzate da $\alpha > 0$

$$(\rho_t, v_t) = \left(\frac{1}{2}(\delta_{\alpha t} + \delta_{-\alpha t}), \alpha \operatorname{sgn}(x) \right)$$

risultano essere soluzioni deboli di SPS che soddisfano anche la condizione (3.4). Esse descrivono la situazione in cui una particella ferma al tempo $t = 0$, si divide istantaneamente in due particelle uguali più piccole. Per evitare questo tipo di fenomeni in [HW01] si considera la condizione ulteriore:

$$\rho_t v_t^2 \rightharpoonup \rho_0 v_0^2$$

che può rappresentare in un certo senso una condizione di preservazione dell'energia. In questo modo, nell'esempio precedente, non possiamo avere le situazioni in cui si dividono le particelle. Infatti Huang e Wang dimostrano che c'è unicità della soluzione che soddisfi la condizione d'entropia e la condizione iniziale sull'energia, con dati iniziali $\rho_0 \geq 0$ in $\mathcal{M}_{loc}(\mathbb{R})$ e v_0 semplicemente limitata; nel caso in cui ρ_0 sia una misura di probabilità con momento

secondo finito, per quanto detto nella sezione precedente, quest'unica soluzione coinciderà con quella trovata mediante il flusso gradiente.

Ulteriori sviluppi potrebbero essere volti a dimostrare che nelle ipotesi più generali in cui abbiamo trovato soluzione col flusso gradiente, c'è unicità della soluzione entropica, e a trovare un'altra definizione di soluzione entropica, più naturale e più fisica. Ipotesi che funzionano nel caso discreto e potrebbero funzionare anche nel caso continuo sono:

- la stessa condizione di Oleinik, ma con una disuguaglianza stretta, dunque richiedere che valga:

$$\frac{u(x_1, t) - u(x_2, t)}{x_1 - x_2} < \frac{1}{t} \quad \text{per } \rho_t\text{-q.o. ogni } x_1 < x_2, \quad \text{per ogni } t > 0.$$

Infatti il problema della condizione di Oleinik senza aggiungere la debole continuità a destra di $v_t^2 \rho_t$, è non funzionare per misure iniziali discrete, caso in cui la condizione è soddisfatta anche quando c'è suddivisione di una particella iniziale;

- ricordando che la soluzione con dato discreto è quella che ha urti totalmente anelastici, potrebbe essere interessante guardare le soluzioni che minimizzano l'energia cinetica, cioè chiamando \mathcal{U}_t l'insieme delle soluzioni deboli della (SPS) definite su $[0, T]$, considerare

$$\mathcal{U}_T^* = \arg \min_{u \in \mathcal{U}_T} \left\{ \int_0^T \left(\int_{\mathbb{R}} v_t^2 d\rho_t \right) dt \right\}$$

e tentare di dimostrare che \mathcal{U}_T^* consiste di un solo elemento u_T , e che $u_{t_0} = u_{t_1}|_{[0, t_0]}$ ogni volta che $t_0 \leq t_1$. Un approccio del genere ricorda un "principio di minima azione" nello spazio di Wasserstein, tecnica che è presente ad esempio in [GNT09], in un caso delle equazioni di Eulero, con un termine di carica costante di background.

3.5 SPS e leggi di conservazione scalare

Come rilevato in [BG98], si può associare ad ogni (SPS) una legge conservativa. Definiamo le funzioni:

$$M(t, x) = \int_0^x d\rho_t \quad A(m) = \int_0^m v_0(F_{\rho_0}(s)) ds = \int_0^m V_0(s) ds$$

cioè $M(t, x)$ è la massa presente prima di x (o in x) all'istante t , mentre $A(m)$ è la quantità di moto totale delle particelle che, partendo da sinistra, nel complesso pesano esattamente m . Nel lavoro citato si afferma che

$$\partial_t M + \partial_x(A(M)) = 0$$

se e solo se la coppia ρ_t, v_t risolve (SPS), dove

$$\rho_t = \partial_x M \quad v_t \rho_t = \partial_x A(M).$$

Più nello specifico, loro dimostrano che se da una soluzione entropica della legge conservativa allora si trova una soluzione entropica di (SPS), mentre se da una soluzione entropica di (SPS) si ha semplicemente una soluzione (non è detto che sia entropica) della legge conservativa. Con i risultati che abbiamo trovato invece riusciamo a dimostrare che:

Teorema 3.13. *Sia $(\rho_t, v_t \rho_t)$ una curva debolmente continua, con ρ_0 a supporto compatto e v_0 limitata superiormente. Allora, definendo M, A come prima, si ha che M è soluzione entropica di:*

$$\partial_t M + \partial_x(A(M)) = 0 \quad (1)$$

se e solo se $(\rho_t, v_t \rho_t)$ è la soluzione di (SPS) descritta finora. In particolare se ho ρ_0^N misure discrete a supporto compatto che convergono a ρ_0 , allora, per v_0 funzione continua, si ha che la soluzione di SPS a cui convergono le evoluzioni “sticky” di $(\rho_0^N, v_0 \rho_0^N)$, genera una soluzione entropica della (1)

Dimostrazione. Consideriamo $\Psi_0(x) = \int_0^x M(0, y) dy$. Essa è una funzione convessa lipschitziana; anche A risulta essere lipschitziana, per la limitatezza di v_0 . Ma allora, posso applicare la seconda formula di Hopf, per stabilire che

$$\Psi(t, x) = (\Psi_0^* + tA)^*$$

è soluzione viscosa dell’equazione di Hamilton-Jacobi

$$\partial_t \Psi + A(\partial_x \Psi) = 0.$$

Ora ricordiamo, come mostrato in appendice, che Ψ è soluzione viscosa dell’equazione di Hamilton-Jacobi se e solo se $\partial_x \Psi = M$ è soluzione entropica della legge di conservazione associata, che risulta essere proprio

$$\partial_t M + \partial_x(A(M)) = 0$$

Chiamiamo ora $X(t, x)$ la pseudoinversa di M (nella variabile spaziale). Definiamo poi $F(t, x) = \int_0^x X(t, s) ds$, e, ricordando le proprietà richiamate nell’appendice sulla trasformata di Legendre, troviamo che

$$\Psi_0^* = F_0 \quad \Psi(t, \cdot)^* = F(t, \cdot)$$

e quindi deduciamo:

$$F(t, x) = (F_0 + tA)^{**},$$

derivando ora in x riconosciamo la formula, già incontrata prima:

$$X(t, \cdot) = \frac{d^+}{dt} (F_0 + tA)^{**} = \mathcal{P}_K(X_0 + tV_0)$$

e dunque effettivamente $(\rho_t, v_t \rho_t)$ è soluzione entropica di (SPS). Notiamo che tutti i passaggi fatti sono dei se e solo se e dunque la tesi del teorema. □

Appendice A

Teoria della misura

Lemma (Ulam). *Sia X spazio polacco e $\mu \in \mathfrak{M}_+(X)$. Per ogni $\varepsilon > 0$ esiste $K \subseteq X$ compatto tale che $\mu(X \setminus K) \leq \varepsilon$.*

Dimostrazione. Essendo X separabile, in quanto polacco, esiste una successione di punti $\{x_i\}_{i \in \mathbb{N}}$ densa in X . Sempre poiché X è polacco, esiste una distanza d che lo rende completo e che induce la topologia di X . Allora per ogni numero naturale k abbiamo che:

$$\bigcup_{i=0}^{\infty} \bar{B}(x_i, \frac{1}{k}) = X$$

Definiamo ora $K_{k,n} = \bigcup_{i=0}^n \bar{B}(x_i, \frac{1}{k})$: allora $K_{k,n} \uparrow X$ al crescere di n , da cui $\mu(K_{k,n}) \uparrow \mu(X)$. Quindi esiste un numero $N(k)$ tale che $\mu(X \setminus K_{k,N(k)}) \leq \varepsilon 2^{-k}$. Consideriamo ora l'insieme:

$$K = \bigcap_{k=1}^{\infty} K_{k,N(k)}$$

Senza altro K così definito è chiuso infatti $K_{k,N(k)}$ è chiuso in quanto unione finita di chiusi, e l'intersezione infinita di chiusi è ancora chiusa. Inoltre risulta essere totalmente limitato, infatti dato $\nu > 0$ esiste k_0 tale che $\nu > \frac{1}{k_0}$ e allora

$$K \subseteq K_{k_0, N(k_0)} \subseteq \bigcup_{i=0}^{N(k_0)} \bar{B}(x_i, \nu)$$

Quindi K risulta compatto e inoltre

$$\mu(K^c) = \mu\left(\bigcup K_{k,N(k)}^c\right) \leq \sum \mu(K_{k,N(k)}^c) \leq \varepsilon,$$

da cui la tesi. □

Lemma A.1. *Sia $\mu \in \mathfrak{M}_+(X)$. Preso un boreliano E e un $\varepsilon > 0$ esistono un compatto K e un aperto A con $K \subseteq E \subseteq A$ tali che $\mu(A \setminus K) \leq \varepsilon$*

Dimostrazione. Dimostriamo prima una cosa più debole, dove al posto di compatto richiediamo che C sia soltanto chiuso.

Definiamo \mathcal{F} la famiglia degli insiemi che soddisfano le tesi del lemma (con C chiuso). Distinguiamo la dimostrazione in 2 parti: nella prima dimostriamo che $\tau \subseteq \mathcal{F}$, cioè la famiglia considerata contiene tutti gli aperti (che sono un π -system). Nella seconda invece dimostriamo che la famiglia \mathcal{F} è una *classe di Dynkin*.

Grazie al teorema di Dynkin (ci dice che se $\mathcal{A} \subseteq \mathcal{B}$ e \mathcal{A} è un π -system e \mathcal{B} è una classe di Dynkin allora $\sigma(\mathcal{A}) \subseteq \mathcal{B}$) concluderemo che tutti i boreliani soddisfano il lemma.

- (1) Dato un aperto $\Omega \in \tau$ possiamo senz'altro prendere $A = \Omega$ come aperto che lo contiene. Definisco poi

$$C_n = \left\{ x \in \Omega \mid d(x, \Omega^c) \geq \frac{1}{n} \right\}$$

I C_n sono chiusi perché la condizione che li descrive è chiusa e inoltre dato un punto $x_0 \in \Omega$ per definizione di aperto in uno spazio metrico, esiste un $n \in \mathbb{N}$ tale che $B(x_0, 1/\bar{n}) \subseteq \Omega$ e quindi $x_0 \in C_{\bar{n}}$.

Dunque ho che $C_n \uparrow \Omega$ che implica $\mu(C_n) \uparrow \mu(\Omega)$ e allora per ogni $\varepsilon > 0$ esiste un n_0 tale che $\mu(A \setminus C_{n_0}) = \mu(\Omega \setminus C_{n_0}) \leq \varepsilon$.

- (2) Una classe di Dynkin è per definizione chiusa per complementare, comprende X ed è stabile per unioni numerabili disgiunte. \mathcal{F} comprende senz'altro X (basta prendere $A = C = X$) e se $E \in \mathcal{F}$ allora $E^c \in \mathcal{F}$: per ogni ε per ipotesi esistono $C \subseteq E \subseteq A$ tali che $\mu(A \setminus C) \leq \varepsilon$, ma allora considero A^c (che è chiuso) e C^c (che è aperto) e ho che $\mu(C^c \setminus A^c) = \mu(A \setminus C) \leq \varepsilon$ e ovviamente $A^c \subseteq E^c \subseteq C^c$.

L'unica verifica più complessa è quella dell'unione numerabile disgiunta: Sia $E = \bigcup E_n$ e gli $\{E_n\}_{n \geq 1}$ sono disgiunti e verificano le ipotesi.

Allora esistono degli A_n aperti tali che $\mu(A_n \setminus E_n) \leq \varepsilon 2^{-n}$, ed esistono C_n chiusi tali che $\mu(E \setminus C_n) \leq \varepsilon 2^{-n}$; inoltre esiste n_0 un intero tale che $\mu(E \setminus \bigcup_{n \leq n_0} E_n) \leq \varepsilon$.

Considero ora $A = \bigcup A_n$ che è un aperto contenente E e $C = \bigcup_{n \leq n_0} C_n$ che è un chiuso contenuto in E . Ora verifico che $\mu(A \setminus C) \leq \varepsilon$:

$$\begin{aligned} \mu(A \setminus C) &= \sum_{n \leq n_0} [\mu(A_n \setminus E_n) + \mu(E_n \setminus C_n)] + \sum_{n \geq n_0} [\mu(A_n \setminus E_n) + \mu(E_n)] \\ &\leq \sum_{n \geq 1} \mu(A_n \setminus E_n) + \sum_{n \leq n_0} \mu(E_n \setminus C_n) + \sum_{n \geq n_0} \mu(E_n) \\ &\leq \sum_{n \geq 1} \varepsilon 2^{-n} + \sum_{n \leq n_0} \varepsilon 2^{-n} + \mu(E \setminus \bigcup_{n \leq n_0} E_n) \\ &\leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon \end{aligned}$$

da cui concludo per l'arbitrarietà di ε .

Quindi abbiamo dimostrato il lemma con l'ipotesi indebolita di C chiuso invece che compatto. Per finire basta usare il lemma di Ulam, che ci dice che $\forall \varepsilon > 0$ esiste un compatto K_ε tale che

$\mu(K_\varepsilon^c) \leq \varepsilon$: dato il mio boreliano E e dato $\varepsilon > 0$ prendo ora un aperto $A_{\varepsilon/2}$ che lo contiene e un chiuso $C_{\varepsilon/2}$ contenuto tali che $\mu(A_{\varepsilon/2} \setminus C_{\varepsilon/2}) \leq \varepsilon/2$ e chiamo $A^\varepsilon = A_{\varepsilon/2}$ e $K^\varepsilon = K_{\varepsilon/2} \cap C_{\varepsilon/2}$ (che stavolta è un compatto contenuto in E); allora vale $\mu(A^\varepsilon \setminus K^\varepsilon) \leq \varepsilon$. \square

Appendice B

Curve assolutamente continue in spazi di Hilbert

Lemma B.1. *Sia $u \in \mathcal{AC}([0, 1]; H)$. Allora esiste il limite forte del rapporto incrementale in quasi ogni punto*

$$s - \lim_{h \rightarrow 0} \frac{u(t+h) - u(t)}{h} = u'(t)$$

e vale il teorema fondamentale del calcolo integrale:

$$\int_s^t u'(r) dr = u(t) - u(s) \quad \text{per ogni } 0 \leq s < t \leq 1$$

Dimostrazione. Si consideri $H_1 = \overline{\text{span}(u([0, 1]))}$ che risulta essere uno spazio separabile. Consideriamo poi una sua base ortonormale e_1, e_2, \dots , costituita da una quantità numerabile di vettori, vista la separabilità dello spazio. Analizziamo la nostra curva nelle sue componenti $u_i(t) = \langle e_i, u(t) \rangle$; anch'esse sono assolutamente continue e dunque sappiamo che possiamo applicare il teorema fondamentale del calcolo integrale per funzioni da \mathbb{R} in \mathbb{R} : possiamo quindi considerare le derivate delle proiezioni $u'_i(t)$, che esistono quasi ovunque, per le quali sappiamo valere il teorema fondamentale del calcolo integrale. Per l'assoluta continuità di $u(t)$, esiste una funzione g tale che

$$\|u(b) - u(a)\| \leq \int_a^b g(t) dt \quad \text{per ogni } 0 \leq s < t \leq 1$$

Prendiamo ora $t_0 \in [0, 1]$ un punto che sia di Lebesgue sia per $g(t)$ che per tutti gli $u'_i(t)$ (questa condizione vale Lebesgue quasi ovunque). Allora si ha

$$\limsup_{h \rightarrow 0} \left\| \frac{u(t_0+h) - u(t_0)}{h} \right\| \leq g(t_0),$$

e quindi ogni sottosuccessione del rapporto incrementale ha un limite debole, che ha norma minore od uguale a $g(t_0)$; chiamiamo w uno di questi limiti deboli. Allora si ha necessariamente

$\langle w, e_i \rangle = u'_i(t_0)$, e quindi tutta la successione converge debolmente al vettore

$$u'(t_0) := \sum_{i=1}^{+\infty} e_i u'_i(t_0),$$

il cui modulo è minore o uguale a $g(t_0)$. Quindi vale $\|u'(t)\| \leq g(t)$ quasi ovunque, ma allora u' è integrabile e si ha:

$$\left\langle \int_a^b u'(t) dt, e_i \right\rangle = \int_a^b u'_i(t) dt = \langle u(b) - u(a), e_i \rangle$$

e quindi vale il teorema fondamentale del calcolo integrale:

$$\int_a^b u'(t) dt = u(b) - u(a).$$

In particolare otteniamo $\int_a^b \|u'\| \geq \|u(b) - u(a)\|$, e quindi possiamo prendere $g(t) = \|u'(t)\|$. Ora, considerando t un punto di Lebesgue di $\|u'\|$ e di differenziabilità debole per u , otteniamo:

$$\|u'(t)\| \leq \liminf \left\| \frac{u(t+h) - u(t)}{h} \right\| \leq \lim \frac{\int_t^{t+h} \|u'(s)\| ds}{h} = \|u'(t)\|.$$

In particolare $\|u'(t)\| = \lim \left\| \frac{u(t+h) - u(t)}{h} \right\|$ e dunque il limite del rapporto incrementale è anche un limite forte e dunque u è derivabile fortemente quasi ovunque con derivata u' . □

Appendice C

Inviluppo convesso di una funzione

Prima di parlare di inviluppo convesso di una funzione, richiamiamo la definizione di trasformata di Legendre:

Definizione C.1 (Trasformata di Legendre). *Data $F : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$, la trasformata di Legendre di F si definisce come*

$$G(y) = F^*(y) = \sup_{x \in \mathbb{R}} \{xy - F(x)\}.$$

La trasformata di Legendre in realtà si può definire per funzioni reali da uno spazio di Banach qualsiasi, ma per i nostri scopi ci restringiamo al caso reale. Molte proprietà della trasformata di Legendre sono interessanti ma noi ci limiteremo a tre particolari proprietà:

- $F(x) + G(y) = yx$ se e solo se $y \in \partial F(x)$;
- $F^{**} = F$ se e solo se F è convessa e s.c.i. ;
- F^{**} è l'inviluppo convesso di F , cioè la massima funzione convessa s.c.i. più piccola di F .

Ricordando che F^* risulta essere sempre una funzione convessa poiché estremo superiore di rette, usando le prime due proprietà, troviamo che se F è convessa s.c.i. allora

$$y \in \partial F(x) \iff x \in \partial G(y)$$

e quindi la derivata di F risulta essere la pseudoinversa della derivata di F^* , e viceversa.

Dimostriamo ora alcune proprietà dell'inviluppo convesso, nel caso in cui F è una funzione definita solo in $[0, 1]$; per definire la trasformata di Legendre immaginiamo che $F(x) = \infty$ se $x \notin [0, 1]$. In questo caso si vede facilmente che anche $F^{**}(x) = \infty$ se $x \notin [0, 1]$, e quindi ha senso considerare anche F^{**} definita solo su $[0, 1]$.

Lemma C.2. *Data F una funzione continua da $[0, 1]$ a valori reali che si annulla in 0. Sia F^{**} il suo inviluppo convesso; allora valgono*

- $F(0) = F^{**}(0)$ e $F(1) = F^{**}(1)$;
- L'insieme $A = \{F(x) > F^{**}(x)\}$ è un aperto e se $(a, b) \subseteq A$ allora $F^{**}|_{(a,b)}$ è una funzione lineare.

Dimostrazione. Innanzitutto ricordiamo le seguenti caratterizzazioni di F^{**} :

$$\begin{aligned} F^{**}(x) &= \sup_{G \text{ convessa}} \{G(x) \mid G(y) \leq F(y) \ \forall y \in [0, 1]\} \\ &= \sup\{ax + b \mid ay + b \leq F(y) \ \forall y \in [0, 1]\} \end{aligned}$$

Il fatto che F e F^{**} coincidano agli estremi segue semplicemente dalla continuità: facciamo vedere in un estremo, ad esempio 0, l'altro caso sarà uguale. Notiamo che ci basta far vedere che $F^{**}(0) \geq F(0)$ poiché l'altra disuguaglianza è ovvia dalla definizione. Chiamiamo m il minimo di F e supponiamo all'inizio che il minimo di $F(0) = m$; allora la retta $y = m$ è nel sottografico $F(x)$ e quindi $F^{**}(0) \geq m = F(0)$. Se invece $F(0)$ non è il minimo allora prendiamo un qualunque $\varepsilon > 0$ tale che $F(0) > m + \varepsilon$ e di seguito prendiamo $\delta > 0$, dato dalla continuità, tale che $F(0) - F(y) > \varepsilon$ ogni volta che $y < \delta$.

Consideriamo ora la retta:

$$y = r(x) = (F(0) - \varepsilon) \left(1 - \frac{x}{\delta}\right) + m \frac{x}{\delta}.$$

Notiamo che $r(0) = F(0) - \varepsilon > m = F(\delta)$ e dunque la retta risulta essere decrescente. In particolare se $y \leq \delta$ abbiamo $r(y) \leq F(0) - \varepsilon < F(y)$ e invece se $y \geq \delta$ abbiamo $r(y) \leq m \leq F(y)$, poiché m è il minimo di F .

Dunque $r(y) \leq F(y)$ per ogni $y \in [0, 1]$ e allora la retta $r(y)$ risulta essere una retta ammissibile e dunque in particolare $F^{**}(0) \geq F(0) - \varepsilon$, e per l'arbitrarietà di ε concludo.

Da questa costruzione inoltre si può dedurre che $F^{**}(x)$ risulta essere continua agli estremi e dunque, essendo convessa, continua ovunque; quindi l'insieme A è ovviamente aperto. Per il secondo punto ci basta dimostrare che F^{**} è lineare negli intervalli chiusi inclusi in A per poi passare al limite.

Consideriamo dunque un intervallo $[a, b] \subseteq A$; sia L la funzione lineare tale che $L(a) = F^{**}(a)$ e $L(b) = F^{**}(b)$. Per la convessità di F^{**} ho che $F^{**} \leq L$ in $[a, b]$ mentre $F^{**} \geq L$ nel complementare. Consideriamo $\widehat{F} = F^{**} \wedge L \geq F^{**}$ che risulta essere convessa e coincide con L in $[a, b]$ e con F^{**} nel complementare.

Sapendo ora che $F - F^{**}$ è una funzione strettamente positiva, essa assumerà un minimo in $[a, b]$, chiamiamolo $m^* > 0$. Chiamiamo invece \widehat{m} il massimo della funzione $\widehat{F} - F^{**}$, realizzato sicuramente in $[a, b]$, poiché altrove le funzioni coincidono. Consideriamo ora una combinazione convessa

$$\widehat{F}_t = t\widehat{F} + (1-t)F^{**} = F^{**} + t(\widehat{F} - F^{**})$$

che risulterà anch'essa una funzione convessa, tale che $\widehat{F}_t \geq F^{**}$. Prendendo ora t_0 tale che $\widehat{m}t_0 \leq m^*$ (osserviamo che t_0 risulta positivo in quanto $m^* > 0$ otteniamo che sicuramente $\widehat{F}_{t_0}(x) \leq F(x)$: infatti nell'intervallo $[a, b]$ si ha

$$\widehat{F}_{t_0} = F^{**} + t_0(\widehat{F} - F^{**}) \leq F^{**} + t_0\widehat{m} \leq F^{**} + m^* \leq F$$

mentre fuori da $[a, b]$ la funzione \widehat{F}_{t_0} coincide con F^{**} che è minore o uguale a F . Quindi, per definizione di involucro convesso, si ha $F^{**} \geq \widehat{F}_{t_0}$ che per costruzione è maggiore di F^{**} da cui $F^{**} = \widehat{F}_{t_0}$ e allora $F^{**} = \widehat{F}$, da cui F^{**} coincide con la funzione lineare L nell'intervallo $[a, b]$.

□

Appendice D

Equazione di Hamilton-Jacobi

Vogliamo, in questo capitolo, introdurre le equazioni di Hamilton-Jacobi e le relative leggi di conservazione scalare. Importante nello studio della fisica, legato alla meccanica Hamiltoniana, è l'equazione

$$\partial_t u(t, x) + H(D_x u(t, x), t, x) = 0$$

detta equazione di Hamilton-Jacobi; la funzione H si dice Hamiltoniana del sistema. In particolare consideriamo qui Hamiltoniane dipendenti solo dalla derivata spaziale, e dunque indipendenti da posizione e tempo, e per dimensione spaziale solo $n = 1$. Il problema che dunque consideriamo è del tipo

$$\begin{cases} \partial_t u(t, x) + H(\partial_x u(t, x)) = 0 & \forall (t, x) \in (0, T) \times \mathbb{R} \\ u(0, x) = u_0(x). \end{cases} \quad (\text{HJ})$$

D.1 Soluzioni viscosi di Hamilton-Jacobi

La nozione di soluzione viscosa nasce similmente alla nozione di soluzione entropica per le leggi di conservazione scalare: non esistono soluzioni lisce per tutti i tempi alle equazioni di Hamilton-Jacobi, ma di soluzioni deboli ce ne sono infinite e dunque bisogna trovare una nozione di soluzione speciale per avere l'unicità.

La definizione tecnica è quella mediante sottosoluzioni e soprassoluzioni (si può trovare in [Eva10]), tuttavia una possibile definizione alternativa di soluzione viscosa è come limite puntuale delle uniche soluzioni u_ε del problema parabolico

$$\begin{cases} \partial_t u(t, x) + H(\partial_x u(t, x)) = \varepsilon \partial_{xx} u(t, x) \\ u(0, x) = u_0(x). \end{cases}$$

Questa caratterizzazione permette di collegare questa soluzione viscosa alla soluzione entropica della legge di conservazione scalare associata

$$\partial_t v(t, x) + \partial_x (H(v(t, x))) = 0,$$

infatti considerando le funzioni $v_\varepsilon = \partial_x u_\varepsilon$, esse soddisfano, derivando spazialmente la (HJ),

$$\begin{cases} \partial_t v(t, x) + \partial_x(H(v(t, x))) = \varepsilon \partial_{xx} v(t, x) \\ v(0, x) = \partial_x u_0(x). \end{cases}$$

e dunque $v_\varepsilon \rightarrow v$, soluzione entropica della legge di conservazione scalare. Se la convergenza delle u_ε è abbastanza regolare, si trova che $\partial_x v = u$ e quindi u è soluzione viscosa dell'equazione di Hamilton-Jacobi se e solo se $\partial_x u$ è soluzione della legge di conservazione scalare associata.

D.2 Prima e seconda formula di Hopf

C'è una classe di soluzioni di (HJ) molto importante, che è quella per cui $\partial_x u$ è costante. Prendendo $\partial_x u \equiv b$, troviamo che $\partial_t u \equiv -H(b)$, e quindi la soluzione

$$u(t, x) = bx - H(b)t + c.$$

Hopf propone in [Hop65] di considerare, tenendo conto anche della condizione iniziale, i seguenti involuppi di soluzioni lineari:

$$u_1(x, t) = \inf_{y \in \mathbb{R}} \sup_{b \in \mathbb{R}} \{bx - H(b)t + u_0(y) - by\}$$

$$u_2(x, t) = \sup_{b \in \mathbb{R}} \inf_{y \in \mathbb{R}} \{bx - H(b)t + u_0(y) - by\}$$

chiamate rispettivamente prima e seconda formula di Hopf.

Queste funzioni sono importanti poiché con opportune ipotesi sull'Hamiltoniana o sulla condizione iniziale, sono le uniche soluzioni *viscose* delle equazioni di Hamilton-Jacobi. La più famosa è la prima formula di Hopf, detta anche di Hopf-Lax, che rappresenta la soluzione viscosa ad esempio nel caso in cui H è una funzione convessa, e, chiamando $L = H^*$ la trasformata di Legendre dell'Hamiltoniana, si può scrivere come

$$u_1(x, t) = \inf_{y \in \mathbb{R}} \left\{ u_0(y) + tL\left(\frac{x-y}{t}\right) \right\}.$$

Tuttavia la soluzione che ci interessa è quella data dalla seconda formula di Hopf, che è soluzione viscosa di (HJ) nel caso in cui u_0 è funzione continua e convessa e H è continua (si veda [LR86, BE84]).

Essa si può scrivere, più comodamente, come:

$$\begin{aligned} u_2(x, t) &= \sup_{b \in \mathbb{R}} \left\{ bx - H(b)t - \sup_{y \in \mathbb{R}} \{by - u_0(y)\} \right\} \\ &= \sup_{b \in \mathbb{R}} \{bx - tH(b) - u_0^*(b)\} \\ &= (u_0^* + tH)^*(x). \end{aligned}$$

Quest'ultima forma è proprio quella che abbiamo usato nell'ultima sezione dell'ultimo capitolo, riguardante le connessioni tra il sistema (SPS) e le leggi di conservazione scalari.

Bibliografia

- [AGS08] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [BE84] M. Bardi and L. C. Evans. On Hopf's formulas for solutions of Hamilton-Jacobi equations. *Nonlinear Anal.*, 8(11):1373–1381, 1984.
- [BG98] Yann Brenier and Emmanuel Grenier. Sticky particles and scalar conservation laws. *SIAM J. Numer. Anal.*, 35(6):2317–2328 (electronic), 1998.
- [BJ95] François Bouchut and François James. Équations de transport unidimensionnelles à coefficients discontinus. *C. R. Acad. Sci. Paris Sér. I Math.*, 320(9):1097–1102, 1995.
- [Bou94] F. Bouchut. On zero pressure gas dynamics. In *Advances in kinetic theory and computing*, volume 22 of *Ser. Adv. Math. Appl. Sci.*, pages 171–190. World Sci. Publ., River Edge, NJ, 1994.
- [Bou00] Laurent Boudin. A solution with bounded expansion rate to the model of viscous pressureless gases. *SIAM J. Math. Anal.*, 32(1):172–193 (electronic), 2000.
- [Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
- [GNT09] W. Gangbo, T. Nguyen, and A. Tudorascu. Euler-Poisson systems as action-minimizing paths in the Wasserstein space. *Arch. Ration. Mech. Anal.*, 192(3):419–452, 2009.
- [Gre95] Emmanuel Grenier. Existence globale pour le système des gaz sans pression. *C. R. Acad. Sci. Paris Sér. I Math.*, 321(2):171–174, 1995.
- [Hop65] Eberhard Hopf. Generalized solutions of non-linear equations of first order. *J. Math. Mech.*, 14:951–973, 1965.
- [HW01] Feimin Huang and Zhen Wang. Well posedness for pressureless flow. *Comm. Math. Phys.*, 222(1):117–146, 2001.

- [LR86] P.-L. Lions and J.-C. Rochet. Hopf formula and multitime Hamilton-Jacobi equations. *Proc. Amer. Math. Soc.*, 96(1):79–84, 1986.
- [Mou08] Octave Moutsinga. Convex hulls, sticky particle dynamics and pressure-less gas system. *Ann. Math. Blaise Pascal*, 15(1):57–80, 2008.
- [MP94] A. Martin and J. Piasecki. One dimensional ballistic aggregation: Rigorous long-time estimates. *J. Stat. Phys*, 76, 1994.
- [NS09] Luca Natile and Giuseppe Savaré. A Wasserstein approach to the one-dimensional sticky particle system. *SIAM J. Math. Anal.*, 41(4):1340–1365, 2009.
- [NT08] Truyen Nguyen and Adrian Tudorascu. Pressureless Euler/Euler-Poisson systems via adhesion dynamics and scalar conservation laws. *SIAM J. Math. Anal.*, 40(2):754–775, 2008.
- [PR97] F. Poupaud and M. Rascle. Measure solutions to the linear multi-dimensional transport equation with non-smooth coefficients. *Comm. Partial Differential Equations*, 22(1-2):337–358, 1997.
- [San10] F. Santambrogio. Gradient flows in wasserstein spaces and applications to crowd movement, 2010. Available on line.
- [Sav93] G. Savaré. Approximation and regularity of evolution variational inequalities. *Rend. Accad. Naz. Sci. XL Mem. Mat. (5)*, 17:83–111, 1993.
- [Sev01] Michael Sever. An existence theorem in the large for zero-pressure gas dynamics. *Differential Integral Equations*, 14(9):1077–1092, 2001.
- [Sob97] A. N. Sobolevskii. The small viscosity method for a one-dimensional system of equations of gas dynamic type without pressure. *Dokl. Akad. Nauk*, 356(3):310–312, 1997.
- [WD97] Zhen Wang and Xiaqi Ding. Uniqueness of generalized solution for the Cauchy problem of transportation equations. *Acta Math. Sci. (English Ed.)*, 17(3):341–352, 1997.
- [Wol07] Gershon Wolansky. Dynamics of a system of sticking particles of finite size on the line. *Nonlinearity*, 20(9):2175–2189, 2007.
- [WRS96] E. Weinan, Yu. G. Rykov, and Ya. G. Sinai. Generalized variational principles, global weak solutions and behavior with random initial data for systems of conservation laws arising in adhesion particle dynamics. *Comm. Math. Phys.*, 177(2):349–380, 1996.
- [Zel70] Y Zeldovich. Gravitational instability: An approximate theory for large density perturbations. *Astron. & Astrophys.*, 5:84–89, 1970.