Appunti per il corso di: Equazioni di Evoluzione, a.a. 2015/16

Marco Veneroni

June 9, 2016

Contents

Premessa: il senso di questi appunti è solo quello di fornire un appoggio e un riferimento scritto a chi ha seguito il corso. Il materiale che segue è principalmente una traduzione di alcuni campioni estratti dai riferimenti in bibliografia, senza nessuna pretesa di completezza, né di precisione o di originalità. Alcuni dei risultati non sono stati presentati a lezione, *non fanno quindi parte del programma d'esame*, ma sono stati comunque inclusi nelle note per completezza.

1 Introduzione

1.1 Flussi gradiente in spazi di Hilbert

Sia H uno spazio di Hilbert, $F : H \to \mathbb{R} \cup \{+\infty\}$ un funzionale, il cui dominio è $D(F) := \{u \in H : F(u) < +\infty\}$. Chiamiamo *flusso gradiente di F in H* una funzione $u : [0, +\infty) \to \mathbb{R}$ che risolva, per ogni $t \geq 0$ l'equazione

$$\frac{d}{dt}u(t) = -\nabla F(u(t)) \qquad \text{in } H. \tag{1.1}$$

Per dare un senso all'equazione, è necessario innanzitutto definire il simbolo ∇F. Vi sono varie possibilità, a seconda del contesto e della regolarità di F. Per esempio, diciamo che F è differenziabile nel senso di Fréchet nel punto $u \in H$ se esiste un operatore lineare $A : D(F) \to \mathbb{R}$ per cui

$$\lim_{h\to 0} \frac{|F(u+h) - F(u) - A(h)|}{\|h\|_H} = 0.$$

In questo caso si pone $\nabla F(u) = A$.

Un'altra possibilità, più debole: diciamo che F ammette derivata nel senso di Gâteaux nel punto $u \in H$, in direzione $h \in H$, se esiste ed è finito

$$\nabla F(u; h) := \lim_{\tau\to 0} \frac{F(u + \tau h) - F(u)}{\tau}. \tag{1.2}$$

Alcune notazioni alternative:

$$\nabla F(u)[h], \qquad dF_u(h), \qquad \frac{d}{d\tau}F(u + \tau h)\Big|_{\tau=0}.$$

L'idea è che il differenziale di una funzione reale si generalizza con la derivata di Fréchet, mentre quella di Gâteaux generalizza la derivata direzionale. Come in analisi 1, se una funzione è differenziabile, allora ha tutte le derivate direzionali e i due concetti coincidono (nel senso: $\nabla f(x) \cdot \nu = \partial_\nu f(x)$), mentre il viceversa è falso. L'applicazione $h \mapsto \nabla F(u; h)$ definita in (1.2) non è necessariamente lineare, né continua. <u>Se</u> (1.2) definisce un'applicazione lineare e continua su H, <u>allora</u> la chiamiamo *differenziale di Gâteaux*. Questo concetto sarà utile nella prossima Sezione.

Le definizioni si trovano, corredate di ampie spiegazioni, esempi ed esercizi, in [1, Section 1].

Esempio 1.1

Sia $\Omega \subset \mathbb{R}^n$ un aperto regolare e si prenda $H = L^2(\Omega)$, $D(F) = H^1_0(\Omega)$ e

$$F(u) := \frac{1}{2}\int_\Omega |\nabla u(x)|^2 \, dx.$$

Sia $\phi \in L^2(\Omega)$, calcoliamo la derivata di Gâteaux di F in un punto u, in direzione ϕ:

$$\begin{aligned}\nabla F(u;\phi) &= \frac{d}{d\tau}\left\{\frac{1}{2}\int_\Omega |\nabla u + \tau\nabla\phi|^2\,dx\right\}\Bigg|_{\tau=0}\\ &= \frac{1}{2}\int_\Omega 2(\nabla u + \tau\nabla\phi)\cdot\nabla\phi\,dx\Bigg|_{\tau=0}\\ &= \int_\Omega \nabla u\cdot\nabla\phi\,dx\\ &= -\int_\Omega (\Delta u)\phi\,dx.\end{aligned}$$

Ora, $\nabla F(u)$ non è un elemento di H, ma un funzionale lineare su H, ovvero un elemento dello spazio duale H':

$$\phi \mapsto \nabla F(u;\phi) = -\int_\Omega (\Delta u)\phi\,dx.$$

Per poter vedere l'equazione (1.1) in H, (assumendo $u \in H^1(\Omega)\cap H^2(\Omega)$), sfruttiamo la struttura dello spazio di Hilbert, e in particolare il Teorema di Riesz, grazie al quale esiste un unico elemento $v \in H$ tale che $(v,\phi) = \nabla F(u;\phi)$ per ogni $\phi \in H$, ovvero, identifichiamo il funzionale $\nabla F(u)$ con la funzione $-\Delta u$, giungendo così all'equazione del calore:

$$\frac{d}{dt}u(t) = \Delta u(t).$$

L'identificazione fatta non è un dettaglio tecnico, ma il nodo cruciale della teoria. Vedremo che diverse identificazioni possono portare ad equazioni diverse.

□ □ □ □ □

1.2 Un punto di vista geometrico

Come viene descritto in [11, Sezione 1] il flusso gradiente può essere considerato anche in un contesto di geometria differenziale: sia M una varietà differenziabile, T_uM lo spazio tangente nel punto u e indichiamo con $g_u : T_uM \times T_uM \to \mathbb{R}$ la metrica. Una curva $u : \mathbb{R} \to M$ è un flusso gradiente di $F : M \to \mathbb{R}$ se

$$g_{u(t)}(u'(t), s) = -dF_{u(t)}(s) \qquad \forall\, s \in T_{u(t)}M,\ \forall\, t \in \mathbb{R} \tag{1.3}$$

dove dF_u indica il differenziale di F in u. Mostriamo ora, formalmente, che due diverse scelte di F e di g possono portare all'equazione del calore. Saremo volutamente imprecisi riguardo alla struttura differenziale della varietà M, in questo momento si cerca solo di dare un'idea, le dimostrazioni rigorose non mancheranno nelle sezioni successive.

Siano

$$M := \left\{\text{funzioni non negative } \rho : \mathbb{R}^n \to \mathbb{R}\ :\ \int \rho = 1\right\}$$

e

$$T_\rho M := \left\{\text{funzioni } s : \mathbb{R}^n \to \mathbb{R}\ :\ \int s = 0\right\}.$$

Il tensore metrico viene definito attraverso un'identificazione dello spazio tangente con un insieme di funzioni $p : \mathbb{R}^n \to \mathbb{R}$ (o precisamente, di classi di funzioni con $p_1 \sim p_2$ se $p_1 - p_2 = cost.$). L'identificazione che coincide con quella fatta nell'esempio 1.1 è

$$(a) \begin{cases} -\Delta p = s, \\ g_\rho(s_1, s_2) = \displaystyle\int \nabla p_1 \cdot \nabla p_2. \end{cases} \tag{1.4}$$

La nuova identificazione, che andiamo ad introdurre è,

$$(b) \begin{cases} -\mathrm{div}(\rho \nabla p) = s, \\ g_\rho(s_1, s_2) = \displaystyle\int (\nabla p_1 \cdot \nabla p_2)\rho. \end{cases} \tag{1.5}$$

Notiamo che integrando per parti la seconda riga di (a) e sostituendo $-\Delta p_1 = s_1$, si ottiene

$$g_\rho(s_1, s_2) = \int \nabla p_1 \cdot \nabla p_2 = -\int \mathrm{div}(\nabla p_1) p_2 = \int s_1 p_2. \tag{1.6}$$

Esattamente lo stesso risultato si ottiene integrando per parti la seconda riga di (b) e sostituendo $-\mathrm{div}(\rho \nabla p_1) = s_1$. Il funzionale per (a) è

$$E(\rho) = \frac{1}{2} \int \rho^2,$$

quello per (b) è

$$E(\rho) = \int \rho \log \rho.$$

Date $\rho \in M$, $s \in T_\rho M$, calcoliamo $dE_\rho(s)$. Facciamo, una volta per tutte, il conto per un'energia generale del tipo $E(\rho) = \int e(\rho)$, con $e \in C^2(\mathbb{R})$. Per definizione di differenziale, si tratta di prendere una curva $\gamma : (-\varepsilon, \varepsilon) \to M$ tale che $\gamma(0) = \rho$, $\gamma'(0) = s$ (per esempio, $\gamma(\tau) = \rho + \tau s$ - con qualche condizione, in modo che $\rho + \tau s \geq 0$) e calcolare

$$dE_\rho(s) = \frac{d}{d\tau} E(\gamma(\tau))\Big|_{\tau=0} = \left\{ \int \frac{d}{d\tau} e(\rho + \tau s) \right\} \Big|_{\tau=0} = \int e'(\rho) s. \tag{1.7}$$

Mostriamo dunque che in entrambi i casi il flusso gradiente (1.3) si traduce nell'equazione del calore. Abbiamo già osservato che

$$g_\rho(\partial_t \rho, s) \overset{(1.6)}{=} \int \partial_t \rho \, p,$$

calcoliamo quindi $dE_\rho(s)$. Nel caso (a), $e'(\rho) = \rho$,

$$dE_\rho(s) = \int \rho \, s \overset{(1.4)}{=} -\int \rho \, \Delta p = -\int (\Delta \rho) p$$

e (1.3) diventa

$$\int (\partial_t \rho - \Delta \rho) p = 0. \tag{1.8}$$

Nel caso (b), $e'(\rho) = 1 + \log\rho$,

$$
\begin{aligned}
dE_\rho(s) = \int (1+\log\rho)\, s \overset{(1.5)}{=} -\int (1+\log\rho)\,\mathrm{div}(\rho\nabla p) &= \int (\nabla(1+\log\rho)\cdot\nabla p)\rho \\
&= \int \left(\frac{\nabla\rho}{\rho}\cdot\nabla p\right)\rho \\
&= \int \nabla\rho\cdot\nabla p \\
&= -\int (\Delta\rho)p
\end{aligned}
$$

e si ottiene ancora (1.8).

Esercizio 1.2
Sia $m \in \mathbb{N}, m \geq 2$. Usando l'identificazione

$$
\text{(b)} \begin{cases} -\mathrm{div}(\rho\nabla p) = s, \\ g_\rho(s_1, s_2) = \int (\nabla p_1 \cdot \nabla p_2)\rho. \end{cases}
$$

e l'energia

$$
E(\rho) = \frac{1}{m-1}\int \rho^m,
$$

dimostrare che il flusso gradiente (1.3) diventa l'equazione dei mezzi porosi:

$$
\partial_t\rho = \Delta(\rho^m). \tag{1.9}
$$

□ □ □ □ □

1.3 Un punto di vista fisico

Seguendo la descrizione di [11, Sezione 2], diamo una breve derivazione fisica dell'equazione dei mezzi porosi (1.9). Nel caso particolare $m = 1$, si otterrà naturalmente l'equazione del calore.

La funzione ρ rappresenta la densità di massa di un gas in un mezzo poroso. Come dominio di ρ e di tutte le funzioni che introdurremo, scegliamo $\mathbb{R}^n$. La prima ipotesi è la conservazione di massa, espressa dall'equazione di continuità

$$
\partial_t\rho + \mathrm{div}(\rho\mathbf{u}) = 0,
$$

dove $\mathbf{u}$ è il vettore velocità (media) del gas. La seconda ipotesi è la legge di Darcy, che in un mezzo *isotropo* e *omogeneo* è

$$
\mathbf{u} = -\nabla p, \tag{1.10}
$$

dove p è la pressione del gas. (Nota: pressione e velocità sono grandezze diverse, per non sommare mele e pere nell'equazione (1.10) stiamo implicitamente assumendo, come solito, di aver opportunamente adimensionalizzato l'equazione.) La terza ipotesi viene dalla termodinamica:

$$p = \frac{\delta E}{\delta \rho}, \tag{1.11}$$

dove E rappresenta l'energia libera e $\delta E/\delta\rho$ la sua derivata funzionale rispetto a ρ. Nel caso di un'energia in forma integrale

$$E(\rho) = \int e(\rho),$$

dove la funzione $z \mapsto e(z)$ descrive la dipendenza della densità di energia da ρ, (1.11) diventa

$$p = e'(\rho). \tag{1.12}$$

(È equivalente a calcolare, come prima in (1.7), il differenziale di E in ρ e quindi identificare l'operatore $s \mapsto dE_\rho(s)$ con la funzione $e'(\rho)$.) Quindi, (1.3), (1.10) e (1.12), si combinano in

$$\partial_t \rho - \operatorname{div}(\rho \nabla\{e'(\rho)\}) = 0. \tag{1.13}$$

Dato che $\rho\nabla\{e'(\rho)\} = \rho e''(\rho)\nabla\rho$, introducendo la funzione $z \mapsto \pi(z)$, che rappresenta la pressione osmotica, e la cui relazione con e è data da

$$\pi(z) = ze'(z) - e(z),$$

si calcola immediatamente che

$$\pi'(z) = ze''(z), \qquad \nabla\{\pi(\rho)\} = \pi'(\rho)\nabla(\rho) = \rho e''(\rho)\nabla\rho,$$

e quindi (1.13) si riscrive più concisamente come

$$\partial_t \rho - \Delta(\pi(\rho)) = 0. \tag{1.14}$$

Ora, (1.14) coincide con l'equazione dei mezzi porosi (1.9) solo se $\pi(z) = z^m$, ovvero solo se

$$e(z) = \begin{cases} \dfrac{1}{m-1} z^m & \text{per } m \neq 1, \\[2ex] z \log z & \text{per } m = 1. \end{cases} \tag{1.15}$$

Troviamo quindi un significato fisico per l'energia E *solo nella nuova formulazione*. Inoltre lungo una soluzione $\rho(t)$, si ha

$$\frac{d}{dt}E(\rho) = dE_\rho(\partial_t\rho) = -g_\rho(\partial_t\rho, \partial_t\rho).$$

Il termine a sinistra rappresenta la velocità di variazione dell'energia libera, il termine a destra la velocità di dissipazione dell'energia cinetica attraverso l'attrito. Nella dinamica le due quantità sono uguali ed essendo $-g_\rho(s,s) \leq 0$, decrescono nel tempo.

1.4 Il risultato di Jordan, Kinderlehrer e Otto

Questa lunga premessa ha lo scopo di suggerire che un'ampia classe di equazioni alle derivate parziali può essere descritta come flusso gradiente di un funzionale energia, attraverso una scelta appropriata della struttura metrica che si usa per collegare il gradiente del funzionale alla derivata temporale. Richard Jordan, David Kinderlehrer e Felix Otto, nel loro fondamentale lavoro "The variational formulation of the Fokker-Planck equation" [10], presentano questa nuova idea e ne danno una dimostrazione rigorosa nel caso dell'equazione di Fokker-Planck.

Precisamente consideriamo l'equazione

$$\frac{\partial \rho}{\partial t} = \operatorname{div}(\rho \nabla \Psi) + \frac{1}{\beta}\Delta\rho, \qquad \rho(x,0) = \rho^0(x), \tag{1.16}$$

dove $\Psi : \mathbb{R}^n \to [0,+\infty)$ è un potenziale regolare, $\beta > 0$ è una costante fissata e $\rho^0(x)$ è una densità di probabilità su $\mathbb{R}^n$. In [10] viene dimostrato che (1.16) è il flusso gradiente del funzionale

$$F(\rho) = E(\rho) + \frac{1}{\beta} S(\rho), \tag{1.17}$$

dove

$$E(\rho) := \int_{\mathbb{R}^n} \Psi(x)\rho(x)\, dx \tag{1.18}$$

gioca il ruolo di un funzionale d'energia, e

$$S(\rho) := \int_{\mathbb{R}^n} \rho(x)\log\rho(x)\, dx \tag{1.19}$$

è il funzionale entropia di Gibbs-Boltzmann, con il segno opposto. La struttura è quella dello spazio metrico $(\mathscr{P}_2(\mathbb{R}^n), W_2)$. Definiamo

$$\mathscr{P}_2(\mathbb{R}^n) = \{\text{misure di probabilità di Borel } \mu \text{ su } \mathbb{R}^n, \text{ tali che } M(\rho) < +\infty\}, \tag{1.20}$$

dove

$$M(\rho) := \int_{\mathbb{R}^n} |x|^2\, d\mu(x)$$

è il momento secondo di μ, e W_2 è la distanza di Kantorovich-Rubinstein-Wasserstein su $\mathscr{P}_2(\mathbb{R}^n)$, definita da

$$\begin{aligned} W_2(\mu_1,\mu_2)^2 = \inf\Big\{ &\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2\, d\gamma(x,y) \;:\; \gamma \in \mathscr{P}(\mathbb{R}^n\times\mathbb{R}^n), \\ &\gamma(A\times\mathbb{R}^n) = \mu_1(A),\ \gamma(\mathbb{R}^n\times A) = \mu_2(A),\ \text{per ogni Boreliano } A \subset \mathbb{R}^n \Big\}. \end{aligned} \tag{1.21}$$

Riprenderemo nel dettaglio la particolare struttura di questo spazio in una sezione successiva.

La dimostrazione di [10] si basa sul seguente schema iterativo discreto:

(1) L'intervallo di tempo $[0,+\infty)$ viene diviso in intervalli di lunghezza $h > 0$,

(2) al tempo $t_0 = 0$ si usa il dato iniziale, definendo $\rho^{(0)} = \rho^0$,

(3) data la soluzione discreta $\rho^{(k-1)}$ al tempo $t_{k-1} = h(k-1)$, la soluzione $\rho^{(k)}$ al tempo $t_k = hk$ si trova come soluzione del problema variazionale:

$$\rho^{(k)} = \arg\min_{\rho \in K} \Phi(\rho; \rho^{(k-1)}, h), \tag{1.22}$$

dove l'insieme K e il funzionale Φ sono definiti da:

$$K = \left\{ \rho : \mathbb{R}^n \to [0, +\infty), \text{ misurabile } : \int_{\mathbb{R}^n} \rho(x)\, dx = 1, \quad M(\rho) < +\infty \right\},$$
$$\Phi(\rho, \rho^{(k-1)}, h) := \frac{1}{2} W_2\big(\rho, \rho^{(k-1)}\big)^2 + hF(\rho).$$

Lo schema proposto è la versione metrica in $(\mathscr{P}_2(\mathbb{R}^n), W_2)$ del metodo di Eulero implicito. Una volta ottenuta la successione di soluzioni approssimate $\rho^{(k)}$ a passo h, si dimostra che l'interpolazione costante a tratti $\rho_h(t)$ converge alla soluzione di (1.16), per $h \to 0$, rispetto alla topologia debole di $L^1(\mathbb{R}^n)$.

Esempio 1.3
Applichiamo lo stesso schema all'equazione del calore. Come nell'Esempio 1.1, sia $\Omega \subset \mathbb{R}^n$ un aperto regolare e si prenda $H = L^2(\Omega)$, $D(F) = H_0^1(\Omega) \cap H^2(\Omega)$ e

$$F(u) := \frac{1}{2} \int_\Omega |\nabla u(x)|^2\, dx.$$

Fissato $h > 0$ e un dato iniziale $u^0 \in D(F)$, lo schema (1.22) si applica cercando il punto di minimo di

$$\Phi(u; u^{(k-1)}, h) := \frac{1}{2} \big\| u - u^{(k-1)} \big\|^2_{L^2(\Omega)} + hF(u).$$

su $K = D(F)$. Se $u^{(k)}$ è il punto di minimo, allora[1] $\nabla\Phi(u^{(k)}) = 0$. Calcoliamo quindi

$$0 = \nabla\Phi(u^{(k)}) = u^{(k)} - u^{(k-1)} - h\Delta u^{(k)},$$

ovvero $u^{(k)}$ è la soluzione discreta al passo k del metodo di Eulero implicito per l'equazione del calore

$$\frac{u^{(k)} - u^{(k-1)}}{h} = \Delta u^{(k)}.$$

□ □ □ □ □

2 Flussi gradiente in spazi di Hilbert

Questa parte del corso segue precisamente lo svolgimento del Capitolo 9.6 "Gradient flows" di [9]. Consiglio quindi di studiare l'argomento direttamente sul testo originale. Riporto qui solo alcune osservazioni ed esempi.

[1] Questo dobbiamo ancora dimostrarlo, l'idea è che, come per funzioni reali, i punti di estremo regolari e interni sono stazionari. Si veda, e.g., [9, Theorem 1 (iii), pag. 524]

Proposizione 2.4. *Sia H uno spazio di Hilbert e sia $F : H \to \mathbb{R} \cup \{+\infty\}$ un funzionale convesso. Se F ha il differenziale di Gâteaux $\nabla F(u)$ in un punto $u \in H$, allora il sottodifferenziale $\partial F(u)$ è ridotto ad un solo elemento e*

$$\nabla F(u) = \partial F(u).$$

Esempio 2.5

1. $f(x) = |x|, \ x \in \mathbb{R}$
$$\Rightarrow \partial f(x) = \begin{cases} -1 & \text{per } x < 0, \\ [-1, 1] & \text{per } x = 0, \\ +1 & \text{per } x > 0. \end{cases}$$

2. $f(x) = |x|, \ x \in \mathbb{R}^n$
$$\Rightarrow \partial f(x) = \begin{cases} \{|x| \leq 1\} & \text{per } x = 0, \\ \dfrac{x}{|x|} & \text{per } x > 0. \end{cases}$$

Dato un insieme $E \subset \mathbb{R}^n$, la funzione *Indicatrice* di E si definisce

$$I_E : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \qquad I_E(x) := \begin{cases} 0 & \text{se } x \in E, \\ +\infty & \text{se } x \notin E. \end{cases}$$

I_E è convessa se e solo se E è convesso. Dato un punto x_0 sulla frontiera ∂E, si chiama *cono normale* a E in x l'insieme

$$\mathcal{N}_E := \{v \in \mathbb{R}^n : v \cdot (x - x_0) \leq 0, \ \forall\, x \in E\}.$$

$\mathcal{N}_E$ è un cono, nel senso che $\lambda v \in \mathcal{N}_E$ per ogni $\lambda > 0$ e $v \in \mathcal{N}_E$, inoltre è chiuso e convesso.

3. Dati $a < b \in \mathbb{R}$, $f := I_{[a,b]}$
$$\Rightarrow \partial f(x) = \begin{cases} (-\infty, 0] & \text{per } x = a, \\ \{0\} & \text{per } x \in (a, b), \\ [0, +\infty) & \text{per } x = b, \\ \emptyset & \text{per } x \in \mathbb{R} \setminus [a, b]. \end{cases}$$

4. Dato $E \subset \mathbb{R}^n$, chiuso e convesso, sia $f := I_E$
$$\Rightarrow \partial f(x) = \begin{cases} \{0\} & \text{per } x \in E, \\ \mathcal{N}_E & \text{per } x \in \partial E, \\ \emptyset & \text{per } x \in \mathbb{R}^n \setminus E. \end{cases}$$

5. Sia $1 < p < +\infty$, $H = L^2(\mathbb{R}^n)$, $D(F) = W^{1,p}(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$,

$$F(u) := \frac{1}{p} \int_{\mathbb{R}^n} |\nabla u|^p \, dx.$$

Allora $\partial F(u)$ (su un certo $D(\partial F) \subset D(F)$) è ridotto ad un solo elemento, detto il p-Laplaciano di u:

$$\partial F(u) = -\Delta_p u := -\mathrm{div}\big(|\nabla u|^{p-2} \nabla u\big).$$

□ □ □ □ □

Esercizio 2.6
Si considerino le funzioni ai punti 1. e 4. dell'esempio precedente. Per ciascuna calcolare, e disegnare quanto possibile

$$x + \lambda \partial f(x), \qquad J_\lambda(x) = (x + \lambda \partial f(x))^{-1}, \qquad A_\lambda(x) = \frac{x - J_\lambda(x)}{\lambda}.$$

□ □ □ □ □

Nota: nella dimostrazione di [9, Theorem 3], si usa il seguente risultato (vedi, e.g., [6, Teorema 7.3]) per stabilire esistenza e unicità di una soluzione del problema di Cauchy $u_\lambda' + A_\lambda(u_\lambda) = 0$, $u_\lambda(0) = u$.

Teorema 2.7 (Cauchy-Lipschitz-Picard)**.** *Sia E uno spazio di Banach, e sia $F : E \to E$ un'applicazione tale che*

$$\|F(u) - F(v)\|_H \leq L \|u - v\|_H \qquad \forall\, u, v \in E,$$

per qualche costante $L \geq 0$. Allora, per ogni dato iniziale $u_0 \in E$ esiste un'unica $u \in C^1([0, +\infty[; E)$ tale che

$$\begin{cases} \dfrac{d}{dt} u = F(u) & \text{su } [0, +\infty[\\ u(0) = u_0. \end{cases}$$

Nota: Si usa anche, seppure un po' implicitamente, il prossimo Teorema 3.11 al momento del passaggio al limite per $\lambda \to 0$ nell'equazione del sottodifferenziale.

2.1 Funzione coniugata

Sia E uno spazio di Banach, E^* il suo duale, indichiamo la dualità fra $x \in E$ e $y \in E^*$ con il simbolo $\langle x, y \rangle$. Sia $\varphi : E \to \mathbb{R} \cup \{+\infty\}$, definiamo la funzione coniugata[2]

$$\varphi^* : E^* \to \mathbb{R} \cup \{+\infty\}, \qquad \varphi^*(y) := \sup_{x \in E} \{\langle x, y \rangle - \varphi(x)\}.$$

Se φ è propria, convessa e semicontinua inferiormente, allora

[2]alla quale son collegati, a seconda del testo e del contesto, i nomi di Legendre, Moreau, Fenchel, Young,...

- φ^* è propria, convessa e semicontinua inferiormente,
- $\partial\varphi^* : E^* \to 2^{E^*}$ è un operatore monotono (multivoco),
- $\varphi(x) + \varphi^*(y) \geq \langle x, y\rangle$ per ogni $x \in E$, per ogni $y \in E^*$.

Lemma 2.8. *Se φ è propria, convessa e semicontinua inferiormente e inoltre E è riflessivo, allora*

(i) $\varphi = \varphi^{**}$,

(ii) sono equivalente le seguenti tre condizioni

$$\text{(a) } y \in \partial\varphi(x) \qquad \text{(b) } x \in \partial\varphi^*(y) \qquad \text{(c) } \varphi(x) + \varphi^*(y) = \langle x, y\rangle.$$

Esempio 2.9

1. Sia E uno spazio di Banach riflessivo, e identifichiamo E^{**} con E. Sia $\varphi(x) := \frac{1}{2}\|x\|_E^2$. Applicando la definizione, si vede che $\varphi(x) = \varphi^*(x)$.
2. Sia $E = \mathbb{R}$, $\varphi(x) = |x|$. Allora φ^* è la funzione indicatrice dell'intervallo $[-1, 1]$. (Si veda l'Esempio 2.5.)

□ □ □ □ □

Esercizio 2.10
Nei due esempi precedenti, verificare l'equivalenza delle tre condizioni date dal Lemma 2.8 (ii). (Per esempio, verificare che (a)⇒(b)⇒(c)⇒(a).)

□ □ □ □ □

2.2 Un'applicazione in meccanica dei continui

Diamo ora una breve descrizione di un modello per la deformazione plastica di un materiale. Per la descrizione del problema usiamo un dominio di riferimento $\Omega \subset \mathbb{R}^3$ occupato dal materiale plastico, e un intervallo di tempo $[0, T]$, con $\Omega_T := \Omega \times (0, T)$. Le variabili dipendenti sono lo spostamento

$$u : \Omega_T \to \mathbb{R}^3,$$

il cui gradiente simmetrizzato $\nabla^s u = \frac{1}{2}(\nabla u + (\nabla u)^t) : \Omega_T \to \mathbb{R}^{3\times 3}$ rappresenta il tensore di deformazione linearizzato, e il tensore degli stress di Cauchy

$$\sigma : \Omega_T \to \mathbb{R}^{3\times 3}.$$

Inoltre, decomponiamo il gradiente simmetrizzato di u in parte elastica e parte plastica

$$\nabla^s u = \varepsilon_{\text{el}} + \varepsilon_{\text{pl}}.$$

Le equazioni del modello sono: la conservazione del momento lineare

$$\varrho \partial_t^2 u - \operatorname{div} \sigma = f, \tag{2.1}$$

dove ϱ è la densità del materiale (dipendente da x, ma non da t) e f è la densità delle forze esterne (per esempio forza di gravità) che agiscono sul corpo. La conservazione del momento angolare, invece, si esprime imponendo che σ sia simmetrico. I materiali elastici sono caratterizzati da una dipendenza lineare fra deformazione $\nabla^s u$ e stress σ. I materiali plastici esibiscono una risposta elastica per deformazioni limitate a una determinata regione di $\mathbb{R}^{3\times 3}$, oltre la quale anche piccole forze possono risultare in grandi deformazioni. Secondo un modello comune, assumiamo che lo stress dipenda solo dalla parte elastica della deformazione e che la loro relazione sia determinata dal tensore di *compliance* C

$$C\sigma = \nabla^s u - \varepsilon_{\mathrm{pl}}. \tag{2.2}$$

Infine, rappresentiamo il comportamento plastico attraverso l'inclusione differenziale

$$\partial_t \varepsilon_{\mathrm{pl}} \in \partial\varphi(\sigma), \tag{2.3}$$

dove $\varphi : \mathbb{R}^{3\times 3} \to \mathbb{R} \cup \{+\infty\}$ è la funzione indicatrice della regione ammissibile degli stress.

È possibile (e molto più pratico) incorporare (2.1)–(2.3) in un'unica formulazione variazionale. Vediamo come questa si può derivare, almeno formalmente, assumendo che tutte le funzioni abbiano la regolarità necessaria perché i passaggi siano giustificati e che siano presenti condizioni al contorno e dati iniziali consistenti. Notazione: per due vettori $x = (x_i)$ e $y = (y_i)$, $i = 1, 2, 3$, denotiamo il prodotto scalare $x \cdot y = \sum_i x_i y_i$. Per due tensori del secondo ordine $A = (a_{ij})$ e $B = (b_{ij})$, $i, j = 1, 2, 3$, denotiamo il prodotto scalare $A : B = \sum_{ij} a_{ij} b_{ij}$. Moltiplichiamo (2.1) per $\partial_t u$, integriamo su Ω e operiamo qualche passaggio

$$\begin{aligned}
0 &= \int_\Omega \varrho \partial_t^2 u \cdot \partial_t u - \operatorname{div}\sigma \cdot \partial_t u - f \cdot \partial_t u \, dx \\
&= \int_\Omega \frac{\varrho}{2} \partial_t |\partial_t u|^2 + \sigma : \partial_t \nabla^s u - f \cdot \partial_t u \, dx \\
&\overset{(2.2)}{=} \int_{\Omega_T} \frac{\varrho}{2} \partial_t |\partial_t u|^2 + \sigma : \partial_t (C\sigma + \varepsilon_{\mathrm{pl}}) - f \cdot \partial_t u \, dx \\
&= \int_\Omega \frac{\varrho}{2} \partial_t |\partial_t u|^2 + \frac{1}{2} \partial_t (\sigma : C\sigma) + \sigma : \partial_t \varepsilon_{\mathrm{pl}} - f \cdot \partial_t u \, dx \\
&= \frac{1}{2} \frac{d}{dt} \left\{ \int_\Omega \varrho |\partial_t u|^2 + \sigma : C\sigma \, dx \right\} + \int_\Omega \sigma : \partial_t \varepsilon_{\mathrm{pl}} - f \cdot \partial_t u \, dx
\end{aligned}$$

Integriamo in tempo sull'intervallo (0,t) e giungiamo all'equazione

$$\begin{aligned}
&\frac{1}{2} \int_\Omega \varrho |\partial_t u(t)|^2 + \sigma(t) : C\sigma(t) \, dx + \int_{\Omega_T} \sigma : \partial_t \varepsilon_{\mathrm{pl}} \, dx \, dt \\
&\qquad = \frac{1}{2} \int_\Omega \varrho |\partial_t u(0)|^2 + \sigma(0) : C\sigma(0) \, dx + \int_{\Omega_T} f \cdot \partial_t u \, dx \, dt
\end{aligned}$$

Infine, grazie al Lemma 2.8 e a (2.3), sostituiamo

$$\sigma : \partial_t \varepsilon_{\text{pl}} = \varphi(\sigma) + \varphi^*(\partial_t \varepsilon_{\text{pl}}), \tag{2.4}$$

ottenendo

$$\begin{aligned} \frac{1}{2} \int_\Omega \varrho |\partial_t u(t)|^2 + \sigma(t) : C\sigma(t)\, dx + \int_{\Omega_T} \varphi(\sigma) + \varphi^*(\partial_t \varepsilon_{\text{pl}})\, dx\, dt \\ = \frac{1}{2} \int_\Omega \varrho |\partial_t u(0)|^2 + \sigma(0) : C\sigma(0)\, dx + \int_{\Omega_T} f \cdot \partial_t u\, dx\, dt. \end{aligned}$$

Quest'equazione rappresenta un bilancio di energia: al tempo t, la somma di energia cinetica $\frac{1}{2} \int \varrho |\partial_t u|^2$, energia elastica immagazzinata $\frac{1}{2} \int \sigma : C\sigma$ ed energia dissipata dalla trasformazione plastica $\int \sigma : \partial_t \varepsilon_{\text{pl}}$ equivale alla somma di energia cinetica ed energia elastica al tempo $t = 0$ più lavoro delle forze esterne $\in f \partial_t u$. In particolare, l'uguaglianza in (2.4), dove per le proprietà generali di φ, φ^* vale sempre un '$\leq$', corrisponde al principio secondo cui l'evoluzione in termodinamica segue la via di massima dissipazione possibile.

3 Problemi di minimo nella topologia debole di L^1

3.1 Semicontinuità inferiore per funzionali integrali

Il seguente Teorema è tratto da [7, Example 1.21, Example 1.23]

Teorema 3.11. *Sia $(\Omega, \mathcal{A}, \mu)$ uno spazio di misure e sia $X := L^p_\mu(\Omega; \mathbb{R}^m)$, con $m \geq 1$ intero e $p \geq 1$ reale. Sia $f : \Omega \times \mathbb{R} \cup \{\pm\infty\}$ una funzione $\mathcal{A} \otimes \mathcal{B}(\mathbb{R}^m)$ misurabile, dove $\mathcal{B}(\mathbb{R}^m)$ è la σ-algebra di Borel di $\mathbb{R}^m$. Se*

(a) *esistono una funzione $a \in L^1_\mu(\Omega)$ e una costante $b \in \mathbb{R}$ tali che*

$$f(x, v) \geq a(x) + b|v|^p \qquad \text{per } \mu\text{-q.o. } x \in \Omega, \quad \forall\, v \in \mathbb{R}^m,$$

(b) *la funzione $v \mapsto f(x, v)$ è semicontinua inferiormente per μ-q.o. $x \in \Omega$,*

allora il funzionale $F : X \to \mathbb{R} \cup \{+\infty\}$

$$F(u) := \int_\Omega f(x, u(x))\, d\mu(x)$$

è semicontinuo inferiormente per la topologia forte di X. Se inoltre

(c) *la funzione $v \mapsto f(x, v)$ è convessa per μ-q.o. $x \in \Omega$,*

allora F è semicontinuo inferiormente per la topologia debole di X.

Proof. Passo I. Semicontinuità forte significa: per ogni successione $\{u_n\} \subset X = L^p_\mu(\Omega; \mathbb{R}^m)$, tale che $u_n \to u$ fortemente, si ha

$$\liminf_{n \to +\infty} \int_\Omega f(x, u_n(x))\, d\mu(x) \geq \int_\Omega f(x, u(x))\, d\mu(x). \tag{3.1}$$

Sia dunque $\{u_n\}$ una successione convergente fissata e sia $\ell := \liminf_{n\to\infty} F(u_n)$. A meno di sottosuccessioni, possiamo assumere che

$$\lim_{n\to\infty} F(u_n) = \ell,$$

e, grazie al teorema della convergenza dominata di Lebesgue, che

$$u_n(x) \to u(x) \qquad \text{per } \mu\text{-q.o. } x \in \Omega.$$

Definiamo la funzione

$$g(x, v) := f(x, v) - a(x) - b|v|^p.$$

Grazie all'ipotesi (a), $g(x, v) \geq 0$ per μ-q.o. $x \in \Omega$, per ogni $v \in \mathbb{R}^m$. Applichiamo il Lemma di Fatou:

$$\liminf_{n\to+\infty} \int_\Omega g(x, u_n(x))\, d\mu(x) \geq \int_\Omega \left\{ \liminf_{n\to+\infty} g(x, u_n(x)) \right\} d\mu(x).$$

Dato che f è s.c.i. e $u_n(x) \to u(x)$ quasi ovunque,

$$\liminf_{n\to+\infty} g(x, u_n(x)) \geq f(x, u) - a(x) - b|u(x)|^p,$$

e quindi

$$\liminf_{n\to+\infty} \int_\Omega g(x, u_n(x))\, d\mu(x) \geq F(u) - \int_\Omega a(x) + b|u(x)|^p\, d\mu(x). \tag{3.2}$$

Ora, dato che $u_n \to u$ in X,

$$\lim_{n\to\infty} \int_\Omega a(x) + b|u_n(x)|^p\, d\mu(x) = \int_\Omega a(x) + b|u(x)|^p\, d\mu(x),$$

e quindi da (3.2) otteniamo (3.1).

Passo II. Per la semicontinuità debole, $u_n \rightharpoonup u$ e non possiamo contare sulla convergenza di $\int |u_n|^p$ (avremo solo semicontinuità, ma questo non serve perchè b non ha segno). Ricordiamo la definizione di epigrafico (applicato al funzionale F):

$$\operatorname{epi}(F) := \big\{(u, y) \in X \times \mathbb{R} \;:\; y \geq F(u)\big\}.$$

Proprietà note dell'epigrafico:

- $\operatorname{epi}(F)$ è convesso se e solo se F è convessa;
- $\operatorname{epi}(F)$ è chiuso rispetto alla topologia forte se e solo se F è semicontinua inferiormente nella topologia forte;
- $\operatorname{epi}(F)$ è chiuso rispetto alla topologia debole se e solo se F è semicontinua inferiormente nella topologia debole.

Un rapido conto mostra che se $v \mapsto f(x, v)$ è convessa, allora il funzionale F è convesso. Sappiamo dal passo I che F è s.c.i. forte, e quindi $\operatorname{epi}(F)$ è convesso e chiuso forte. Per il Teorema di Mazur, allora $\operatorname{epi}(F)$ è chiuso debole e possiamo concludere che F è s.c.i. per la topologia debole di X. $\square$

3.2 Uniforme integrabilità e compattezza debole in L^1

Questa sezione è tratta da [5, Section 4.5]. Sia $(X, \mathcal{A}, \mu)$ uno spazio di misura con una misura non negativa μ (finita o con valori in $[0, +\infty]$).

Definizione 3.12
Un insieme di funzioni $\mathcal{F} \subset L^1(\mu)$ si dice *uniformemente integrabile* se

$$\lim_{c \to +\infty} \sup_{f \in \mathcal{F}} \int_{\{|f| \geq c\}} |f(x)| \, d\mu(x) = 0.$$

Esempio 3.13
In $X = \mathbb{R}$:

1. (Travelling) $\mu_n = \chi_{[n,n+1]}(x)\mathcal{L}^1$ (non tight, uniformemente integrabili, misura ∞)
2. (Concentration) $\mu_n = \frac{n}{2}\chi_{[-1/n,1/n]}(x)\mathcal{L}^1$ (tight, non unif. integrabili, misura $< \infty$)
3. (Diffusion) $\mu_n = \frac{1}{n}\chi_{[0,n]}(x)\mathcal{L}^1$ (non tight, uniformemente integrabili, misura ∞)

Nessuna delle tre successioni converge debolmente in $L^1(\mathbb{R})$. (La seconda converge in misura, ma il limite non è una funzione.) Congettura: tightness + uniforme integrabilità = compattezza debole in L^1.

□ □ □ □ □

Definizione 3.14
Un insieme di funzioni $\mathcal{F} \subset L^1(\mu)$ ha *integrali uniformemente assolutamente continui* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che

$$\int_A |f(x)| \, d\mu(x) \leq \varepsilon \qquad \forall\, f \in \mathcal{F}, \quad \forall\, A \in \mathcal{A} : \mu(A) \leq \delta.$$

Nota, in $(\mathbb{R}^n, \mathcal{L}^n)$, uniforme assoluta continuità degli integrali e uniforme integrabilità sono equivalenti. Nel caso di misure più generali (non atomless), l'uniforme assoluta continuità degli integrali non dice nulla sulla limitatezza degli integrali, che va quindi chiesta a parte.

Teorema 3.15. [5, Theorem 4.7.18] *Sia μ una misura finita e sia $\mathcal{F}$ un insieme di funzioni μ-integrabili. Sono equivalenti:*

(i) $\mathcal{F}$ è uniformemente integrabile rispetto a μ;

(ii) $\mathcal{F}$ ha chiusura compatta per la topologia debole di $L^1(\mu)$.

Proof. Dimostriamo solo l'implicazione (i) $\Rightarrow$ (ii). Sia $\mathcal{F}$ un insieme uniformemente integrabili, allora $\mathcal{F}$ è limitato in $L^1(\mu)$. Consideriamo l'immersione $L^1(\mu) \hookrightarrow L^\infty(\mu)^*$:

$$g \in L^1(\mu) \mapsto G \in L^\infty(\mu)^* : \quad G(\phi) := \int_X g(x)\phi(x)\, d\mu.$$

Sia $\mathcal{H}$ la chiusura di $\mathcal{F}$ rispetto alla topologia debole* di $L^\infty(\mu)^*$. Per il Teorema di Banach-Alaoglu, un chiuso in $L^\infty(\mu)^*$ è compatto per la topologia debole*. Mostriamo che $\mathcal{H} \subset L^1(\mu)$. Per costruzione $G \in \mathcal{H}$ è un funzionale lineare e continuo su $L^\infty(\mu)$, per il quale esiste una successione $\{f_n\} \subset \mathcal{F}$ tale che

$$G(\phi) = \lim_{n\to+\infty} \int_X f_n(x)\phi(x)\, dx.$$

Indicando con χ_A la funzione caratteristica dell'insieme A, definiamo la funzione d'insieme

$$\nu(A) := G(\chi_A) = \lim_{n\to+\infty} \int_A f_n(x)\, d\mu.$$

Grazie all'equivalenza fra uniforme integrabilità e assoluta continuità degli integrali, l'insieme $\mathcal{F}$ ha integrali uniformemente assolutamente continui (rispetto a μ). Si ha dunque

$$|\nu(A)| \leq \limsup_{n\to+\infty} \left| \int_A f_n(x)\, d\mu \right| \leq \limsup_{n\to+\infty} \int_A |f_n(x)|\, d\mu \leq \varepsilon$$

per ogni $A \in \mathcal{A} : \mu(A) \leq \delta$. Concludiamo che ν è assolutamente continua rispetto a μ, e quindi ν è σ-additiva (si veda, a questo proposito, [5, Proposition 1.3.3]) e dunque è una misura. Per il Teorema di Radon-Nikodym esiste infine $f \in L^1(\mu)$ tale che

$$\nu(A) = \int_A f(x)\, d\mu.$$

A questo punto G coincide con un funzionale integrale sulle funzioni semplici di X e

$$G(\phi) = \lim_{n\to+\infty} \int_X \phi(x) f_n(x)\, d\mu = \int_X \phi(x) f(x)\, d\mu$$

per ogni $\phi \in L^\infty(\mu)$. □

In molti casi, un risultato valido per misure finite si può estendere a misure σ-finite grazie al seguente trucco:

Proposizione 3.16. [5, Proposition 2.6.2] *Per ogni misura σ-finita μ esiste una funzione ϱ strettamente positiva, μ-integrabile, con immagine numerabile, tale che:*

(i) una funzione f è μ-integrabile se e solo se la funzione f/ϱ è integrabile rispetto alla misura $\nu = \varrho \cdot \mu$ definita da

$$\nu(A) := \int_A \varrho(x)\, d\mu(x), \qquad A \in \mathcal{A};$$

(ii) la misura ν è finita;

(iii) vale:

$$\int_X f(x)\, d\mu(x) = \int_X \frac{f(x)}{\varrho(x)}\, d\nu(x).$$

La numerabilità della funzione ϱ può essere utile nel caso di misure complicate. Nel caso della misura di Lebesgue su $\mathbb{R}^n$, invece, può essere utile rinunciare alla numerabilità in cambio di una funzione continua, per esempio

$$\varrho(x) = \frac{e^{-\frac{|x|^2}{2}}}{\sqrt{(2\pi)^n}} \qquad \Rightarrow \qquad \nu(\mathbb{R}^n) = 1. \tag{3.3}$$

Un'osservazione: non si pensi che la proposizione appena enunciata permetta di trasferire automaticamente proprietà da μ a ν. Per esempio, l'insieme $\mathcal{F}$ dato dalla successione *travelling* dell'esempio 3.13 è un caso di insieme uniformemente integrabile, ma non debolmente relativamente compatto. Allora, l'insieme $\mathcal{F}/\varrho = \{f/\varrho : f \in \mathcal{F}\}$ non può essere uniformemente integrabile rispetto a $\nu = \varrho \cdot \mu$, altrimenti, per il Teorema 3.15 $\mathcal{F}/\varrho$ sarebbe debolmente compatto, e quindi per la Proposizione 3.16 sarebbe compatto $\mathcal{F}$ (falso!). Vediamo a mano:

Esempio 3.17
Dati

$$\mathcal{F} = \left\{ \chi|_{[n,n+1]} : n \in \mathbb{N} \right\}, \qquad \varrho(x) = \frac{e^{-x^2}}{\sqrt{\pi}}, \qquad \nu = \varrho \cdot \mathcal{L}^1$$

chiedere che $\mathcal{F}/\varrho$ sia uniformemente integrabile rispetto a ν significa chiedere che

$$\begin{aligned} 0 &= \lim_{c\to+\infty} \sup_{g\in\mathcal{F}/\varrho} \int_{\{g\geq c\}} |g(x)|\, d\nu(x) \\ &= \lim_{c\to+\infty} \sup_{f\in\mathcal{F}} \int_{\{f/\varrho\geq c\}} |f/\varrho(x)|\, \varrho(x) dx \\ &= \lim_{c\to+\infty} \sup_{n\in\mathbb{N}} \int_{\{f_n\geq c\varrho(x)\}} f_n(x)\, dx. \end{aligned}$$

Dato che $\varrho \to 0$ per $x \to +\infty$, per ogni $c > 0$ si può trovare un $n > 0$ tale che $c\varrho(n) < 1$, per cui

$$\sup_{n\in\mathbb{N}} \int_{\{f_n\geq c\varrho(x)\}} f_n(x)\, dx = 1.$$

Conclusione: $\mathcal{F}/\varrho$ non è uniformemente integrabile.

□ □ □ □ □

Teorema 3.18 (Criterio di De la Valleé-Poussin)**.** *Sia μ una misura* finita *non negativa. Una famiglia $\mathcal{F}$ di funzioni μ-integrabili è uniformemente integrabile se e solo se esiste una funzione crescente non negativa $G : [0, +\infty) \to [0, +\infty)$ tale che*

$$\lim_{t\to+\infty} \frac{G(t)}{t} = +\infty \qquad e \qquad \sup_{f\in\mathcal{F}} \int_X G(|f(x)|) d\mu(x) < +\infty.$$

Proof. 1. Dimostriamo che l'esistenza di tale G implica l'uniforme integrabilità di $\mathcal{F}$. Sia $\varepsilon > 0$ fissato, e sia

$$M := \sup_{f\in\mathcal{F}} \int_X G(|f(x)|)d\mu(x).$$

Per la condizione di superlinearità di G, esiste una $c > 0$ tale che, se $t \geq c$, allora $G(t)/t \geq M/\varepsilon$, e quindi, per ogni $f \in \mathcal{F}$

$$|f(x)| \leq \frac{\varepsilon}{M} G(|f(x)|) \qquad \forall\, x : |f(x)| \geq c.$$

Per tale c stimiamo

$$\int_{\{|f|\geq c\}} |f(x)|d\mu(x) \leq \frac{\varepsilon}{M} \int_{\{|f|\geq c\}} G(|f(x)|)d\mu(x) \leq \varepsilon.$$

Dato che la stima è uniforme su $\mathcal{F}$ ed ε è arbitrario, abbiamo dimostrato che $\mathcal{F}$ è uniformemente integrabile.

2. Sia $\mathcal{F}$ uniformemente integrabile, costruiamo una G nelle ipotesi del Teorema. L'idea è di definire

$$G(t) = \int_0^t g(s)\, ds,$$

per una funzione g a scala, definita attraverso i coefficienti α_n, $n \in \mathbb{N}$, cioè

$$g(t) = \alpha_n, \quad \text{per } t \in [n, n+1[.$$

Per $f \in \mathcal{F}$, definiamo

$$\mu_n(f) := \mu\,(x : f(x) \geq n)\,.$$

Per l'uniforme integrabilità, possiamo trovare una sottosuccessione di numeri interi $\{C_n\}$ tale che

$$\sup_{f\in\mathcal{F}} \int_{\{|f|\geq C_n\}} |f(x)|d\mu \leq \frac{1}{2^n}. \tag{3.4}$$

Per $n \in \mathbb{N}$ stimiamo

$$\int_{\{|f|\geq C_n\}} |f(x)|d\mu \geq \sum_{k=C_n}^{+\infty} \mu_k(f),$$

e quindi, grazie a (3.4)

$$\sum_{n=1}^{+\infty} \sum_{k=C_n}^{+\infty} \mu_k(f) \leq 1 \tag{3.5}$$

per ogni $f \in \mathcal{F}$. Definiamo $\alpha_n = 0$ se $n < C_1$ e $\alpha_n := \max\{k \in \mathbb{N} : C_k \leq n\}$. Allora si ha

$$\begin{aligned}
\int_X G(|f(x)|)d\mu &= \alpha_1\mu\big(x:\ 1 < |f(x)| \leq 2\big) \\
&\quad + (\alpha_1 + \alpha_2)\mu\big(x:\ 2 < |f(x)| \leq 3\big) \\
&\quad + (\alpha_1 + \alpha_2 + \alpha_3)\mu\big(x:\ \ \ldots \\
&= \sum_{n=1}^{+\infty} \alpha_n \mu_n(f) \\
&= \sum_{n=1}^{+\infty} \sum_{k=C_n}^{+\infty} \mu_k(f) \leq 1.
\end{aligned}$$

Le proprietà di positività, monotonia, superlinearità (e convessità) di G seguono facilmente dalle proprietà di g. $\square$

Si noti che la condizione di finitezza della misura è fondamentale: in $\mathbb{R}$ non è difficile costruire una successione di funzioni caratteristiche che sia uniformemente integrabile, ma non debolmente relativamente compatta.

Nell'applicazione allo schema discreto iterativo di [10] è utile il seguente

Lemma 3.19. *Per $t \geq 0$ definiamo $G(t) := \max\{t \log t, 0\}$. Sia $\{f_j\}_j \subset L^1(\mathbb{R}^n)$ una successione di funzioni tali che*

$$\text{(i)} \quad f_j(x) \geq 0 \quad \forall\, x \in \mathbb{R}^n; \tag{3.6}$$

$$\text{(ii)} \quad \sup_j \int_{\mathbb{R}^n} |x|^2 f_j(x)\, dx < +\infty; \tag{3.7}$$

$$\text{(iii)} \quad \sup_j \int_{\mathbb{R}^n} G(f_j(x))\, dx < +\infty. \tag{3.8}$$

Allora esiste una sottosuccessione $\{f_{j_k}\}_k$ e una funzione $f \in L^1(\mathbb{R}^n)$, tale che $f_{j_k} \rightharpoonup f$ in $L^1(\mathbb{R}^n)$.

Proof. Siano

$$\varrho(x) := e^{-|x|^2}, \quad \nu = \varrho \cdot \mathcal{L}^n, \quad \mathcal{F} = \{f_j\}_j, \qquad G(t) = \max\{t \log t, 0\}$$

Passo I. $\mathcal{F}/\varrho$ soddisfa le ipotesi del Teorema di De la Vallée-Poussin rispetto a ν, con funzione superlineare G. Calcoliamo prima

$$\begin{aligned}\int_{\mathbb{R}^n} \left(\frac{f_j}{\varrho}\right) \log\left(\frac{f_j}{\varrho}\right) \varrho\, dx &= \int_{\mathbb{R}^n} f_j \left(\log f_j - \log \varrho\right)\, dx \\ &= \int_{\mathbb{R}^n} f_j \left(\log f_j - \log \varrho\right)\, dx \\ &= \int_{\mathbb{R}^n} f_j \log f_j\, dx + \int_{\mathbb{R}^n} |x|^2 f_j\, dx.\end{aligned}$$

Quindi

$$\begin{aligned}\sup_j \int_{\mathbb{R}^n} G(f_j/\varrho)\, d\nu &= \sup_j \int_{\mathbb{R}^n} \max\{f_j \log f_j + |x|^2 f_j, 0\}\, dx \\ &\leq \sup_j \int_{\mathbb{R}^n} \max\{f_j \log f_j, 0\}\, dx + \sup_j \int_{\mathbb{R}^n} |x|^2 f_j\, dx < +\infty.\end{aligned}$$

Passo II. Dato che ν è una misura finita, per il Teorema di De la Vallée-Poussin $\mathcal{F}/\varrho$ è uniformemente integrabile rispetto a ν.

Passo III. Per il Teorema 3.15, $\mathcal{F}/\varrho$ è debolmente relativamente compatto in $L^1(\nu)$, ovvero, esistono una sottosuccessione $\{f_{j_k}\}$ e una funzione $g \in L^1(\nu)$ tali che $f_{j_k}/\varrho \rightharpoonup g$ in $L^1(\nu)$. Cioè

$$\lim_{k \to +\infty} \int_{\mathbb{R}^n} \frac{f_{j_k}}{\varrho} \phi\, d\nu = \int_{\mathbb{R}^n} g\phi\, d\nu \qquad \forall\, \phi \in L^\infty(\nu).$$

Definendo $f := g\varrho \in L^1(\mathbb{R}^n)$, dato che $L^\infty(\nu) = L^\infty(\mathbb{R}^n)$, si ha

$$\lim_{k \to +\infty} \int_{\mathbb{R}^n} f_{j_k} \phi\, dx = \int_{\mathbb{R}^n} f\phi\, dx \qquad \forall\, \phi \in L^\infty(\mathbb{R}^n),$$

cioè la tesi. $\square$

4 Lo spazio delle misure di probabilità

4.1 Compattezza e semicontinuità in $\mathscr{P}(X)$

Questa sezione è tratta da [3, Section 1].

Sia (X, d) uno spazio metrico separabile, indichiamo con $\mathscr{P}(X)$ l'insieme delle misure di probabilità $\mu : \mathcal{B}(X) \to [0, 1]$, dove $\mathcal{B}(X)$ è la σ-algebra di Borel. Il supporto di μ è l'inseme chiuso

$$\operatorname{supp}(\mu) := \{x \in X \ : \ \mu(B_r(x)) > 0, \forall\, r > 0\}.$$

($B_r(x)$ è la palla aperta di centro x e raggio r.) Quando X è un boreliano dello spazio euclideo $\mathbb{R}^m$, poniamo

$$M_2(\mu) := \int_X |x|^2 \, d\mu.$$

È pratica comune, per un tale X, identificare

$$\mathscr{P}(X) = \{\mu \in \mathscr{P}(\mathbb{R}^m) : \mu(\mathbb{R}^m \setminus X) = 0\}.$$

Indichiamo con $\mathscr{P}_2(X)$ il sottospazio di $\mathscr{P}(X)$ composto dalle misure di momento secondo finito:

$$\mathscr{P}_2(X) := \{\mu \in \mathscr{P}(X) : M_2(\mu) < +\infty\}.$$

Indichiamo con $\mathscr{L}^m$ la misura di Lebesgue in $\mathbb{R}^m$ e quando $X \subset \mathcal{B}(\mathbb{R}^m)$ poniamo

$$\mathscr{P}_2^a(X) := \{\mu \in \mathscr{P}_2(X) : \mu \ll \mathscr{L}^m\},$$

cioè, il sottospazio di $\mathscr{P}_2(X)$ delle misure assolutamente continue rispetto alla misura di Lebesgue. ($\mu \in \mathscr{P}_2^a(X) \Leftrightarrow \mu \in \mathscr{P}_2(X)$ e $\mu(A) = 0$ per ogni $A \subset \mathcal{B}(\mathbb{R}^m)$ tale che $\mathscr{L}^m(A) = 0$.)

Come in Probabilità, diciamo che una successione $(\mu_k) \subset \mathscr{P}(X)$ converge *strettamente* a $\mu \in \mathscr{P}(X)$ per $n \to \infty$ se

$$\lim_{n\to\infty} \int_X f(x) \, d\mu_n(x) = \int_X f(x) \, d\mu(x) \tag{4.1}$$

per ogni funzione $f \in C_b^0(X)$, lo spazio delle funzioni continue e limitate su X a valori reali. (Detta anche "narrow convergence".)

Teorema 4.20 (Prokhorov)**.** *[3, Theorem 1.1] Se un insieme $K \subset \mathscr{P}(X)$ è* tight*, cioè*

$$\forall\, \varepsilon > 0 \quad \exists\, K_\varepsilon \text{ compatto contenuto in } X \text{ tale che } \quad \mu(\mathbb{R}^n \setminus K_\varepsilon) \le \varepsilon \quad \forall\, \mu \in K, \tag{4.2}$$

allora K è relativamente compatto in $\mathscr{P}(X)$. Viceversa, se X è uno spazio Polacco, ogni sottoinsieme relativamente compatto di $\mathscr{P}(X)$ è tight*.*

(Uno spazio topologico si dice "Polacco" se è omeomorfo a uno spazio metrico completo e separabile. L'insieme vuoto è solitamente incluso nella classe degli spazi Polacchi.)

Lemma 4.21 (Un criterio integrale per la tightness)**.** *La condizione* (4.2) *è equivalente alla seguente: Esiste una funzione* $\varphi : X \to [0, +\infty]$*, i cui sottolivelli* $\{x \in X : \varphi(x) \leq c\}$ *sono compatti, tale che*

$$\sup_{\mu \in K} \int_X \varphi(x) d\mu(x) < +\infty. \tag{4.3}$$

Proof. 1. (4.2) $\Rightarrow$ (4.3)
Sia $\{\varepsilon_n\}_n$ una successione tale che $\sum_{n=0}^{+\infty} \varepsilon_n < +\infty$ e sia $K_n = K_{\varepsilon_n}$ una successione di compatti che soddisfa (4.2). Definiamo la funzione

$$\varphi(x) := \inf\{n \geq 0 : x \in K_n\} = \sum_{n=0}^{+\infty} \chi_{X \setminus K_n}(x).$$

Allora $\varphi \geq 0$, i sottolivelli di φ sono esattamente gli insiemi K_n, compatti, e

$$\sup_{\mu \in K} \int_X \varphi(x) d\mu(x) = \sup_{\mu \in K} \sum_{n=0}^{+\infty} \mu(X \setminus K_n) \leq \sum_{n=0}^{+\infty} \varepsilon_n < +\infty.$$

2. (4.3) $\Rightarrow$ (4.2)
La disuguaglianza di Chebyshev applicata a φ è

$$\mu\left(\{x \in X : \varphi(x) \geq t\}\right) \leq \frac{1}{t} \int_X \varphi(x)\, d\mu(x) \qquad \forall\, t > 0.$$

Sia M il valore del sup in (4.3). Dato $\varepsilon > 0$ scegliamo $t = M/\varepsilon$, allora $K_\varepsilon := \{\varphi \leq M/\varepsilon\}$ è un sottolivello di φ, e quindi compatto e inoltre, per ogni $\mu \in K$

$$\mu(X \setminus K_\varepsilon) = \mu\left(\{x \in X : \varphi(x) > M/\varepsilon\}\right) \leq \frac{\varepsilon}{M} \int_X \varphi(x)\, d\mu(x) \leq \varepsilon.$$

□

Quando serve passare al limite in espressioni come (4.1), ma con funzione integranda f non limitata, o solo *semi*continua, sono utili le seguenti proprietà. Ci serve questa definizione: diciamo che una funzione Borel-misurabile $g : X \to [0, +\infty]$ è uniformemente integrabile rispetto a un determinato insieme $\mathcal{K} \subset \mathscr{P}(X)$ se

$$\sup_{\mu \in \mathcal{K}} \lim_{k \to \infty} \int_{\{g \geq k\}} g(x)\, d\mu(x) = 0, \tag{4.4}$$

dove $\{g \geq k\}$ è un'abbreviazione per $\{x \in X : g(x) \geq k\}$. Nel caso particolare in cui g è una potenza della distanza in X da un punto fissato, $g(x) = d(x, \bar{x})^p$, per $p > 0$, ovvero se

$$\sup_{\mu \in \mathcal{K}} \lim_{k \to \infty} \int_{X \setminus B_k(\bar{x})} d(x, \bar{x})^p\, d\mu(x) = 0, \tag{4.5}$$

diciamo che $\mathcal{K} \subset \mathscr{P}(X)$ ha i *momenti* p*-esimi uniformemente integrabili*, oppure che è p*-uniformemente integrabile.*

Esercizio 4.22

(a) Dimostrare che se $\mathcal{K} \subset \mathscr{P}(X)$ ha i momenti p-esimi uniformemente integrabili rispetto a un punto base $x^* \in X$, allora ha i momenti p-esimi uniformemente integrabili rispetto a ogni $\bar{x} \in X$.

(b) Dimostrare che se

$$\sup_{\mu \in \mathcal{K}} \int_X d(x, \bar{x})^{p_1} \, d\mu(x) < +\infty,$$

allora $\mathcal{K} \subset \mathscr{P}(X)$ ha i momenti p-esimi uniform. integrabili per ogni $0 < p < p_1$.

□ □ □ □ □

Il Lemma seguente fornisce una caratterizzazione delle famiglie p-uniformemente integrabili, estendendo la validità di (4.1) a funzioni non limitate, ma con crescita 'p', ovvero a funzioni $f : X \to \mathbb{R}$ tali che

$$|f(x)| \leq A + B d(x, \bar{x})^p \qquad \forall\, x \in X,$$

per certi valori di $A, B \geq 0$ e $\bar{x} \in X$.

Lemma 4.23. *Sia μ_n una successione in $\mathscr{P}(X)$, strettamente convergente a $\mu \in \mathscr{P}(X)$. Sia $g : X \to (-\infty, +\infty]$ semicontinua inferiormente, tale che g^- sia uniformemente integrabile rispetto a $\{\mu_n\}$. Allora*

$$\liminf_{n \to \infty} \int_X g(x) \, d\mu_n(x) \geq \int_X g(x) \, d\mu(x) > -\infty. \tag{4.6}$$

Sia $f : X \to \mathbb{R}$ continua, tale che $|f|$ sia uniformemente integrabile rispetto a $\{\mu_n\}$. Allora

$$\liminf_{n \to \infty} \int_X f(x) \, d\mu_n(x) = \int_X f(x) \, d\mu(x). \tag{4.7}$$

Viceversa, se $f : X \to \mathbb{R}$ è continua, μ_n-integrabile e

$$\limsup_{n \to \infty} \int_X f(x) \, d\mu_n(x) \leq \int_X f(x) \, d\mu(x) < +\infty, \tag{4.8}$$

allora f è uniformemente integrabile rispetto a $\{\mu_n\}$. In particolare, una famiglia $\{\mu_n\} \subset \mathscr{P}(X)$ è p-uniformemente integrabile se e solo se (4.7) è verificata per ogni funzione continua $f : X \to \mathbb{R}$ con crescita p.

4.2 Trasporto di misure

Sezione tratta da [3, Section 1.1]. Siano X_1, X_2 spazi metrici separabili. Sia $\mu \in \mathscr{P}(X_1)$ e $\mathbf{t} : X_1 \to X_2$ una funzione Borel-misurabile. Indichiamo con $\mathbf{t}_\# \mu \in \mathscr{P}(X_2)$ il *push-forward* di μ attraverso $\mathbf{t}$, definito da

$$\mathbf{t}_\# \mu(B) := \mu(\mathbf{t}^{-1}(B))$$

per ogni Boreliano $B \subset X_2$. Più in generale, per ogni funzione Borel-misurabile e limitata $f : X_2 \to \mathbb{R}$, vale

$$\int_{X_2} f(y) \, d\mathbf{t}_\# \mu(y) = \int_{X_1} f(\mathbf{t}(x)) \, d\mu(x).$$

Si verifica facilmente che, per ogni coppia di misure $\mu, \nu \in \mathscr{P}(X_1)$, se $\mu \ll \nu$, allora $\mathbf{t}_{\#}\mu \ll \mathbf{t}_{\#}\nu$. Vale inoltre la seguente regola di composizione: date due applicazioni Borel-misurabili $\mathbf{t} : X_1 \to X_2$, $\mathbf{s} : X_2 \to X_3$,

$$(\mathbf{s} \circ \mathbf{t})_{\#}\mu = \mathbf{s}_{\#}(\mathbf{t}_{\#}\mu).$$

Indichiamo con π^i, $i = 1, 2$, gli operatori di proiezione sullo spazio prodotto $X := X_1 \times X_2$, definiti da

$$\begin{aligned} \pi^1 : X &\to X_1 \\ (x, y) &\mapsto x \end{aligned} \qquad\qquad \begin{aligned} \pi^2 : X &\to X_2 \\ (x, y) &\mapsto y. \end{aligned}$$

Data $\mu \in \mathscr{P}(X)$, le marginali di μ sono le misure di probabilità

$$\mu^1 := \pi^1_{\#}\mu \in \mathscr{P}(X_1), \qquad \mu^2 := \pi^2_{\#}\mu \in \mathscr{P}(X_2).$$

Date $\mu \in \mathscr{P}(X_1)$ e $\mu \in \mathscr{P}(X_2)$, la classe $\Gamma(\mu, \nu)$ dei piani di trasporto fra μ e ν è definita da

$$\Gamma(\mu, \nu) := \left\{\gamma \in \mathscr{P}(X) \ : \ \pi^1_{\#}\gamma = \mu, \quad \pi^2_{\#}\gamma = \nu\right\}.$$

Si noti che se μ (o ν) è una delta di Dirac in un punto $\bar{x}$, allora c'è un solo elemento nella classe dei piani di trasporto

$$\Gamma(\delta_{\bar{x}}, \nu) = \{\delta_{\bar{x}} \times \nu\}.$$

Sia $\mathbf{t} : X_1 \to X_2$ un applicazione misurabile, alla coppia di misure $\mu \in \mathscr{P}(X_1)$, $\nu = \mathbf{t}_{\#}\mu$, possiamo associare un particolare piano di trasporto $\gamma \in \Gamma(\mu, \nu)$, detto il piano indotto da $\mathbf{t}$:

$$\gamma := (id, \mathbf{t})_{\#}\mu,$$

dove id è l'identità su X_1. Rappresentando γ attraverso un integrale, si ha

$$\int_{X_1 \times X_2} f(x, y)\, d\gamma(x, y) = \int_{X_1} f(x, \mathbf{t}(x))\, d\mu(x),$$

per ogni $f : X_1 \times X_2 \to \mathbb{R}$ Borel-misurabile e limitata. Si noti che il supporto del piano indotto da $\mathbf{t}$ è concentrato sul grafico di $\mathbf{t}$ in $X_1 \times X_2$.

Esercizio 4.24
Sia Q il quadrato in $\mathbb{R}^2$ con vertici in $(0, a)$, $(a, 2a)$, $(2a, a)$, $(a, 0)$ per $a > 0$. Si consideri la misura $\mu \in \mathscr{P}(\mathbb{R}^2)$

$$\mu := \frac{1}{a\sqrt{2}} \chi_Q \mathcal{L}^2.$$

1. Calcolare le marginali di μ;
2. Trovare tre esempi di elementi di $\Gamma(\pi^1_{\#}\mu, \pi^2_{\#}\mu)$;
3. Per ciascuna delle misure γ_j al punto 2, scrivere esplicitamente $\int f(x, y)\, d\gamma_j$ e disegnare il supporto di γ_j.

□ □ □ □ □

Esercizio 4.25
Dimostrare che $\Gamma(\mu, \nu)$ è chiuso rispetto alla convergenza stretta.

□ □ □ □ □

4.3 Il problema di trasporto ottimo

Questa sezione è tratta da [3, Section 2.2]. Siano X e Y spazi metrici completi e separabili e sia $c : X \times Y \to [0, +\infty]$ una funzione Borel-misurabile. Date $\mu \in \mathscr{P}(X)$ e $\nu \in \mathscr{P}(Y)$, il problema di trasporto ottimo, nella formulazione di Monge, chiede di determinare

$$\inf \left\{ \int_X c(x, \mathbf{t}(x))\, d\mu(x) \ : \ \mathbf{t}_{\#}\mu = \nu \right\}, \tag{4.9}$$

dove l'estremo inferiore si intende su tutte le applicazioni misurabili $\mathbf{t} : X \to Y$ che trasportano μ in ν. Il problema, in questa formulazione, potrebbe non essere nemmeno ben posto, poiché se μ è una delta di Dirac e ν non lo è, non cè nessun trasporto misurabile fra μ e ν.

Esempio 4.26
Inoltre, la condizione $\mathbf{t}_{\#}\mu = \nu$ non è chiusa rispetto a nessuna topologia debole ragionevole. Per esempio, si consideri la successione $f_n(x) = f(nx)$, dove $f : \mathbb{R} \to \mathbb{R}$ è 1-periodica, $f(x) = 1$ se $x \in [0, 1/2[$, $f(x) = -1$ se $x \in [1/2, 1[$, e le misure $\mu := \chi_{[0,1]}\mathcal{L}^1$ e $\nu := \frac{1}{2}(\delta_1 + \delta_{-1})$. Si vede facilmente che $(f_n)_{\#}\mu = \nu$ per ogni $n \in \mathbb{N}$, che $f_n \rightharpoonup 0$ debolmente in $L^p(0, 1)$, e che $0_{\#}\mu = \delta_0 \neq \nu$.

□ □ □ □ □

Esercizio 4.27
Se $\mu = \delta_{\bar{x}}$ e $\nu = \delta_{\bar{y}}$, per certi $\bar{x} \in X$, $\bar{y} \in Y$, quanto vale l'inf?

□ □ □ □ □

La formulazione di Kantorovich del problema di trasporto ottimo chiede di determinare

$$\min \left\{ \int_{X\times Y} c(x, y)\, d\gamma(x, y) \ : \ \gamma \in \Gamma(\mu, \nu) \right\}. \tag{4.10}$$

Questo risolve il problema della buona posizione, dato che $\Gamma(\mu, \nu)$ contiene almeno un elemento (cioè $\mu \times \nu$). Inoltre, con il metodo diretto del calcolo delle variazioni, si può dimostrare che esiste un piano che realizza il minimo, non appena si suppone che la funzione costo c sia semicontinua inferiormente. Vediamo come: definiamo

$$J(\gamma) := \int_{X\times Y} c(x, y)\, d\gamma(x, y)$$

e poniamo

$$m := \inf_{\gamma\in\Gamma(\mu,\nu)} J(\gamma).$$

Poiché $c(x, y) \geq 0$, si ha $m \geq 0$. Sia $\{\gamma_n\}$ una successione minimizzante di piani di trasporto, cioè una successione tale che

$$\lim_{n\to\infty} J(\gamma_n) = m.$$

Grazie al teorema di tightness di Ulam e alle ipotesi su X e Y, μ e ν sono tight. Quindi l'insieme $\Gamma(\mu, \nu)$ è tight e quindi esistono una sottosuccessione $\{\gamma_{n_k}\}$ e un elemento

$\gamma \in \overline{\Gamma(\mu, \nu)}$ tali che γ_{n_k} converge strettamente a γ. Dato che le marginali μ, ν sono fissate, l'insieme $\Gamma(\mu, \nu)$ è anche chiuso rispetto alla convergenza stretta, e quindi $\gamma \in \Gamma(\mu, \nu)$. Grazie alla semicontinuità di c e a (4.6), otteniamo che

$$\liminf_{n \to \infty} J(\gamma_{n_k}) \geq J(\gamma).$$

Quindi $m \geq J(\gamma)$, e dunque γ realizza il minimo di J. Notazione: indichiamo con $\Gamma_o(\mu, \nu)$ l'insieme dei piani di trasporto che realizzano il minimo in (4.10).

La problema di Kantorovich si può considerare una formulazione debole di quello di Monge, nel senso che se c è limitata e continua e c non ha atomi, allora il "min" di (4.10) coincide con l'inf di (4.9).

Nel caso particolare $X = Y$ e $c(x, y) = d(x, y)^2$, dove d è la distanza in X, il problema (4.10) viene usato per definire la distanza di Kantorovich-Rubinstein-Wasserstein:

$$W_2(\mu, \nu) := \left(\int_{X \times X} |x - y|^2 \, d\gamma(x, y) \right), \qquad \gamma \in \Gamma_o(\mu, \nu). \tag{4.11}$$

Teorema 4.28. *Sia X uno spazio metrico completo e separabile. Allora W_2 definisce una distanza in $\mathscr{P}_2(X)$, e $(\mathscr{P}_2(X), W_2)$ è uno spazio metrico completo e separabile. Inoltre, data una successione $\{\mu_n\} \subset \mathscr{P}_2(X)$, vale*

$$\lim_{n \to \infty} W_2(\mu_n, \mu) = 0 \qquad \Leftrightarrow \begin{cases} \mu_n \rightharpoonup \mu \text{ strettamente} \\ \{\mu_n\} \text{ ha momenti secondi uniformemente integrabili} \end{cases}$$

Qui dimostreremo solo che W_2 è una distanza. La dimostrazione dell'enunciato completo si può trovare su [2, Proposition 7.1.5]. Premettiamo due risultati utili.

Teorema 4.29 (Disintegrazione di misure)**.** *Siano X, Y spazi metrici separabili e localmente compatti e sia $\pi : X \to Y$ una funzione Borel-misurabile. Siano $\mu \in \mathscr{P}(X)$, $\nu := \pi_\# \mu \in \mathscr{P}(Y)$. Allora, esiste una famiglia di misure $\{\mu_y\}$ tali che*

(i) $y \mapsto \mu_y$ è Borel-misurabile e $\mu_y \in \mathscr{P}(X)$ per ν-q.o y,

(ii) $\mu = \mu_y \otimes \nu$, cioè:

$$\mu(A) = \int_Y \mu_y(A) \, d\nu(y), \qquad \int_X f(x) \, d\mu(x) = \int_Y \left(\int_{\pi^{-1}(y)} f(x) \, d\mu_y(x) \right) d\nu(y)$$

(iii) $\mu_y(X \backslash \pi^{-1}(y)) = 0$ per ν-q.o. $y \in Y$.

Il teorema vale più in generale se $\mu \in [\mathcal{M}(X)]^m$ e $\mu = \pi_\# |\mu|$. Inoltre, se sostituiamo $X \times Y$ al posto di X nell'enunciato, cioè se abbiamo $\pi : X \times Y \to Y$, $\mu \in \mathscr{P}(X \times Y)$, $\nu := \pi_\# \mu \in \mathscr{P}(Y)$, possiamo identificare ciascuna fibra $\pi^{-1}(y)$ con X stesso, trovare $\mu_y \in \mathscr{P}(X)$, e decomporre

$$\int_{X \times Y} f(x, y) \, d\mu(x, y) = \int_Y \left(\int_X f(x, y) \, d\mu_y(x) \right) d\nu(y).$$

Il risultato appare probabilmente più chiaro in questa seconda forma, ma può essere utile conoscere la versione enunciata, che è più generale.

Lemma 4.30. *Siano X, Y, Z tre spazi metrici separabili e completi e siano $\gamma_1 \in \mathscr{P}(X \times Y)$, $\gamma_2 \in \mathscr{P}(Y \times Z)$ tali che $\pi^Y_\# \gamma_1 = \pi^Y_\# \gamma_2$. Allora esiste una misura $\gamma \in \mathscr{P}(X \times Y \times Z)$ tale che*

$$\pi^{X\times Y}_\# \gamma = \gamma_1, \qquad \pi^{Y\times Z}_\# \gamma = \gamma_2.$$

Proof. Sia $\nu := \pi^Y_\# \gamma_1 = \pi^Y_\# \gamma_2$, allora, grazie al Teorema di Decomposizione esistono $(\gamma_1)_y \in \mathscr{P}(X), (\gamma_2)_y \in \mathscr{P}(Z)$ tali che

$$\gamma_1 = (\gamma_1)_y \otimes \nu, \qquad \gamma_2 = (\gamma_2)_y \otimes \nu.$$

La tesi è soddisfatta dalla misura $\gamma := ((\gamma_1)_y \times (\gamma_2)_y) \otimes \nu \in \mathscr{P}(X \times Y \times Z)$, che verifica

$$\int_{X\times Y\times Z} f(x,y,z)\, d\gamma(x,y,z) = \int_Y \left(\int_X \int_Z f(x,y,z) d\gamma_2(z)\, d\gamma_1(x) \right) d\nu(y).$$

□

Dimostriamo ora che W_2 è una distanza su $\mathscr{P}_2(X)$.

Proof – Teorema 4.28. Dalla definizione, si vede che $W_2(\mu,\nu) = W_2(\nu,\mu)$. Se $\mu = \nu$, come piano si può scegliere $(id, id)_\# \mu$ (quello indotto da $\mathbf{t} = id$), e

$$W_2(\mu,\nu) \leq \int_{X\times X} d(x, id(x))^2 d\mu(x) = \int_{X\times X} d(x,x)^2 d\mu(x) = 0.$$

Per dimostrare che $W_2(\mu,\nu) = 0 \Rightarrow \mu = \nu$, si scelga un piano γ ottimale. Se

$$\int_{X\times X} d(x,y)^2 d\gamma = 0,$$

allora γ è concentrata sulla diagonale $\{y = x\} \subset X \times X$, quindi

$$\mu = \pi^1_\# \gamma = \pi^2_\# \gamma = \nu.$$

Siano ora $\mu, \nu, \sigma \in \mathscr{P}_2(X)$, vogliamo dimostrare la disuguaglianza triangolare, cioè che

$$W_2(\mu,\sigma) \leq W_2(\mu,\nu) + W_2(\nu,\sigma).$$

Siano $\gamma \in \Gamma_o(\mu,\nu)$ e $\eta \in \Gamma_o(\nu,\sigma)$. Dato che

$$\pi^2_\# \gamma = \nu = \pi^1_\# \eta,$$

per il Lemma 4.30, esiste una misura $\lambda \in \mathscr{P}(X \times X \times X)$ tale che

$$(\pi^1, \pi^2)_\# \lambda = \gamma, \qquad (\pi^2, \pi^3)_\# \lambda = \eta.$$

Allora $(\pi^1, \pi^3)_\# \lambda \in \Gamma(\mu,\sigma)$, infatti

$$\pi^1_\# (\pi^1, \pi^3)_\# \lambda = \pi^1_\# \lambda = \pi^1_\# \gamma = \mu, \qquad \pi^3_\# (\pi^1, \pi^3)_\# \lambda = \pi^3_\# \lambda = \pi^3_\# \eta = \sigma.$$

(o anche, ragionando sulla misura di un insieme $A \subset X$,

$$\mu(A) = \gamma(A \times X) = \lambda(A \times X \times X) = \pi^1_\# \lambda(A)$$
$$\sigma(A) = \eta(X \times A) = \lambda(X \times X \times A) = \pi^3_\# \lambda(A).\ )$$

Quindi

$$\begin{aligned}
W_2(\mu,\sigma) &\le \left(\int_{X\times X} d(x,z)^2\, d(\pi^1,\pi^3)_\#\lambda(x,z)\right)^{\frac12}\\
&= \left(\int_{X\times X\times X} d(x,z)^2\, d\lambda(x,y,z)\right)^{\frac12}\\
&\le \left(\int_{X\times X\times X} \big(d(x,y)+d(y,z)\big)^2\, d\lambda(x,y,z)\right)^{\frac12}\\
&\le \left(\int_{X\times X\times X} d(x,y)^2\, d\lambda(x,y,z)\right)^{\frac12} + \left(\int_{X\times X\times X} d(y,z)^2\, d\lambda(x,y,z)\right)^{\frac12}\\
&= \left(\int_{X\times X} d(x,y)^2\, d\gamma(x,y)\right)^{\frac12} + \left(\int_{X\times X} d(y,z)^2\, d\eta(y,z)\right)^{\frac12}\\
&= W_2(\mu,\nu)+W_2(\nu,\sigma).
\end{aligned}$$

(Si noti che nel quarto passaggio si è usata la disuguaglianza triangolare per la norma di $d(x,y)+d(y,z)$ in $L^2(\lambda)$.) Infine, dimostriamo che $W_2(\mu,\nu) < +\infty$. Grazie alla disuguaglianza triangolare si ha

$$\begin{aligned}
W_2(\mu,\nu) &\le W_2(\mu,\delta_{x_0})+W_2(\delta_{x_0},\nu)\\
&= \left(\int_X d(x,x_0)^2\, d\mu(x)\right)^{\frac12} + \left(\int_X d(x,x_0)^2\, d\nu(x)\right)^{\frac12},
\end{aligned}$$

e avendo supposto che $\mu,\nu \in \mathscr{P}_2(X)$, i momenti secondi di μ,ν sono finiti. $\square$

Anche se in molti casi i piani ottimali sono indotti da una funzione $\mathbf{t}$, la formulazione di Kantorovich rimane utile per maggiorare W_2. Per esempio, siano $\mu,\nu,\sigma \in P_2(X)$, supponiamo che $\mathbf{t}_\#\sigma=\mu$, $\mathbf{s}_\#\sigma=\nu$, per certe funzioni $\mathbf{t},\mathbf{s}: X\to X$, allora

$$W_2(\mu,\nu) \le \int_X d(\mathbf{t}(x),\mathbf{s}(x))^2\, d\sigma(x),$$

dato che $(\mathbf{t},\mathbf{s})_\#\sigma \in \Gamma(\mu,\nu)$ e che

$$\int_{X\times X} d(x,y)^2\, d(\mathbf{t},\mathbf{s})_\#\sigma(x,y) = \int_X d(\mathbf{t}(x),\mathbf{s}(x))^2\, d\sigma(x).$$

Il seguente risultato fornisce esistenza e unicità della mappa $\mathbf{t}$, usata per trasportare μ in $\nu=\mathbf{t}_\#\mu$ nella formulazione di Monge del problema di trasporto ottimo. L'ipotesi cruciale è che la misura μ di partenza sia in $\mathscr{P}_2^a(\mathbb{R}^m)$. Ricordiamo la definizione:

$$\mathscr{P}_2^a(\mathbb{R}^m) := \left\{\mu\in\mathscr{P}(\mathbb{R}^m): \int_{\mathbb{R}^m}|x|^2 d\mu < +\infty, \quad \mu \ll \mathscr{L}^m\right\}.$$

Dato che $\mu \ll \mathscr{L}^m$, esiste una densità nonnegativa $\rho\in L^1(\mathbb{R}^m)$ tale che $\mu=\rho\mathcal{L}^m$. La definizione di $\mathscr{P}_2^a$ si può quindi riscrivere come

$$\mathscr{P}_2^a(\mathbb{R}^m) := \left\{\mu=\rho\mathcal{L}^m : \rho\in L^1(\mathbb{R}^m), \quad \rho(x)\ge 0 \text{ q.o.}, \quad \int_{\mathbb{R}^m}|x|^2\rho(x)dx<+\infty\right\}.$$

Teorema 4.31. *Per ogni $\mu \in \mathscr{P}_2^a(\mathbb{R}^m)$, $\nu \in \mathscr{P}_2(\mathbb{R}^m)$ il problema di trasporto ottimo di Kantorovich (4.10) con $c(x,y) = |x-y|^2$ ha un'unica soluzione $\gamma \in \Gamma_o(\mu,\nu)$. Inoltre*

(i) *γ è indotto da una mappa di trasporto $\mathbf{t} : \mathbb{R}^m \to \mathbb{R}^m$, ovvero, $\gamma = (id, \mathbf{t})_{\#}\mu$. In particolare, $\mathbf{t}$ è l'unica soluzione del problema di trasporto ottimo di Monge (4.9).*

(ii) *La mappa $\mathbf{t}$ coincide μ-q.o. con il gradiente di una funzione convessa $\varphi : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$, il cui dominio $D(\varphi)$ ha interno non vuoto e*

$$\mu(\mathbb{R}^m \backslash D(\varphi)) = \mu(\mathbb{R}^m \backslash D(\nabla\varphi)) = 0.$$

(iii) *Se anche $\nu \in \mathscr{P}_2^a(\mathbb{R}^m)$, indicando con $\mathbf{s}$ la mappa di trasporto ottimo fra ν e μ, allora*

$$\mathbf{s} \circ \mathbf{t} = id \quad \text{-q.o.,} \qquad \mathbf{t} \circ \mathbf{s} = id \quad \text{-q.o.}$$

Infine, indicando con ρ, ρ' le densità di μ e ν, si ha (quasi ovunque)

$$\rho' = \frac{\rho}{\det \nabla^2 \varphi} \circ \mathbf{t}^{-1}$$

4.4 Curve assolutamente continue e derivata metrica

Sia (X,d) uno spazio metrico, siano $I \subset \mathbb{R}$ un intervallo aperto e $u : I \to X$ una curva in X. Diciamo che u è *p-assolutamente continua* se esiste $m \in L^p(I)$, $p \in [1,+\infty]$, tale che

$$d(u(s), u(t)) \leq \int_s^t m(\tau)\, d\tau \qquad \forall\, s,t \in I,\ s \leq t. \tag{4.12}$$

Indichiamo tali curve con $\mathrm{AC}^p(I;X)$. Nel caso $p = 1$ indichiamo le curve assolutamente continue semplicemente con $\mathrm{AC}(I;X)$. Ogni curva assolutamente continua è anche uniformemente continua e si può quindi estendere alla chiusura dell'intervallo I. Diciamo che $u : I \to X$ ammette *derivata metrica* in $t \in I$ se esiste, finito

$$|u|'(t) := \lim_{h\to 0} \frac{d(u(t+h), u(t))}{|h|}. \tag{4.13}$$

Teorema 4.32. *Sia $p \in [1,+\infty]$. Se $u \in \mathrm{AC}^p(I;X)$, allora u ammette derivata metrica per q.o. $t \in I$. Inoltre, la funzione $t \mapsto |u|'(t)$*

(i) *appartiene a $L^p(I)$,*

(ii) *verifica*

$$d(u(s), u(t)) \leq \int_s^t |u|'(\tau)\, d\tau \qquad \forall\, s,t \in I,\ s \leq t.$$

(ii) *è minimale, nel senso che*

$$|u|'(t) \leq m(t)$$

per ogni m che soddisfi (4.12).

4.5 Geodetiche in $\mathscr{P}_2(\mathbb{R}^m)$

Sia (X,d) uno spazio metrico. Una geodetica parametrizzata a velocità costante è una funzione $v:[0,T]\to X$ tale che

$$d(v(s),v(t))=\frac{t-s}{T}d(v(0),v(T)) \qquad \text{per ogni } 0\le s\le t\le T.$$

L'idea è che per ogni intervallo $[s,t]\subset[0,T]$ la lunghezza della curva coincida con la distanza fra gli estremi e che, ancora per ogni intervallo, la velocità media della curva su $[s,t]$ sia costante:

$$\frac{d(v(s),v(t))}{t-s}=const=\frac{d(v(0),v(T))}{T-0} \qquad \text{per ogni } 0\le s\le t\le T.$$

Nella definizione, in realtà, è sufficiente controllare che valga il $\le$. Infatti, supponiamo che

$$d(v(s),v(t))\le\frac{t-s}{T}d(v(0),v(T)) \qquad \text{per ogni } 0\le s\le t\le T$$

e che esistano s,t tali che

$$d(v(s),v(t))<\frac{t-s}{T}d(v(0),v(T)),$$

allora, per la disuguaglianza triangolare,

$$\begin{aligned}d(v(0),v(T))&\le d(v(0),v(s))+d(v(s),v(t))+d(v(t),v(T))\\&<\frac{s}{T}d(v(0),v(T))+\frac{t-s}{T}d(v(0),v(T))+\frac{T-t}{T}d(v(0),v(T)) \quad =d(v(0),v(T)),\end{aligned}$$

assurdo. Usando questa osservazione, possiamo costruire una geodetica a velocità costante in $(\mathscr{P}_2(X),W_2)$, dove $X=\mathbb{R}^m$. Date $\mu,\nu\in\mathscr{P}_2(X)$, si tratta di costruire una curva $\mu:[0,1]\to\mathscr{P}_2(X)$ tale che

$$\mu(0)=\mu,\quad \mu(1)=\nu,\quad W_2(\mu(s),\mu(t))\le(t-s)W_2(\mu,\nu) \quad \text{per ogni } 0\le s\le t\le T. \tag{4.14}$$

Prendiamo allora $\gamma\in\Gamma_o(\mu,\nu)$ e definiamo

$$\mu(t):=\left(\pi^1+t\left(\pi^2-\pi^1\right)\right)_{\#}\gamma. \tag{4.15}$$

Controlliamo ora che $\mu(t)$ verifichi (4.14). Si vede subito che

$$\mu(0)=\left(\pi^1\right)_{\#}\gamma=\mu, \qquad \mu(1)=\left(\pi^2\right)_{\#}\gamma=\nu.$$

Per comodità di notazione, per $t\in[0,1]$, definiamo

$$\begin{aligned}f_t: X\times X &\to X\\ (x,y) &\mapsto x+t(y-x).\end{aligned}$$

Allora, $\mu(t)=(f_t)_{\#}\gamma$. Definiamo $\gamma_{st}:=(f_s,f_t)_{\#}\gamma\in\mathscr{P}(X\times X)$. Testando con una funzione ϕ che dipende solo dalla prima variabile si vede

$$\int_{X\times X}\phi(x)\,d\gamma_{st}(x,y)=\int_{X\times X}\phi(f_s(x,y))\,d\gamma(x,y)=\int_X\phi(x)\,d(f_s)_{\#}\gamma(x)$$

e quindi $\pi^1_{\#}\gamma_{st} = (f_s)_{\#}\gamma = \mu(s)$. Analogamente, si vede che $\pi^2_{\#}\gamma_{st} = (f_t)_{\#}\gamma = \mu(t)$, e quindi $\gamma_{st} \in \Gamma(\mu(s), \mu(t))$, cioè è un piano ammissibile per calcolare la distanza fra $\mu(s)$ e $\mu(t)$. Dato che non sappiamo se sia anche ottimale, vale la disuguaglianza

$$\begin{aligned}
W_2(\mu(s), \mu(t)) &\leq \left(\int_{X\times X} |x-y|^2 d\gamma_{st}(x,y) \right)^{\frac{1}{2}} \\
&= \left(\int_{X\times X} |f_s(x,y) - f_t(x,y)|^2 d\gamma(x,y) \right)^{\frac{1}{2}} \\
&= \left(\int_{X\times X} |x + s(y-x) - x - t(y-x)|^2 d\gamma(x,y) \right)^{\frac{1}{2}} \\
&= \left(\int_{X\times X} |(s-t)(y-x)|^2 d\gamma(x,y) \right)^{\frac{1}{2}} \\
&= |s-t| \left(\int_{X\times X} |y-x|^2 d\gamma(x,y) \right)^{\frac{1}{2}} \\
&= (t-s) W_2(\mu, \nu),
\end{aligned}$$

dove, nell'ultimo passaggio, si è usato che γ sia un piano ottimale, e quindi realizzi la distanza fra le sue marginali μ e ν. Abbiam così fatto vedere che la curva definita in (4.15) è una geodetica a velocità costante in $\mathscr{P}_2(\mathbb{R}^m)$.

4.6 L'equazione di continuità

In questa sezione raccogliamo alcuni risultati sulla cosiddetta equazione di continuità

$$\partial_t \mu_t + \operatorname{div}(\mathbf{v}_t \mu_t) = 0 \qquad \in \mathbb{R}^m \times (0,T). \tag{4.16}$$

Notazione e significato: qui $\mu : (0,T) \to \mathscr{P}(\mathbb{R}^m)$ è una curva di misure di probabilità. È notazione comune indicare $\mu_t = \mu(t)$. Chiediamo che il campo di vettori $\mathbf{v} : \mathbb{R}^m \times (0,T) \to \mathbb{R}^m$ sia Borel-misurabile. Analogamente, scriviamo $\mathbf{v}_t(x) = \mathbf{v}(x,t)$ (per sottolineare la funzione svolta in $\mathbb{R}^m$ da μ_t e $\mathbf{v}_t$, che pensiamo parametrizzati da $t \in (0,T)$). Chiediamo inoltre che $\mathbf{v}$ soddisfi

$$\int_0^T \int_{\mathbb{R}^m} |\mathbf{v}_t(x)| \, d\mu_t(x) \, dt < +\infty, \tag{4.17}$$

e che l'equazione (4.16) sia soddisfatta nel senso delle distribuzioni $\mathcal{D}'(\mathbb{R}^m \times (0,T))$, ovvero che

$$\int_0^T \partial_t \phi(x,t) + \nabla_x \phi(x,t) \cdot \mathbf{v}_t(x) \, d\mu_t(x) \, dt = 0 \qquad \forall \, \phi \in \mathcal{D}(\mathbb{R}^m \times (0,T)).$$

(In realtà, dato che compaiono solo derivate del primo ordine, basta testare con funzioni in $C^1_c(\mathbb{R}^m \times (0,T))$. Per dimostrare esistenza, unicità e una formula di rappresentazione per l'equazione di continuità, ricordiamo un risultato di ODE.

Lemma 4.33 (Il sistema caratteristico di ODE)**.** *Sia $\mathbf{v}_t$ un campo vettoriale Borel-misurabile tale che per ogni compatto $B \subset \mathbb{R}^m$*

$$\int_0^T \left(\sup_B |\mathbf{v}_t| + \operatorname{Lip}(\mathbf{v}_t, B) \right) dt < +\infty. \tag{4.18}$$

Allora, per ogni $x \in \mathbb{R}^m$, s $in[0,T]$ esiste un'unica soluzione massimale $X : [0,T] \to \mathbb{R}^m$ del problema di Cauchy

$$\begin{cases} \frac{d}{dt} X_t = \mathbf{v}_t(X_t) \\ X_s = x, \end{cases} \tag{4.19}$$

definita in un intervallo I relativamente aperto in $[0,T]$ e contenente s.

Notazione: indichiamo la dipendenza dal tempo con $X_t = X(t)$, il dato iniziale è al generico tempo $t = s$, anziché $t = 0$. Inoltre, per indicare la dipendenza della soluzione dal dato x al tempo s, denotiamo $X_t = X_t(x,s)$ e per indicare la dipendenza dell'intervallo massimale I dal dato x al tempo s, denotiamo $I = I(x,s)$. Indichiamo con $\tau(x)$ la lunghezza dell'intervallo massimale $I(x,0)$.

Lemma 4.34. *Sia $\mathbf{v}_t$ un campo di vettori Borel-misurabile che soddisfi (4.18). Sia $\mu_0 \in \mathscr{P}(\mathbb{R}^m)$ e sia X_t la soluzione del problema (4.19) con dato $(x,0)$. Supponiamo che per qualche $\bar{t} \in (0,T]$*

$$\tau(x) > \bar{t} \qquad \text{per } \mu_0\text{-q.o. } x \in \mathbb{R}^m.$$

Definiamo $\mu_t := (X_t)_\# \mu_0$. Se $\mathbf{v}_t$ soddisfa (4.17), allora $t \mapsto \mu_t$ è una soluzione continua dell'equazione di continuità (4.16).

Proposizione 4.35. *Viceversa, se μ_t è una soluzione (strettamente) continua di (4.16), rispetto a un campo di vettori $\mathbf{v}_t$ Borel-misurabile che soddisfa (4.17) e (4.18), allora per μ_0-q.o. $x \in \mathbb{R}^m$ il problema caratteristico (4.19) ha una soluzione globale X_t in $[0,T]$, e vale*

$$\mu_t = (X_t)_\# \mu_0$$

Se inoltre

$$\int_0^T \int_{\mathbb{R}^m} |\mathbf{v}_t(x)|^2 \, d\mu_t(x) \, dt < +\infty, \tag{4.20}$$

allora il campo $\mathbf{v}_t$ è la derivata temporale di X_t in L^2, ovvero

$$\lim_{h \to 0} \int_0^{T-h} \int_{\mathbb{R}^m} \left| \frac{X_{t+h}(x,0) - X_t(x,0)}{h} - \mathbf{v}_t(X_t(x,0)) \right|^2 d\mu_0(x) \, dt \tag{4.21}$$

Nota: (4.20) significa che $\|\mathbf{v}_t\|_{L^2(\mu_t;\mathbb{R}^m)} \in L^1(0,T)$. Colleghiamo adesso queste nozioni con quelle di curva assolutamente continua e di derivata metrica. Il campo di vettori $\mathbf{v}_t$ gioca il ruolo di vettore tangente alla curva $t \mapsto \mu_t$.

Teorema 4.36. *Sia I un intervallo aperto di $\mathbb{R}$, sia $\mu : I \mapsto \mathscr{P}_2(\mathbb{R}^m)$ una curva assolutamente continua e sia $|\mu'| \in L^1(I)$ la sua derivata metrica. Allora esiste un campo di vettori Borel-misurabile $\mathbf{v} : \mathbb{R}^m \times I \to \mathbb{R}^m$ tale che*

$$\mathbf{v}_t \in L^2(\mu_t;\mathbb{R}^m), \qquad \|\mathbf{v}_t\|_{L^2(\mu_t;\mathbb{R}^m)} \leq |\mu'|(t) \quad \text{per q.o. } t \in I,$$

e l'equazione di continuità

$$\partial_t \mu_t + \operatorname{div}(\mathbf{v}_t \mu_t) = 0$$

è verificata nel senso delle distribuzioni in $\mathbb{R}^m \times I$.

Viceversa, se una curva strettamente continua $\mu : I \to \mathscr{P}_2(\mathbb{R}^m)$ soddisfa l'equazione di continuità per qualche campo di vettori Borel-misurabile $\mathbf{w}_t$, tale che

$$\|\mathbf{w}_t\|_{L^2(\mu_t;\mathbb{R}^m)} \in L^1(I),$$

allora

- *μ è assolutamente continua,*
- *$|\mu'|(t) \leq \|\mathbf{w}_t\|_{L^2(\mu_t;\mathbb{R}^m)}$.*

In particolare,

$$\|\mathbf{v}_t\|_{L^2(\mu_t;\mathbb{R}^m)} = |\mu'|(t) \quad \textit{per q.o. } t \in I.$$

Note, la stessa curva μ_t può risolvere l'equazione di continuità per diversi vettori, però ce n'è uno solo che in norma è più piccolo della derivata metrica. O, equivalentemente, ce n'è uno solo che ha norma = alla derivata metrica. (Come per la coniugata, visto che l'altra disuguaglianza vale sempre, posso chiedere una disuguaglianza al posto dell'uguale.)

Non lo stiamo dimostrando, ma l'idea sottostante è che $\mathbf{v}_t$ è il vettore tangente a μ_t (e si può anche caratterizzare lo spazio tangente a μ).

4.7 Funzionali convessi in $\mathscr{P}_2(\mathbb{R}^m)$

Sia $\Phi : \mathscr{P}_2(\mathbb{R}^m) \to \mathbb{R} \cup +\{\infty\}$ un funzionale definito sullo spazio di misure di probabilità con momento secondo finito. La nozione di convessità più utile in questo ambiente è quella di convessità lungo le geodetiche, che sostituiscono l'usuale segmento $\theta x + (1-\theta)y \in \mathbb{R}^m$. Date $\mu_1, \mu_2 \in \mathscr{P}_2(\mathbb{R}^m)$, indichiamo con $\mu : [0,1] \to \mathscr{P}_2(\mathbb{R}^m)$ una geodetica che le unisce, secondo la definizione (4.15). Diciamo che ϕ è *convesso lungo le geodetiche* se per ogni coppia di misure $\mu_1, \mu_2 \in \mathscr{P}_2(\mathbb{R}^m)$ la funzione $t \mapsto \phi(\mu(t)$ è convessa, cioè se

$$\phi(\mu(t)) \leq (1-t)\phi(\mu_1) + t\phi(\mu_2) \qquad \forall\, t \in [0,1].$$

Una nozione più generale, usata in questo contesto, è quella di λ-convessità: Diciamo che ϕ è λ-*convesso lungo le geodetiche* se esiste $\lambda \in \mathbb{R}$ tale che per ogni coppia di misure $\mu_1, \mu_2 \in \mathscr{P}_2(\mathbb{R}^m)$

$$\phi(\mu(t)) \leq (1-t)\phi(\mu_1) + t\phi(\mu_2) - \frac{\lambda}{2}t(1-t)W_2^2(\mu_1,\mu_2) \qquad \forall\, t \in [0,1].$$

In $\mathbb{R}^m$, per una funzione $\phi \in C^2$,

$$\phi \text{ è } \lambda\text{-convesso} \qquad \Rightarrow \qquad H\phi - \lambda Id \text{ è semidefinita positiva.}$$

(E, analogamente, se la matrice Hessiana di ϕ meno λ volte la matrice identità è definita positiva, allora ϕ è λ-convessa.)

Esercizio 4.37
Siano $v_0 \in \mathbb{R}^m$ e $\phi : \mathbb{R}^m \to \mathbb{R}$ la funzione definita da $\phi(v) = \frac{1}{2}|v - v_0|^2$. Determinare il massimo dei valori di λ per cui ϕ è convessa.

Un esempio di funzionale λ-convesso in $\mathscr{P}_2(\mathbb{R}^m)$ è la cosiddetta energia potenziale: sia $V : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ una funzione propria, semicontinua inferiormente, la cui parte negativa abbia crescita quadratica:

$$V(x) \geq -A - B|x|^2 \tag{4.22}$$

per qualche $A, B \geq 0$. Costruiamo il funzionale $\mathcal{V} : \mathscr{P}_2(\mathbb{R}^m) \to \mathbb{R} \cup \{+\infty\}$

$$\mathcal{V}(\mu) := \int_{\mathbb{R}^m} V(x)\, d\mu(x).$$

Dato che V è proprio, esiste $x_0 \in \mathbb{R}^m$ tale che $V(x_0) < +\infty$, allora $\mathcal{V}(\delta_{x_0}) = V(x_0) < +\infty$, e quindi $\mathcal{V}$ è proprio. Ora, se $\mu_n \to \mu$ in $\mathscr{P}_2(\mathbb{R}^m)$ (si veda il Teorema 4.28), i momenti secondi di μ_n sono uniformemente integrabili, quindi per (4.22) V^- è uniformemente integrabile rispetto a $\{\mu_n\}$, e quindi per (4.6) $\mathcal{V}$ è semicontinuo inferiormente. Dimostriamo ora che se V è λ-convessa, allora $\mathcal{V}$ è λ-convesso: siano $\mu_1, \mu_2 \in D(\mathcal{V})$ e sia μ_t una geodetica che le unisce:

$$\mu_t := (\pi^1 + t(\pi^2 - \pi^1))_{\#}\gamma, \qquad \gamma \in \Gamma_o(\mu_1, \mu_2).$$

Allora:

$$\begin{aligned}
\mathcal{V}(\mu_t) &= \int_{\mathbb{R}^m} V(x)\, d\mu_t(x) \\
&= \int_{\mathbb{R}^m \times \mathbb{R}^m} V(x + t(y-x))\, d\gamma(x,y) \\
&\leq \int_{\mathbb{R}^m \times \mathbb{R}^m} (1-t)V(x) + tV(y) - \frac{\lambda}{2} t(1-t)|x-y|^2\, d\gamma(x,y) \\
&= (1-t)\int_{\mathbb{R}^m} V(x)\, d\mu_1(x) + t\int_{\mathbb{R}^m} V(y)\, d\mu_2(y) - \frac{\lambda}{2} t(1-t)\int_{\mathbb{R}^m \times \mathbb{R}^m} |x-y|^2\, d\gamma(x,y) \\
&= (1-t)\mathcal{V}(\mu_1) + t\mathcal{V}(\mu_2) - \frac{\lambda}{2} t(1-t) W_2^2(\mu_1, \mu_2).
\end{aligned}$$

Un altro esempio è dato dalla cosiddetta Energia interna: sia $F : [0, +\infty) \to \mathbb{R} \cup \{+\infty\}$ una funzione propria, semicontinua inferiormente, e convessa, tale che

$$F(0) = 0, \qquad \liminf_{s \to 0} \frac{F(s)}{s^\alpha} > -\infty, \quad \text{per qualche } \alpha > \frac{m}{m+2}.$$

(Quest'ultima condizione serve a garantire che la parte negativa di $F(u(x))$ sia integrabile in $\mathbb{R}^m$.) Definiamo il funzionale $\mathcal{F} : \mathscr{P}_2(\mathbb{R}^m) \to \mathbb{R} \cup \{+\infty\}$

$$\mathcal{F}(\mu) := \begin{cases} \displaystyle\int_{\mathbb{R}^m} F(u(x))\, dx & \text{se } \mu \ll \mathcal{L}^m, \\ +\infty & \text{altrimenti,} \end{cases}$$

dove u è la parte assolutamente continua di μ rispetto alla misura di Lebesgue $\mathcal{L}^m$. Se F ha crescita superlineare a $+\infty$ e

$$\text{la funzione} \quad s \mapsto s^m F(s^{-m}) \quad \text{è convessa e non crescente in } (0, +\infty), \tag{4.23}$$

allora $\mathcal{F}$ è convesso lungo le geodetiche (per la dimostrazione, si veda [3, Proposition 3.11]). Per esempio, la funzione

$$F(s) = s \log s$$

soddisfa la condizione (4.23).

Un caso interessante di energia interna è l'Entropia relativa: siano $\mu, \nu \in \mathscr{P}(\mathbb{R}^m)$, l'entropia relativa di μ rispetto a ν è

$$\mathcal{H}(\mu|\nu) := \begin{cases} \displaystyle\int_{\mathbb{R}^m} \frac{d\mu}{d\nu} \log\left(\frac{d\mu}{d\nu}\right) d\nu & \text{se } \mu \ll \nu, \\ +\infty & \text{altrimenti.} \end{cases}$$

Introduciamo la funzione non negativa, semicontinua inferiormente e convessa

$$H(s) := \begin{cases} s(\log(s) - 1) + 1 & \text{se } s > 0, \\ 1 & \text{se } s = 0, \\ 0 & \text{altrimenti,} \end{cases}$$

e osserviamo che, per $\mu \ll \nu$,

$$\begin{aligned} \int_{\mathbb{R}^m} H\left(\frac{d\mu}{d\nu}\right) d\nu &= \int_{\mathbb{R}^m} \frac{d\mu}{d\nu} \log\left(\frac{d\mu}{d\nu}\right) - \frac{d\mu}{d\nu} + 1\, d\nu \\ &= \int_{\mathbb{R}^m} \frac{d\mu}{d\nu} \log\left(\frac{d\mu}{d\nu}\right) d\nu - \int_{\mathbb{R}^m} 1\, d\mu + \int_{\mathbb{R}^m} 1\, d\nu \\ &= \mathcal{H}(\mu|\nu). \end{aligned}$$

Dato che $H \geq 0$, questo mostra che $\mathcal{H} \geq 0$ e $\mathcal{H} = 0 \Leftrightarrow \mu = \nu$. Infine, sia γ è una misura di Borel su $\mathbb{R}^m$ (non necessariamente finita), e sia $V : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ è una funzione misurabile tale che

$$V^+ \text{ ha crescita al più quadratica}, \qquad \tilde{\gamma} := e^{-V}\gamma \in \mathscr{P}(\mathbb{R}^m).$$

Allora, per misure $\mu \in \mathscr{P}_2(\mathbb{R}^m)$ l'entropia relativa rispetto a γ è ben definita dalla formula

$$\mathcal{H}(\mu|\gamma) := \mathcal{H}(\mu|\tilde{\gamma}) - \int_{\mathbb{R}^m} V(x)\, d\mu(x).$$

In particolare, quando γ è la misura di Lebesgue m-dimensionale, e $\mu = \rho\mathcal{L}^m$, questa formula si riduce a

$$\mathcal{H}(\mu|\mathcal{L}^m) := \mathcal{H}(\mu|\tilde{\gamma}) - \int_{\mathbb{R}^m} V(x)\, d\mu(x).$$

e ritroviamo il solito funzionale di entropia introdotto in (1.19):

$$\gamma = \mathcal{L}^m, \quad \tilde{\gamma} = e^{-V}\mathcal{L}^m, \quad \mu = \rho\mathcal{L}^m = (\rho e^V)e^{-V}\mathcal{L}^m,$$

$$\begin{aligned} \mathcal{H}(\mu|\tilde{\gamma}) &= \int_{\mathbb{R}^m} (\rho e^V) \log(\rho e^V) e^{-V}\, dx \\ &= \int_{\mathbb{R}^m} \rho\big(\log(\rho) + V\big)\, dx \\ &= \int_{\mathbb{R}^m} \rho \log(\rho)\, dx + \int_{\mathbb{R}^m} V(x)\rho(x)\, dx. \end{aligned}$$

Esercizio 4.38

La distanza di Wasserstein al quadrato, in dimensione $m \geq 2$, non è λ-convessa lungo le geodetiche. Sia $\mu_0 \in \mathscr{P}_2(\mathbb{R}^2)$ una misura fissata, definiamo

$$\Phi(\mu) := W_2^2(\mu, \mu_0).$$

1. Siano
$$\mu_1 := \frac{1}{2}(\delta_{(0,0)} + \delta_{(2,1)}), \qquad \mu_2 := \frac{1}{2}(\delta_{(0,0)} + \delta_{(-2,1)}).$$
Scrivere esplicitamente una geodetica μ_t in $\mathscr{P}_2(\mathbb{R}^2)$ che connette μ_1 a μ_2.

2. Calcolando esplicitamente $\Phi(\mu_t)$ per una scelta particolare di μ_0, dimostrare che $t \mapsto \Phi(\mu_t)$ non è λ-convessa per nessun $\lambda \in \mathbb{R}$. (Suggerimento: $\mu_0 := \frac{1}{2}(\delta_{(0,0)} + \delta_{(0,-2)})$.)

□ □ □ □ □

4.8 Sottodifferenziali in $\mathscr{P}_2(\mathbb{R}^m)$

Sia H uno spazio di Hilbert. La controparte metrica del differenziale di Fréchet introdotto nella Sezione 1 è la *pendenza metrica*, che per un funzionale $\phi : H \to \mathbb{R} \cup \{+\infty\}$ è definita da

$$|\partial\phi|(v) := \limsup_{w \to v} \frac{(\phi(v) - \phi(w))^+}{\|v - w\|} \qquad \text{per } v \in D(\phi), \tag{4.24}$$

e può essere caratterizzata da uno sviluppo asintotico, per $s \geq 0$

$$s \geq |\partial\phi|(v) \qquad \Leftrightarrow \qquad \phi(w) \geq \phi(v) - s\|w - v\| + o(\|w - v\|) \quad \text{per } w \to v.$$

Lo sviluppo ricorda la nota caratterizzazione del sottodifferenziale:

$$u \in \partial\phi(v) \qquad \Leftrightarrow \qquad \phi(w) \geq \phi(v) + (u, w - v) \quad \forall\, w \in H. \tag{4.25}$$

Nota: la nozione di sottodifferenziale usata nella Sezione 2 ha un carattere globale (vale $\forall w \in H$), adatta a funzioni convesse. La definizione (4.24), invece, è una nozione locale, che generalizza il differenziale locale standard. Nel caso di funzionali λ-convessi, vi sono altre due proprietà utili, corrispondenti alla caratterizzazione (4.25) e alla monotonia di $\partial\phi$. Se ϕ è λ-convesso, allora, per ogni $v, v_1, v_2 \in D(\phi)$

$$\begin{aligned} u \in \partial\phi(v) \qquad &\Leftrightarrow \qquad \phi(w) \geq \phi(v) + (u, w - v) + \frac{\lambda}{2}\|v - w\|^2 \quad \forall\, w \in H, \\ u_i \in \partial\phi(v_i) \qquad &\Leftrightarrow \qquad (u_2 - u_1, v_2 - v_1) \geq \lambda\|v_2 - v_1\|^2. \end{aligned}$$

Inoltre, nel caso in cui $\partial\phi(v)$ è un insieme, la pendenza metrica seleziona la minima norma degli elementi di $\partial\phi(v)$:

$$|\partial\phi|(v) = \min\big\{\|u\| \ : \ u \in \partial\phi(v)\big\}.$$

(Con la notazione della Sezione 2, indicheremmo $|\partial\phi|(v)$ con $|A^0 v|$, dove $A = \partial\phi$.)

Cerchiamo ora di riprodurre questi concetti nell'ambito di $(\mathscr{P}_2(\mathbb{R}^m), W_2)$. Sia $\phi : \mathscr{P}_2(\mathbb{R}^m) \to \mathbb{R} \cup \{+\infty\}$, definiamo la pendenza metrica di ϕ:

$$|\partial\phi|(\mu) := \limsup_{\nu\to\mu} \frac{(\phi(\nu) - \phi(\mu))^+}{W_2(\nu,\mu)} \qquad \text{per } \mu \in D(\phi). \tag{4.26}$$

Per semplificare alcune ipotesi supponiamo che ϕ sia funzionale proprio, semicontinuo inferiormente e tale che $D(|\partial\phi|) \subset \mathscr{P}_2^a(\mathbb{R}^m)$. Sotto tali ipotesi, per $\mathbf{v} \in L^2(\mu;\mathbb{R}^m)$ si ha:

$$\mathbf{v} \in \partial\phi(\mu) \quad \Leftrightarrow \quad \phi(\nu) \geq \phi(\mu) + \int_{\mathbb{R}^m} \mathbf{v}(x) \cdot (\mathbf{t}(x) - x)\, d\mu(x) + \frac{\lambda}{2} W_2^2(\mu,\nu), \quad \forall\, \nu \in D(\phi)$$

dove $\mathbf{t} : \mathbb{R}^m \to \mathbb{R}^m$ è la (unica) mappa di trasporto ottimo fra μ e ν (cioè quella per cui

$$\mathbf{t}_{\#}\mu = \nu \quad \text{e} \quad \int_{\mathbb{R}^m} |x - \mathbf{t}(x)|^2\, d\mu = W_2^2(\mu,\nu),$$

la cui esistenza è assicurata dal Teorema 4.31.) Inoltre, se $\mathbf{v}_i \in \partial\phi(\mu_i)$, $i = 1, 2$, e $\mathbf{t}$ è la mappa di trasporto ottimo fra μ_1 e μ_2, allora

$$\int_{\mathbb{R}^m} \big(\mathbf{v}_2(\mathbf{t}(x)) - \mathbf{v}_1(x)\big) \cdot (\mathbf{t}(x) - x)\, d\mu_1(x) \geq \lambda W_2^2(\mu_1,\mu_2).$$

Informalmente, si può pensare al seguente meccanismo per tradurre un enunciato in uno spazio di Hilbert nel corrispondenti in $(\mathscr{P}_2(\mathbb{R}^m), W_2)$: il punto di riferimento u diventa una misura μ, il prodotto scalare $(\ ,\ )$ deve essere inteso nello spazio $L^2(\mu;\mathbb{R}^m)$ (che contiene le velocità $\mathbf{v}_t$ a curve AC), e i vettori spostamento $w - v$ diventano le mappe di trasporto $\mathbf{t} - id$, con $\mathbf{t}_{\#}\mu = \nu$, che esistono non appena $\mu \in \mathscr{P}_2^a(\mathbb{R}^m)$.

4.9 Gradient flows

In questa sezione definiamo il flusso gradiente nello spazio metrico $(\mathscr{P}_2(\mathbb{R}^m), W_2)$ ed enunciamo il principale risultato di esistenza e unicità. Nel seguito assumiamo che $\phi : \mathscr{P}_2(\mathbb{R}^m) \to \mathbb{R} \cup \{+\infty\}$ sia un funzionale proprio e semicontinuo inferiormente, e che

(i) ϕ sia λ-geodeticamente convesso,

(ii) $D(|\partial\phi|) \subset \mathscr{P}_2^a(\mathbb{R}^m)$,

Diciamo che una curva $\mu_t \in AC^2_{\mathrm{loc}}((0,+\infty); \mathscr{P}_2(\mathbb{R}^m))$ è una soluzione dell'equazione di gradient flow di ϕ se esiste un campo vettoriale Borel-misurabile $\mathbf{v}_t$ tale che

$$\text{(i)} \quad \|\mathbf{v}_t\|_{L^2(\mu_t;\mathbb{R}^m)} = |\mu'|(t) \in L^2_{\mathrm{loc}}(0,+\infty), \tag{4.27}$$

$$\text{(ii)} \quad \partial_t\mu_t + \mathrm{div}(\mathbf{v}_t\mu_t) = 0 \qquad \text{in } \mathcal{D}'(\mathbb{R}^m \times (0,+\infty)), \tag{4.28}$$

$$\text{(iii)} \quad -\mathbf{v}_t \in \partial\phi(\mu_t) \qquad \text{per q.o. } t > 0. \tag{4.29}$$

Il seguente teorema è tratto da [3, Theorem 5.8].

Teorema 4.39 (Esistenza dei gradient flow). *Sia $\phi : \mathscr{P}_2(\mathbb{R}^m) \to \mathbb{R}\cup\{+\infty\}$ un funzionale proprio e semicontinuo inferiormente, supponiamo inoltre che* *(i)* ϕ *sia* λ*-geodeticamente convesso,*

(ii) $D(|\partial\phi|) \subset \mathscr{P}_2^a(\mathbb{R}^m)$,

(iii) per ogni $C > 0$ *siano compatti i sottolivelli*

$$\{\mu \in \mathscr{P}_2(\mathbb{R}^m) : \phi(\mu) \le C,\ M_2(\mu) \le C\}\,.$$

Allora, per ogni $\mu_0 \in \overline{D(\phi)}$ *esiste un'unica soluzione* μ_t *del gradient flow* (4.29) *con dato iniziale che verifica*

$$\lim_{t\to 0} \mu_t = \mu_0 \qquad \textit{in } \mathscr{P}_2(\mathbb{R}^m).$$

Nota: l'ipotesi (*iii*) si può sostituire con un rafforzamento della convessità di ϕ lungo le geodetiche. È soddisfatta dagli esempi standard di funzionali di energia interna, potenziale e di interazione, non appena questi sono λ-convessi lungo le geodetiche.

5 Approfondimenti

1. Teoria dei semigruppi lineari: Capitolo II, pagine 47–58 di [8], in particolare il Diagramma 1.14 p. 58. (Nota: sarà utile, se non necessario, leggere anche il Capitolo I.)

2. Teorema di Hille-Yosida lineare: [6, Teorema VII.7]. (Nota: sarà necessario leggere anche la parte precedente del capitolo VII sugli operatori massimali monotoni.)

3. Teorema di Ball sulle misure di Young [4]. https://people.maths.ox.ac.uk/ball/Articles%20in%20Conference%20Proceedings%20and%20Books/Ball%20-%20Version%20of%20the%20Fundamental%20Theorem%20for%20Young%20Measures.pdf

4. Esistenza per un passo dello schema iterativo di [10]: Introduzione dell'articolo e dimostrazione di [10, Proposition 4.1].

5. Convergenza dello schema iterativo di [10]: dimostrazione di [10, Theorem 5.1].

6. Relazione fra equazione di continuità e ODE caratteristica: dimostrazione del Lemma 4.34 e della Proposizione 4.35. (Si veda [3, Lemma 2.10 e Proposition 2.12].)

6 Soluzioni di alcuni esercizi

4.22 (b) Per ogni $\mu \in \mathcal{K}$, per ogni $0 < p < p_1$ si ha

$$\begin{aligned} C &> \int_X d(x,\bar{x})^{p_1} d\mu \ge \int_{B_k(\bar{x})} d(x,\bar{x})^{p_1} d\mu = \int_{B_k(\bar{x})} d(x,\bar{x})^p d(x,\bar{x})^{p_1-p} d\mu \\ &\ge \int_{B_k(\bar{x})} d(x,\bar{x})^p k^{p_1-p} d\mu, \end{aligned}$$

cioè

$$\sup_{\mu\in\mathcal{K}}\int_{B_k(\bar{x})} d(x,\bar{x})^p d\mu \leq \frac{C}{k^{p_1-p}}.$$

Passando al limite per $k\to+\infty$, si ottiene l'uniforme integrabilità voluta.

4.24 1. Per $x\in(0,a)$, dalla definizione si ha

$$\pi^1_\#\mu([0,x]) = \mu\big((\pi^1)^{-1}([0,x])\big) = \frac{\mathcal{L}^2(Q\cap([0,x]\times\mathbb{R}))}{a\sqrt{2}} = \frac{x^2}{a\sqrt{2}}.$$

Indicando con ρ la densità di $\pi^1_\#\mu$ rispetto alla misura di Lebesgue, si ha

$$\rho(x) = \frac{d}{dx}\frac{x^2}{a\sqrt{2}} = \frac{2x}{a\sqrt{2}}, \qquad x\in[0,a].$$

Per $x\in[a,2a]$, per simmetria, $\rho(x) = \frac{-2}{a\sqrt{2}}(x-2a)$.

2. $\gamma_1=\mu$, $\quad\gamma_2 = (\rho\mathcal{L}^1)\times(\rho\mathcal{L}^1)$, $\quad \gamma_3 = (id\times\mathbf{t})_\#(\rho\mathcal{L}^1)$, con $\mathbf{t}(x)=id(x)=x$ (dato che le due marginali sono identiche).

3.

$$\int_{\mathbb{R}^2} f(x,y)\,d\gamma_1 = \int_Q f(x,y)\,dx\,dy$$
$$\int_{\mathbb{R}^2} f(x,y)\,d\gamma_2 = \int_{\mathbb{R}^2} f(x,y)\rho(x)\rho(y)\,dx\,dy$$
$$\int_{\mathbb{R}^2} f(x,y)\,d\gamma_3 = \int_{\mathbb{R}} f(x,x)\rho(x)\,dx$$

4.25 Sia $\{\gamma_n\}\subset\Gamma(\mu,\nu)\subset\mathscr{P}(X\times Y)$ una successione di piani, tale che $\gamma_n\rightharpoonup\gamma\in\mathscr{P}(X\times Y)$, strettamente. Sia $f\in C^0_b(X)$, usando, rispettivamente, $\pi^1_\#\gamma_n=\mu$, la definizione di marginale, la definizione convergenza stretta, e di nuovo la definizione di marginale

$$\begin{aligned}\int_X f(x)\,d\mu(x) &= \lim_{n\to\infty}\int_X f(x)\,d(\pi^1_\#\gamma_n)(x)\\ &= \lim_{n\to\infty}\int_{X\times Y} f\left(\pi^1(x,y)\right)\,d\gamma_n(x,y)\\ &= \int_{X\times Y} f\left(\pi^1(x,y)\right)\,d\gamma(x,y)\\ &= \int_X f(x)\,d(\pi^1_\#\gamma)(x).\end{aligned}$$

Dato che f è arbitraria, concludiamo che $\mu=\pi^1_\#\gamma$.

4.27 $c(\bar{x},\bar{y})$.

4.38 Si veda [3, Example 3.3 pag. 38]

References

[1] A. Ambrosetti and G. Prodi. *A primer of nonlinear analysis*, volume 34 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Corrected reprint of the 1993 original.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics. ETH Zürich. Birkhauser, Basel, 2nd edition, 2008.

[3] L. Ambrosio and G. Savaré. Gradient flows of probability measures. In *Handbook of differential equations: evolutionary equations. Vol. III*, Handb. Differ. Equ., pages 1–136. Elsevier/North-Holland, Amsterdam, 2007.

[4] J. M. Ball. A version of the fundamental theorem for Young measures. In *PDEs and continuum models of phase transitions (Nice, 1988)*, volume 344 of *Lecture Notes in Phys.*, pages 207–215. Springer, Berlin, 1989.

[5] V. I. Bogachev. *Measure theory. Vol. I.* Springer-Verlag, Berlin, 2007.

[6] H. Brezis. *Analisi funzionale. Teoria e applicazioni.* Liguori, Napoli, 1986.

[7] G. Dal Maso. *An Introduction to* Γ*-Convergence*, volume 8 of *Progress in Nonlinear Differential Equations and Their Applications.* Birkhäuser, Boston, first edition, 1993.

[8] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[9] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, first edition, 2002.

[10] R. Jordan, D. Kinderlehrer, and F. Otto. The Variational Formulation of the Fokker-Planck Equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

[11] F. Otto. The geometry of dissipative evolution equations: The porous medium equation. *Communications in Partial Differential Equations*, 26(1&2):101–174, 2001.