Rendiconti
Accademia Nazionale delle Scienze detta dei XL
Memorie di Matematica e Applicazioni
123° (2005), Vol. XXIX, fasc. 1, pagg. 15-34

L. AMBROSIO (*)

Flussi gradiente in spazi metrici e nello spazio di Wasserstein delle misure di probabilità

1. - Introduzione

In questo articolo riassumiamo i principali risultati del libro [4], dedicato ad una teoria dei flussi gradiente in spazi metrici generali e nello spazio delle misure di probabilità munito della distanza di Wasserstein quadratica. La sintesi che faremo dei risultati (largamente ispirata a [3] con qualche piccola aggiunta, modifica e aggiornamento) riflette la struttura del libro, con due parti che possono essere lette quasi indipendentemente: nella prima studiamo i flussi gradiente in spazi metrici (chiamandoli curve di massima pendenza, dato che solo il concetto di pendenza, e non quello di gradiente, ha senso in spazi metrici), seguendo con piccole varianti le idee sviluppate in [12] (si veda anche [2]).

Troviamo anche condizioni generali che assicurano la convergenza dello schema di discretizzazione implicito a una curva di massima pendenza. Queste condizioni sono tipicamente soddisfatte se il funzionale è convesso lungo le geodetiche (parametricamente percorse a velocità costante). Troviamo anche una nuova condizione di convessità, che combina il comportamento del funzionale e della distanza, e che assicura sia l'unicità delle curve di massima pendenza che stime esplicite dell'errore per lo schema di discretizzazione implicito (e qui sviluppiamo ulteriormente le idee introdotte in [6, 21]). Questi risultati estendono altri già noti in spazi con curvatura non positiva (NPC) (si veda [19], [17]), ma si possono applicare anche ad alcuni spazi con curvatura non negativa (PC), come ad esempio lo spazio di Wasserstein delle misure di probabilità.

Nella seconda parte descriveremo lo spazio di Wasserstein delle misura di probabilità in uno spazio di Hilbert separabile X e la sua struttura "differenziale", recuperando in un

(*) Indirizzo dell'Autore: Scuola Normale Superiore, Piazza dei Cavalieri, 7, 56100 Pisa (Italy), e-mail: l.ambrosio@sns.it

contesto molto più generale il calcolo formale sviluppato da Otto in [22] e la rappresentazione della distanza di Wasserstein come distanza “Riemanniana” in [7]. Una delle principali caratteristiche del nostro approccio è che non richiediamo né che X sia di dimensione finita né che le misure in questione siano assolutamente continue rispetto a una misura fissata (ad esempio quella di Lebesgue). La nostra derivazione della struttura “Riemanniana” (i.e. il fibrato tangente e la metrica sulle fibre) nasce da una caratterizzazione variazionale delle curve assolutamente continue a valori nello spazio di Wasserstein. Dopo la costruzione del fibrato tangente esamineremo in dettaglio le proprietà di differenziabilità, lungo la curve assolutamente continue, della distanza di Wasserstein da una misura fissata, e caratterizziamo il vettore tangente alla curva attraverso il comportamento asintotico dei piani di trasporto tra punti vicini sulla curva. In questo contesto è possibile introdurre in maniera naturale anche i concetti di sottodifferenziale e di flusso gradiente (si veda anche [10], dove concetti analoghi appaiono nel contesto dei cosiddetti spazi di lunghezza quasi-Riemanniani) e si trova che i flussi gradiente coincidono con le curve di massima pendenza. Questo porta in molti casi a risultati di esistenza per i flussi gradiente, alla convergenza dello schema di discretizzazione implicita (si veda anche [16], [9], [1]) e, sotto un'ipotesi leggermente più restrittiva di convessità, a stime esplicite dell'errore per lo schema di discretizzazione implicita.

2. - Flussi gradiente in spazi metrici

2.1. *Nozioni preliminari*

Iniziamo con l'introdurre alcune nozioni di base che hanno senso in ogni spazio metrico completo $(\mathscr{S}, d)$ (si veda [2, 5] per una trattazione più completa di questo argomento).

Definizione 2.1: *Sia $v : (a,b) \to \mathscr{S}$ una curva; diremo che v appartiene a $AC^p(a,b;\mathscr{S})$, per $p \in [1,+\infty]$, se esiste $m \in L^p(a,b)$ tale che*

$$d(v(s), v(t)) \leq \int_s^t m(r)\,dr \qquad \forall\, a < s \leq t < b. \tag{2.1}$$

Nel caso $p = 1$ lo spazio coincide con quello delle curve assolutamente continue e lo indicheremo semplicemente con $AC(a,b;\mathscr{S})$.

Teorema 2.2: (Derivata metrica). *Sia $p \in [1,+\infty]$. Allora, per ogni $v \in AC^p(a,b;\mathscr{S})$ il limite*

$$|v'|(t) := \lim_{s\to t} \frac{d(v(s), v(t))}{|s-t|} \tag{2.2}$$

esiste per $\mathscr{L}^1$*-q.o.* $t \in (a,b)$*. Inoltre la funzione* $t \mapsto |v'|(t)$ *appartiene* $L^p(a,b)$*, è un integrando ammissibile per il membro destro della* (2.1), *ed è minimo in questo senso:*

(2.3) $|v'|(t) \leq m(t)$ per $\mathscr{L}^1$-*q.o.*$t \in (a,b)$, per ogni funzione m soddisfacente (2.1)

Definizione 2.3: (Pendenza locale). *La pendenza locale di un funzionale* ϕ *nel punto* $v \in D(\phi)$ *è 0 se* v *è un punto isolato di* $\mathscr{S}$*, altrimenti è definito da*

$$|\partial\phi|(v) := \limsup_{w \to v} \frac{\big(\phi(v) - \phi(w)\big)^+}{d(v,w)}. \tag{2.4}$$

Indicheremo con $D(\partial\phi)$ *l'insieme dei* $v \in \mathscr{S}$ *tali che* $|\partial\phi|(v) < +\infty$*.*

Avendo in mente queste definizioni, possiamo formulare come in [12] (si veda anche [12, 18, 2]) il concetto di flusso gradiente in spazi metrici generali.

Definizione 2.4: (Curve di massima pendenza). *Diremo che una mappa localmente assolutamente continua* $u : (a,b) \to \mathscr{S}$ *è una* curva di massima pendenza *del funzionale* ϕ *se* $\phi \circ u$ *è* $\mathscr{L}^1$*-q.o. uguale a una mappa non decrescente* φ *e*

$$\varphi'(t) \leq -\frac{1}{2}|u'|^2(t) - \frac{1}{2}|\partial\phi|^2(u(t)) \quad \text{per } \mathscr{L}^1\text{-}q.o.\ t \in (a,b). \tag{2.5}$$

Diremo che $\tilde{u}$ *è il punto iniziale della curva* u *se* $\lim_{t \to a} u(t) = \tilde{u}$*.*

Per illustrare le idee euristiche alla base della precedente definizione, incominciamo con il caso classico del flusso gradiente

$$u'(t) = -\nabla\phi(u(t)) \tag{2.6}$$

in uno spazio di Hilbert. Se, in ambo i membri, prendiamo i moduli, otteniamo l'equazione $|u'|(t) = |\nabla\phi(u(t))|$ che ha senso in spazi metrici, interpretando il membro sinistro come la derivata metrica e il membro destro come la pendenza locale. Tuttavia, nel passare da (2.6) a un'equazione scalare abbiamo chiaramente una perdita di informazione. Questa informazione può essere recuperata guardando alla derivata dell'energia:

$$\frac{d}{dt}\phi(u(t)) = \langle u'(t), \nabla\phi(u(t))\rangle = -|u'(t)||\nabla\phi(u(t))| = -\frac{1}{2}|u'|^2(t) - \frac{1}{2}|\nabla\phi(u(t))|^2.$$

La seconda uguaglianza vale se e solo se u' e $-\nabla\phi(u)$ sono parallele e la terza uguaglianza vale se e solo se $|u'|$ e $|\nabla\phi(u)|$ sono uguali, cosicché possiamo riscrivere (2.6) come segue:

$$\frac{1}{2}|u'|^2(t) + \frac{1}{2}|\nabla\phi(u(t))|^2 = -\frac{d}{dt}\phi(u(t)).$$

Questo argomento mostra che la formulazione metrica è consistente con il classico caso Hilbertiano.

2.2. *Discretizzazione implicita nel tempo*

La costruzione classica delle curve di massima pendenza è basata su una discretizzazione temporale, che fornisce soluzioni discrete dipendenti da un passo temporale $\tau > 0$, che approssimano la curva di massima pendenza per $\tau \to 0$.

Le soluzioni discrete sono costruite attraverso uno schema variazionale ricorsivo nel modo seguente: fissato $\tau > 0$ e il dato iniziale U_τ^0, scegliamo U_τ^{n+1} nell'insieme

$$J_\tau[U_\tau^n] := \operatorname{argmin}\Big\{\phi(\,\cdot\,) + \frac{1}{2\tau}d^2(U_\tau^n,\cdot)\Big\}. \tag{2.7}$$

Definizione 2.5: (Soluzioni discrete). *La soluzione discreta di passo τ è la funzione costante a tratti $\bar{U}_\tau$ definita da:*

$$\bar{U}_\tau(t) := U_\tau^n \qquad \textit{if } \ t \in [n\tau,(n+1)\tau).$$

Dato che vogliamo mostrare la convergenza delle soluzioni discrete a una curva limite, dobbiamo imporre delle proprietà di compattezza dello spazio metrico (o dei sottolivelli di ϕ). Naturalmente si può richiedere compattezza rispetto alla topologia indotta dalla distanza d, ma questo requisito sarebbe troppo restrittivo per alcune applicazioni. Per questa ragione supporremo che esista una topologia σ con le seguenti proprietà:

2.1a: Topologia debole. — σ è una topologia di Hausdorff su $\mathscr{S}$ compatibile con d nel senso che σ è più debole della topologia indotta da d e d è σ-semicontinuo inferiormente per successioni:

$$(u_n,v_n) \overset{\sigma}{\rightharpoonup} (u,v) \quad \Longrightarrow \quad \liminf_{n\to\infty} d(u_n,v_n) \geq d(u,v). \tag{2.8}$$

De Giorgi ha introdotto in [11] la nozione di limite di soluzioni discrete in senso astratto (dove si considerano perturbazioni più generali del quadrato della distanza), coniando il termine *movimenti minimizzanti*. Ricordiamo la sua definizione.

Definizione 2.6: (Movementi minimizzanti). *Per un dato funzionale ϕ e un dato iniziale $u_0 \in \mathscr{S}$, diremo che una curva $u : [0,+\infty) \to \mathscr{S}$ è un* movimento minimizzante *per ϕ con dato iniziale u_0 se per ogni $\tau > 0$ sufficientemente piccolo esistono soluzioni discrete $\bar{U}_\tau$ definite come in* (2.5) *e tali che*

$$\lim_{\tau\downarrow 0}\phi(U_\tau^0) = \phi(u_0), \quad \lim_{\tau\downarrow 0} d(U_\tau^0,u_0) = 0, \quad \bar{U}_\tau(t) \overset{\sigma}{\rightharpoonup} u(t) \quad \forall\, t \in [0,+\infty). \tag{2.9}$$

Indichiamo con $MM(\phi;u_0)$ l'insieme di tutti i movimenti minimizzanti per ϕ con dato iniziale u_0.

Analogamente, diremo che una curva $u : [0,+\infty) \to \mathscr{S}$ è un movimento minimizzante generalizzato *per ϕ con dato iniziale u_0 se esiste una successione $\tau_k \downarrow 0$ e una corrispondente successione di soluzioni discrete $\bar{U}_{\tau_k}$ definite come in* (2.5) *tali che*

$$\lim_{k\to\infty}\phi(U_{\tau_k}^0) = \phi(u_0), \quad \lim_{k\to\infty} d(U_{\tau_k}^0,u_0) = 0, \quad \bar{U}_{\tau_k}(t) \overset{\sigma}{\rightharpoonup} u(t) \quad \forall\, t \in [0,+\infty). \tag{2.10}$$

Indichiamo con $GMM(\phi; u_0)$ *l'insieme di tutti i movimenti minimizzanti generalizzati per* ϕ *con dato iniziale* u_0 .

Per essere sicuri dell'esistenza di soluzioni discrete, e per mostrare che i movimenti minimizzanti (generalizzati) sono curv di massima pendenza, dovremo imporre alcune condizioni sul funzionale ϕ.

2.1b: Semicontinuità inferiore. — Supponiamo che ϕ sia σ-semicontinuo inferiormente per successioni sugli insiemi d-limitati:

$$\sup_{n,m} d(u_n, u_m) < +\infty, \quad u_n \overset{\sigma}{\rightharpoonup} u \quad \Longrightarrow \quad \liminf_{n\to\infty} \phi(u_n) \geq \phi(u). \tag{2.11a}$$

2.1c: Coercività. — Esistono $\tau_* > 0$ e $u_* \in \mathscr{S}$ tali che

$$m^* := \inf_{v\in\mathscr{S}} \phi(v) + \frac{1}{2\tau} d^2(v, u_*) > -\infty. \tag{2.11b}$$

2.1d: Compattezza. — Ogni insieme d-limitato contenuto in un sottolivello di ϕ è relativamente σ-compatto per successioni:

$$\begin{gathered}\text{ogni successione } (u_n) \subset \mathscr{S} \text{ con } \sup_n \phi\,(u_n) < +\infty, \quad \sup_{n,m}\; d(u_n, u_m) < +\infty \\ \text{ha una sottosuccessione } \sigma\text{-convergente.}\end{gathered} \tag{2.11c}$$

2.1e: Semicontinuità della pendenza locale.

$$|\partial\phi|(u) = \inf\left\{ \liminf_{n\to+\infty} |\partial\phi|(u_n) \;:\; u_n \overset{\sigma}{\rightharpoonup} u,\; \sup_n\{d(u_n, u), \phi(u_n)\} < +\infty \right\}. \tag{2.11d}$$

In pratica le ipotesi b, c assicurano l'esistenza di soluzioni discrete, d è necessaria per trovare una curva limite e e è necessaria per mostrare che questo limite è una curva di massima pendenza.

Proposizione 2.7: (Convergenza dello schema di discretizzazione implicita). *Supponiamo che le ipotesi* 2.1a, b, c, d *valgano e siano date una successione* $(\tau_n) \downarrow 0$ *e una corrispondente successione di dati iniziali* $\{U^0_{\tau_n}\}$ *soddisfacenti*

$$U^0_{\tau_n} \overset{\sigma}{\rightharpoonup} u_0, \quad \phi(U^0_{\tau_n}) \to \phi(u_0) < +\infty \quad \text{as } n \to +\infty, \quad \sup_n d(U^0_{\tau_n}, u_0) < +\infty. \tag{2.12}$$

Allora esistono una sottosuccessione $n(k)$ e $u \in AC^2_{\mathrm{loc}}([0, +\infty); \mathscr{S})$ tali che

$$\bar{U}_{\tau_{n(k)}}(t) \overset{\sigma}{\rightharpoonup} u(t) \qquad \forall\, t \in [0, +\infty). \tag{2.13}$$

In particolare $u \in GMM(\phi; u_0)$, che risulta un sottoinsieme non vuoto di $AC^2_{\mathrm{loc}}([0, +\infty); \mathscr{S})$.

Teorema 2.8: (Le curve limite sono di massima pendenza). *Supponiamo che le condizioni* 2.1*a*, *b*, *c*, *e valgano e che* ϕ *soddisfi la proprietà di continuità*

$$\sup_{n\in\mathbb{N}}\left\{|\partial\phi|(v_n), d(v_n, v_0), \phi(v_n)\right\} < +\infty, \quad v_n \overset{\sigma}{\rightharpoonup} v \quad \Longrightarrow \quad \phi(v_n) \to \phi(v). \tag{2.14}$$

Allora ogni $u \in GMM(\phi; u_0)$ *è una curva di massima pendenza per* ϕ.

Alcuni risultati più forti (i.e. l'identità dell'energia e la convergenza di varie quantità discrete alla loro controparte continua) possono essere ottenuti sostituendo l'ipotesi di continuità (2.14) con l'ipotesi che $|\partial\phi|$ sia un upper gradient di ϕ nel senso di Heinonen e Koskela (si veda [15] e la (2.15)).

Teorema 2.9: *Supponiamo che le condizioni* 2.1*a*, *b*, *c*, *e valgano e che* $|\partial\phi|$ *abbia la seguente proprietà: per ogni curva assolutamente continua* $v : [0, 1] \to \mathscr{S}$ *la funzione* $|\partial\phi| \circ v$ *è di Borel e*

$$|\phi(v(t)) - \phi(v(s))| \leq \int_t^s |\partial\phi|(v(r))|v'(r)|\, dr \quad \forall 0 < s \leq t < 1. \tag{2.15}$$

Allora ogni $u \in GMM(\phi; u_0)$ è una curva di massima pendenza per ϕ for ϕ e u soddisfa l'identità dell'energia

$$\frac{1}{2}\int_0^T |u'|^2(t)\, dt + \frac{1}{2}\int_0^T |\partial\phi|^2(u(t))\, dt + \phi(u(T)) = \phi(u_0) \qquad \forall T > 0. \tag{2.16}$$

Inoltre, se $\{\bar{U}_{\tau_n}\}_{n\in\mathbb{N}}$ è una successione di soluzioni discrete soddisfacenti (2.12) *e* (2.13), abbiamo

$$\lim_{n\to\infty} \phi(\bar{U}_{\tau_n}(t)) = \phi(u(t)) \quad \forall\, t \in [0, +\infty), \tag{2.17}$$

$$\lim_{n\to\infty} |\partial\phi|(\bar{U}_{\tau_n}) = \lim_{n\to\infty} |U'_{\tau_n}| = |u'| = |\partial\phi|(u) \quad \text{in } L^2_{\text{loc}}([0, +\infty\ )), \tag{2.18}$$

dove $|U'_\tau|$ *è definita da*

$$|U'_\tau|(t) = \frac{d(U_\tau^{n-1}, U_\tau^n)}{\tau} \quad \text{se } t \in ((n-1)\tau, n\tau) \tag{2.19}$$

e $|\partial^-\phi|$ *il membro destro in* (2.11*d*).

2.3. *Funzionali convessi lungo le geodetiche*

Le condizioni astratte date qui sopra sono soddisfatte quando il funzionale è convesso lungo le geodetiche (percorse a velocità costante). Questo caso è rilevante per molte applicazioni, si veda per esempio [17, 20, 22, 23] e gli esempi menzionati nella prossima sezione.

Definizione 2.10: (Convessità lungo le curve). *Un funzionale* $\phi : \mathscr{S} \to (-\infty + \infty]$ *è detto convesso lungo la curva* $\gamma : [0,1] \to \mathscr{S}$ *se*

$$\phi(\gamma_t) \leq (1-t)\phi(\gamma_0) + t\phi(\gamma_1) \tag{2.20}$$

Definizione 2.11: (Geodetiche a velocità costante e spazi di lunghezza). *Una curva* $\gamma : [0,1] \to \mathscr{S}$ *è una geodetica a velocità costante congiungente due punti* $v_0,\ v_1 \in \mathscr{S}$ *se* $\gamma_i = v_i$, $i = 0,\ 1$ *e*

$$d(\gamma_s, \gamma_t) = (t-s)d(v_0, v_1) \qquad \forall s,\ t \in [0,1],\ \ s \leq t. \tag{2.21}$$

Uno spazio metrico $\mathscr{S}$ *è detto* spazio di lunghezza *se per ogni coppia di punti* $x,\ y \in \mathscr{S}$ *esiste almeno una (non necessariamente unica) geodetica che li congiunge.*

Definizione 2.12: (Funzionali convessi lungo le geodetiche). *Un funzionale* $\phi : \mathscr{S} \to$ $\to (-\infty, +\infty]$ *su uno spazio di lunghezza* $\mathscr{S}$ *è detto* convesso lungo le geodetiche *se per ogni* $x,\ y \in \mathscr{S}$ *esiste una geodetica a velocità costante* γ *tale che* $\gamma_0 = x$, $\gamma_1 = y$ *e* $t \mapsto \phi(\gamma_t)$ *è una funzione convessa in* $[0,1]$.

Per funzionali convessi lungo le geodetiche la pendenza locale ammette la seguente semplice rappresentazione:

$$|\partial\phi|(v) = \sup_{w \neq v} \left(\frac{\phi(v) - \phi(w)}{d(v,w)} \right)^+, \tag{2.22}$$

usando la quale si può mostrare che $|\partial\phi|$ ha la proprietà di upper gradient formulata nel Teorema 2.9 e che 2.1e vale con la topologia σ uguale a quella indotta da d. Si ottiene quindi il seguente risultato.

Teorema 2.13: *Supponiamo che* ϕ *sia convesso lungo le geodetiche e che tutte le ipotesi* 2.1a, b, c, d *valgano. Allora ogni* u_0 *tale che* $\phi(u_0) < +\infty$ *è il punto iniziale di una curva di massima pendenza per* ϕ *e valgono tutte le conclusioni del Teorema* 2.9.

La proprietà di convessità garantisce inoltre alcune proprietà puntuali di differenziabilità, osservate per la prima volta da Brezis [8] in un contesto Hilbertiano. Sono raccolte nel seguente teorema.

Teorema 2.14: (Proprietà puntuali). *Supponiamo che le ipotesi* 2.1a, b *valgano e che* ϕ *sia convesso lungo le geodetiche. Se* $\phi(u_0) < +\infty$ *allora ogni* $u \in GMM(\phi; u_0)$ *è localmente Lipschitz in* $(0, +\infty)$ *e soddisfa le seguenti proprietà:*

(i) *La derivata metrica destra*

$$|u'_+|(t) := \lim_{s \downarrow t} \frac{d(u(s), u(t))}{s - t} \tag{2.23}$$

esiste e $|\partial\phi|(u(t)) < +\infty$ *per ogni* $t > 0$.

(ii) *La funzione* $t \mapsto \phi(u(t))$ *è convessa e la funzione* $t \mapsto |\partial\phi|(u(t))$ *è non crescente e continua a destra.*

(iii) *L'equazione*

$$\frac{d}{dt_+}\phi(u(t)) = -|\partial\phi|^2(u(t)) = -|u'_+|^2(t) = -|\partial\phi|(u(t))\,|u'_+|(t) \tag{2.24}$$

vale per ogni $t \in (0, +\infty)$.

Anche se la convessità geodetica assicura l'esistenza di curve di massima pendenza e molte utili proprietà, compresa l'identità dell'energia, l'unicità delle curve di massima pendenza con un fisato dato iniziale è un problema aperto. Il problema è irrisolto anche in un contesto lineare, per esempio quando $\mathscr{S}$ è uno spazio di Banach riflessivo, e sembra dipendere dalla mancanza di una struttura Hilbertiana. Alcuni risultati parziali sono discussi nella prossima sottosezione.

2.4. *Un nuovo tipo di convessità*

Stime esplicite dell'errore per il metodo di discretizzazione implicita e l'unicità delle curve di massima pendenza possono essere dimostrare se si richiede la convessità non solo del funzionale, ma anche della distanza.

Ipotesi 2.15: *Supponiamo che per ogni scelta di* $w, v_0, v_1 \in \mathscr{S}$ *esista una curva* $\gamma : [0,1] \to \mathscr{S}$ *tale che* $\gamma_i = v_i$, $i = 0, 1$, ϕ *è convessa lungo la curva e*

$$d^2(w, \gamma_t) \leq (1-t)d^2(w, v_0) + td^2(w, v_1) - t(1-t)d^2(v_0, v_1) \tag{2.25}$$

Questa ipotesi è in un certo senso più forte della conessità geodetica, dato che stiamo richiedendo una condizione di uniforme convessità anche sulla distanza; d'altro canto è più debole, dato che non imponiamo che la curva γ sia una geodetica, né che sia a velocità costante. Nelle applicazioni descritte nella prossima sezione questo grado di libertà risulterà estremamente utile.

Si noti anche che la condizione (2.25) lungo le geodetiche è la definizione di Non Positively Curved (NPC) spazio metrico nel senso di Alexandroff (per spazi metrici Riemanniani la condizione è equivalente alla non positività di tutte le curvature sezionali). Nel contesto degli spazi NPC l'unicità dei flussi gradiente è stata mostrata in [19] (si veda anche [17]), ma il nostro più generale risultato segue da una linea diversa di ragionamento, fortemente legata a quello usato in [6, 21].

D'ora in poi noi identifichiamo la topologia σ con quella indotta dalla distanza d e lasciamo cadere l'ipotesi di compattezza 2.1d: infatti, l'Ipotesi (2.15) garantisce sia l'esistenza del punto di minimo nel problema discreto 2.7 che la convergenza delle soluzioni discrete a una curva di massima pendenza. I principali risultati sono contenuti nel seguente teorema, nel quale noi usiamo la notazione $\phi_\tau(u)$ per la funzione $\inf \phi(\,\cdot\,) + d^2(\cdot, u)/2\tau$.

Teorema 2.16: *Supponiamo che ϕ sia un funzionale semicontinuo inferiormente e coercivo (in base alla* 2.1*c) e che l'Ipotesi* 2.15 *sia soddisfatta. Allora*

i) Convergenza e formula esponenziale: *per ogni* $u_0 \in \overline{D(\phi)}$ *esiste un unico* $u = S[u_0]$ *in* $MM(\Phi; u_0)$ *che può quindi essere espresso tramite la formula esponenziale*

$$u(t) = S[u_0](t) = \lim_{n\to\infty} U^n_{t/n}(u_0). \tag{2.26}$$

ii) Effetto regolarizzante: *u è una curva localmente Lipschitz di massima pendenza con* $u(t) \in D(\partial\phi) \subset D(\phi)$ *per* $t > 0$*; in particolare, valgono le seguenti stime a priori:*

$$\phi(u(t)) \le \phi_t(u_0), \quad |\partial\phi|^2(u(t)) \le |\partial\phi|^2(v) + \frac{1}{t^2} d^2(v, u_0) \quad \forall\, v \in D(\phi). \tag{2.27}$$

iii) Unicità e disuguaglianze variazionali di evoluzione: *u è l'unica soluzione della disuguaglianza variazionale di evoluzione*

$$\frac{1}{2}\frac{d}{dt} d^2(u(t), v) + \phi(u(t)) \le \phi(v) \quad \mathscr{L}^1\text{-a.e.in } (0, +\infty), \quad \forall\, v \in D(\phi). \tag{2.28}$$

iv) Semigruppo di contrazioni: *La mappa* $t \mapsto S[u_0](t)$ *è un semigruppo non espansivo, i.e.*

$$d\big(S[u_0](t), S[v_0](t)\big) \le d(u_0, v_0) \quad \forall\, u_0, v_0 \in \overline{D(\phi)}. \tag{2.29}$$

v) Stima ottimale a priori: *se* $u_0 \in D(\phi)$ *allora*

$$d^2\Big(S[u_0](t), U^n_{t/n}(u_0)\Big) \le \frac{t}{n}\Big(\phi(u_0) - \phi_{t/n}(u_0)\Big) \le \frac{t^2}{2n^2} |\partial\phi|^2(u_0). \tag{2.30}$$

3. - Flussi gradiente nello spazio delle misure di probabilità

3.1. *Nozioni di base e notazioni*

Questa sezione è dedicata allo studio della struttura differenziale dello spazi delle misure di probabilità su uno spazio di Hilbert separabile X (nel seguito indicato con $\mathscr{P}(X)$), munito della distanza di Wasserstein W_2 (si veda [4] per estensioni al caso delle metriche "di Finsler" W_p con $p \neq 2$).

Iniziamo col richiamare alcune nozioni di base.

Definizione 3.1: (Trasporto di misure). *Sia μ una misura di probabilità su X e sia $r : X \to X$ una mappa di Borel. La misura immagine $r_\#\mu \in \mathscr{P}(X)$ di μ attraverso r è definita da*

$$r_\#\mu(B) := \mu(r^{-1}(B)) \qquad \textit{per ogni insieme di Borel } B \subset X. \tag{3.1}$$

Più in generale, l'integrale rispetto a μ e l'integrale rispetto a $r_{\#}\mu$ sono legati da

$$\int_X f(r(x))d\mu(x) = \int_X f(y)dr_{\#}\mu(y) \tag{3.2}$$

per ogni funzione di Borel f, positiva o limitata.

Definizione 3.2: (Piani di trasporto). *Date due misure μ^1, $\mu^2 \in \mathscr{P}(X)$, l'insieme dei piani di trasporto tra μ_1 e μ_2 è definito da:*

$$\Gamma(\mu^1, \mu^2) := \left\{ \boldsymbol{\mu} \in \mathscr{P}(X \times X) : \pi^1_{\#}\,\boldsymbol{\mu} = \mu^1,\ \pi^2_{\#}\,\boldsymbol{\mu} = \mu^2 \right\}, \tag{3.3}$$

dove $\pi^i : X \times X \to X$, $i = 1, 2$, sono rispettivamente le proiezioni sulla prima e sulla seconda coordinata; si noti che questo insieme è sempre non vuoto, visto che contiene certamente $\mu^1 \times \mu^2$.

La famiglia dei piani di trasporto include in un certo senso quella delle mappe di trasporto: infatti ogni mappa r induce un piano $\boldsymbol{\mu}$ tra μ^1 e $\mu_2 := r_{\#}\mu^1$ definito da $(Id \times r)_{\#}\mu^1$, dove $(Id \times r)(x) = (x, r(x))$.

La distanza di Wasserstein è definita sulla famiglie delle misure in $\mathscr{P}(X)$ il cui secondo momento è finito, i.e.:

$$\mathscr{P}_2(X) := \left\{ \mu \in \mathscr{P}(X) : \int_X |x|^2 d\mu(x) < +\infty \right\}. \tag{3.4}$$

Definizione 3.3: (Il problema del trasporto ottimale). *Date μ^1, $\mu^2 \in \mathscr{P}_2(X)$, la loro distanza di Wasserstein è definita da*

$$W_2^2(\mu^1, \mu^2) := \min \left\{ \int_{X^2} |x_1 - x_2|^2 \, d\boldsymbol{\mu}(x_1, x_2) : \boldsymbol{\mu} \in \Gamma(\mu^1, \mu^2) \right\}. \tag{3.5}$$

Non è difficile mostrare che il minimo è sempre raggiunto e che la funzione definita sopra è una distanza: indicheremo con $\Gamma_o(\mu^1, \mu^2)$ il sottoinsieme di $\Gamma(\mu^1, \mu^2)$ dove il minimo è raggiunto, i.e.

$$\boldsymbol{\mu} \in \Gamma_o(\mu^1, \mu^2) \Longleftrightarrow \int_{X \times X} |x_1 - x_2|^2 \, d\boldsymbol{\mu}(x_1, x_2) = W_2^2(\mu^1, \mu^2). \tag{3.6}$$

Notiamo che $\Gamma_o(\mu^1, \mu^2)$ è un sottoinsieme chiuso e convesso di $\Gamma(\mu^1, \mu^2)$.

Definizione 3.4: (Misure Gaussiane e insiemi Gauss nulli). *Data $\mu \in \mathscr{P}(X)$, diremo che μ è una misura Gaussiana non degenere in X se per ogni $f \in X'$ la misura immagine $f_{\#}\mu \in \mathscr{P}(\mathbb{R})$ è una Gaussiana non degenere sulla retta, i.e. esistono $m = m(f) \in \mathbb{R}$ e*

$\sigma = \sigma(f) > 0$ *tali che*

$$\mu(\{x \in X : -\infty < f(x) < a\}) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{a} e^{-|t-m|^2/2\sigma^2}\, dt \quad \forall a \in \mathbb{R}.$$

Un insieme $B \in \mathscr{B}(X)$ *(dove* $\mathscr{B}(X)$ *indica la* σ*-algebra di Borel) è un insieme Gauss nullo se* $\mu(B) = 0$ *per ogni misura Gaussiana non degenere* μ *in* X*.*

Nel caso in cui μ^1 si annulla su ogni insieme Gauss nullo (nel caso $X = \mathbb{R}^d$ questa condizione equivale all'assoluta continuità di μ^1 rispetto alla misura di Lebesgue $\mathscr{L}^d$) vale il seguente teorema, dovuto nel caso finito dimensionale a Knott–Smith e Brenier (si veda anche [13] per un analogo risultato negli spazi di Wiener e [14] per un risultato più preciso in dimensione finita).

Teorema 3.5: (Esistenza e unicità della mappa di trasporto ottimale). *Siano* μ^1, $\mu^2 \in \mathscr{P}_2(X)$ *e supponiamo che* μ^1 *si annulli su ogni insieme Gauss nullo. Allora* $\Gamma_o(\mu^1, \mu^2)$ *contiene un solo elemento* $\boldsymbol{\mu}$*, inoltre* $\boldsymbol{\mu}$ *è indotto da una mappa di Borel* r *e* r *è il gradiente, nel senso di Gateaux, di una funzione convessa.*

Indicheremo con $\mathscr{P}_2^r(X)$ il sottoinsieme di tutte le misure in $\mathscr{P}_2(X)$ che si annullano su tutti gli insiemi Gauss nulli, e per ogni $\mu \in \mathscr{P}_2^r(X)$ indicheremo con T_μ^σ la mappa data dal teorema precedente; notiamo che $(T_\mu^\sigma)_\# \mu = \sigma$ e che $\int_X |T_\mu^\sigma - Id|^2\, d\mu = W_2^2(\mu, \sigma)$.

Richiamiamo infine le proprietà fondamentali della convergenza in $(\mathscr{P}_2(X), W_2)$; ricordiamo che un sottoinsieme $\mathscr{K}$ di $\mathscr{P}(X)$ è detto *teso* se

$$\lim_{R \to +\infty} \sup_{\mu \in \mathscr{K}} \mu(X \setminus B_R) = 0, \tag{3.7}$$

e *uniformemente 2-integrabile* se

$$\lim_{R \to +\infty} \sup_{\mu \in \mathscr{K}} \int_{X \setminus B_R} |x|^2 d\mu(x) = 0. \tag{3.8}$$

Teorema 3.6: (Compattezza e convergenza). $(\mathscr{P}_2(X), W_2)$ *è uno spazio metrico completo e separabile. Inoltre, un insieme* $\mathscr{K} \subset \mathscr{P}_2(X)$ *è relativamente compatto se e solo se è teso e uniformemente 2-integrabile. Infine, data una successione* $(\mu_n) \subset \mathscr{P}_2(X)$ *e* $\mu \in \mathscr{P}_2(X)$ *vale:*

$$\lim_{n \to \infty} W_2(\mu_n, \mu) = 0 \iff \begin{cases} \mu_n \text{ converge debolmente a } \mu\,, \\ \{\mu_n\} \text{ è uniformemente 2-integrabile}, \end{cases} \tag{3.9}$$

dove la convergenza debole è intesa nella dualità con le funzioni continue e limitate (a supporto compatto, se X *ha dimensione finita).*

3.2. *La struttura Riemanniana di* $(\mathscr{P}_2(X), W_2)$

In questa sezione analizzeremo la struttura “Riemanniana” di $(\mathscr{P}_2(X), W_2)$, sviluppando un calcolo differenziale che rende rigoroso l'approccio seguito in [22] ed è consistente con la formula di Benamou–Brenier (3.15) di rappresentazione della distanza di Wasserstein.

Il nostro punto di partenza è di derivare la struttura metrico-differenziale dello spazio (i.e. il fibrato tangente e la metrica su di esso) a partire dalla caratterizzazione delle curve assolutamente continue, definite usando solo la distanza (globale) dello spazio; questo punto di vista può essee adottato anche in contesti classici, ad esempio varietà Riemanniane finito-dimensionali immerse isometricamente in uno spazio Euclideo.

Nell'enunciato del seguente teorema faremo uso della classe delle funzioni *cilindriche* in X, della forma $\varphi \circ \pi$ dove $\varphi \in C_c^\infty(\mathbb{R}^d)$ e $\pi : X \to \mathbb{R}^d$ è la mappa naturale indotta da d vettori ortonormali in X. Se $I \subset \mathbb{R}$ è un intervallo aperto, la classe delle funzioni test cilindriche in $X \times I$ è definita analogamente, usando funzioni $\varphi \in C_c^\infty(\mathbb{R}^d \times I)$.

Teorema 3.7: (Curve assolutamente continue e equazione di continuità). *Sia $I \subset \mathbb{R}$ un intervallo aperto, sia $\mu_t : I \to \mathscr{P}_2(X)$ una curva assolutamente continua e sia $|\mu'| \in L^1(I)$ la sua derivata metrica. Allora esistono campi vettoriali di Borel $v_t(x)$ tali che*

$$v_t \in L^2(\mu_t, X), \quad \|v_t\|_{L^2(\mu_t)} \le |\mu'|(t) \qquad \text{for } \mathscr{L}^1\text{-q.o. } t \in I \tag{3.10}$$

e l'equazione di continuità

$$\frac{\partial \mu}{\partial t} + \nabla \cdot (v_t \mu_t) = 0 \qquad \text{in } X \times (0,1) \tag{3.11}$$

vale nel senso delle distribuzioni, i.e.

$$\int_0^1 \int_X \Big(\partial_t \phi(x,t) + \langle v_t(x), \nabla_x \phi(x,t) \rangle \Big)\, d\mu_t(x)\, dt = 0 \tag{3.12}$$

per ogni funzione cilindrica φ in $X \times I$. Inoltre, per $\mathscr{L}^1$-q.o. $t \in I$ v_t appartiene alla chiusura in $L^2(\mu_t, X)$ dello spazio dei gradienti delle funzioni cilindriche in X.

Viceversa, se $\mu_t : I \to \mathscr{P}_2(X)$ soddisfa l'equazione di continuità per un certo campo di Borel v_t tale che $\|v_t\|_{L^2(\mu_t)} \in L^1(I)$, allora μ_t è una curva assolutamente continua e $|\mu'|(t) \le \|v_t\|_{L^2(\mu_t)}$ per $\mathscr{L}^1$-q.o. $t \in I$.

Naturalmente, per una data curva μ_t non vi è un unico campo di velocità v_t per il quale vale l'equazione di continuità (3.11): scelti dei campi w_t tali che $\nabla \cdot (w_t \mu_t) = 0$ i campi $v_t + w_t$ soddisfano ancora (3.11). Tuttavia, le seguenti considerazioni suggeriscono una scelta canonica di v_t:

- i vettori v_t agiscono solo su $\nabla\varphi$, con φ cilindrica in X;

- la seconda implicazione nel teorema precedente mostra che il vettore v_t in $L^2(\mu_t, X)$ maggiora sempre la derivata metrica $|\mu'|(t)$ per $\mathscr{L}^1$-q.o. $t \in I$;
- la prima implicazione mostra che è possibile avere

$$\|v_t\|_{L^2(\mu_t,X)} = |\mu'|(t) \qquad \text{per } \mathscr{L}^1\text{-}q.o.\ t \in I; \tag{3.13}$$

- la linearità di (3.11) e la stretta convessità della norma implicano l'unicità dei vettori soddisfacenti (3.13).

Quindi è naturale considerare come vettori tangenti quei campi di velocità v_t per i quali vale sia (3.11) che (3.13). Questo porta alla seguente definizione.

Definizione 3.8: (Fibrato tangente). *Sia $\mu \in \mathscr{P}_2(X)$. Definiamo*

$$\mathrm{Tan}_\mu \mathscr{P}_2(X) := \overline{\{\nabla\varphi : \ \varphi \ \textit{cilindrica in } X\}}_{L^2(\mu)}.$$

Osservazione 3.9: *Come osservato in precedenza, i vettori in* $\mathrm{Tan}_\mu \mathscr{P}_2(X) \subset L^2(\mu, X)$ *possono anche essere caratterizzati dal seguente principio variazionale: $v \in \mathrm{Tan}_\mu \mathscr{P}_2(X)$ se e solo se*

$$\|v + w\|_{L^2(\mu)} \geq \|v\|_{L^2(\mu)} \quad \forall\, w \in L^2(\mu; X) \text{ tale che } \nabla \cdot (w\mu) = 0. \tag{3.14}$$

Lo spazio tangente nel punto μ, essendo un sottospazio chiuso di $L^2(\mu, X)$, è munito di un prodotto scalare naturale $\langle \cdot, \cdot \rangle_\mu$; tuttavia lo spazio $(\mathscr{P}_2(X), W_2)$ *non* è una varietà Riemanniana infinito-dimensionale (neanche nel caso Euclideo $X = \mathbb{R}^d$), dato che non è possibile definite una mappa esponenziale da un intorno dell'origine in $\mathrm{Tan}_\mu \mathscr{P}_2(X)$ in $\mathscr{P}_2(X)$ che sia un omeomorfismo. Questo è dovuto al fatto che la caratterizzazione delle geodetiche dello spazio dà

$$\exp_\mu(w) := (Id + w)_{\#}\mu$$

se e solo se $x \mapsto x + w(x)$ è il gradiente di una funzione convessa. D'altro canto esistono $v \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ la cui derivata seconda è illimitata dal basso, quindi $x \mapsto x + hv(x)$ non è ottimale per alcun $h > 0$.

Nonostante queste difficoltà manteniamo la terminologia *struttura Riemanniana*, dato che vi sono molte analogie con il caso Riemanniano (queste analogie hanno indotto gli autori di [10] al concetto di *spazio di lunghezza Riemanniano*): la più rilevante è che comunque la distanza di Wasserstein è la distanza indotta dalla metrica Riemanniana sul fibrato tangente. Questo fatto è stato esplicitato per la prima volta da Benamou and Brenier [7] e Otto [22], nel caso in cui $X = \mathbb{R}^d$ e tutte le misure sono assolutamente continue rispetto a $\mathscr{L}^d$.

Teorema 3.10: (Formula di Benamou–Brenier). *Dati μ^0, $\mu^1 \in \mathscr{P}_2(X)$, vale*

$$W_2^2(\mu^0, \mu^1) = \inf \int_0^1 \|v_t\|^2_{L^2(\mu_t)} dt, \tag{3.15}$$

dove l'estremo inferiore è preso tra tutte le curve assolutamente continue $\mu_t : [0,1] \to \to \mathscr{P}_2(X)$ *tali che* $\mu_i = \mu^i$ *for* $i = 0, 1$ *e* v_t *soddisfa l'equazione di continuità* (3.11).

Il prossimo teorema mostra come i vettori tangenti lungo una curva possono essere ottenuti guardando il comportamento asintotico dei piani ottimali di trasporto lungo la curva. È interessante notare che al limite si ottiene un piano $(Id \times v_t)_\# \mu_t$ indotto da una mappa di trasporto anche quando μ_t non è necessariamente regolare.

Teorema 3.11: (Piani ottimali lungo le curve assolutamente continue). *Sia* $\mu_t : I \to \to \mathscr{P}_2(X)$ *una curva assolutamente continua e sia* $v_t \in \mathrm{Tan}_{\mu_t}\mathscr{P}_2(X)$ *il suo campo tangente di velocità. Allora per* $\mathscr{L}^1$*-q.o.* $t \in I$ *vale la seguente proprietà: per ogni scelta di* $\boldsymbol{\mu}_h \in \Gamma_0(\mu_t, \mu_{t+h})$ *si ha*

$$\lim_{h\to 0} \left(\pi^1, \frac{1}{h}(\pi^2 - \pi^1)\right)_\# \boldsymbol{\mu}_h = (Id \times v_t)_\# \, \mu_t \qquad \text{in } \mathscr{P}_2(X \times X) \tag{3.16}$$

e

$$\lim_{h\to 0} \frac{W_2(\mu_{t+h}, (Id + hv_t)_\# \mu_t)}{|h|} = 0. \tag{3.17}$$

In particolare, per $\mathscr{L}^1$*-q.o.* $t \in I$ *tale che* $\mu_t \in \mathscr{P}_2^r(X)$ *si ha*

$$\lim_{h\to 0} \frac{1}{h}(T_{\mu_t}^{\mu_{t+h}} - Id) = v_t \qquad \text{in } L^2(\mu_t, X). \tag{3.18}$$

Le nozioni e i risultati appena presentati conducono in modo naturale al concetto di differenziabilità; tutte le definizioni e risultati elencati nel seguito possono essere adattati al caso di misure non regolari (si veda [4]) ma, per semplicità, li enunciamo solo per misure regolari. Nel caso generale è necessario sostituire mappe di trasporto con piani di trasporto, o anche con oggetti più complessi definiti in X^3.

Definizione 3.12: (Differenziabilità per misure regolari). *Diremo che* $\phi : \mathscr{P}_2(X) \to \to \mathscr{P}_2(X)$ *è differenziabile nel punto* $\mu \in \mathscr{P}_2^r(X)$ *se esiste* $v \in \mathrm{Tan}_\mu \mathscr{P}_2(X)$ *tale che*

$$\lim_{\mu' \to \mu} \frac{\phi(\mu') - \phi(\mu) - \langle v, T_\mu^{\mu'} - Id\rangle_\mu}{W_2(\mu', \mu)} = 0.$$

Il vettore v *se esiste è unico, e verrà chiamato differenziale di* ϕ *in* μ.

Un risultato utile e interessante è il fatto che, per ogni $\sigma \in \mathscr{P}_2(X)$, la funzione $\mu \to W_2^2(\mu, \sigma)$ è differenziabile in ogni $\mu \in \mathscr{P}_2^r(X)$.

Teorema 3.13: (Differenziale di $W_2^2(\cdot, \sigma)$). *La funzione* $\mu \mapsto W^2(\mu, \sigma)$ *è differenziabile per ogni* $\mu \in \mathscr{P}_2^r(X)$. *Il suo differenziale è* $v := 2(Id - T_\mu^\sigma)$.

Si noti che il Teorema 3.11 implica che se μ_t è una curva assolutamente continua e ϕ è una funzione differenziabile in μ_t per $\mathscr{L}^1$-q.o. t, allora la mappa $t \mapsto \phi(\mu_t)$ è $\mathscr{L}^1$-q.o. differenziabile e la sua derivata vale $\langle w_t, v_t \rangle_{\mu_t}$, dove w_t è il differenziale di ϕ in μ_t e v_t è il vettore velocità della curva. In particolare, per ogni $t \mapsto \mu_t \in \mathscr{P}_2^r(X)$ e ogni $\sigma \in \mathscr{P}_2(X)$ la derivata della funzione $t \mapsto W_2^2(\mu_t, \sigma)$ esiste $\mathscr{L}^1$-q.o. e vale $2\langle Id - T_{\mu_t}^{\sigma}, v_t \rangle_{\mu_t}$. Si può mostrare, usando ancora il Teorema 3.11, che questa proprietà di differenziabilità è ancora vera per mappe $t \mapsto \mu_t \in \mathscr{P}_2(X)$ e, in questo caso, la derivata può essere calcolata usando i piani ottimali di trasporto tra μ_t e σ.

3.3. *Sottodifferenziale e flussi gradiente*

Seguendo le stesse idee della precedente sezione si vede subito che la definizione più naturale di sottodifferenziale $\partial^- \phi(\mu)$ di una funzione ϕ nel punto $\mu \in \mathscr{P}_2^r(X)$ è

$$v \in \partial^- \phi(\mu) \quad \text{se e solo se} \quad \phi(\sigma) \geq \phi(\mu) + \langle v, T_\mu^\sigma - Id \rangle_\mu \quad \forall \sigma \in \mathscr{P}_2(X). \tag{3.19}$$

Tuttavia, per ragioni tecniche (ad esempio per considerare anche misure non regolari), è conveniente generalizzare questo concetto.

Nella prossima definizione considereremo piani multipli di trasporto, i.e. misure di probabilità in $X \times X \times X$; indicheremo con $\pi^{i,j} : X^3 \to X^2$ le proiezioni canoniche, per $1 \leq i, j \leq 3$.

Definizione 3.14: (Sottodifferenziale). *Sia $\phi : \mathscr{P}_2(X) \to (-\infty, +\infty]$ e sia $\mu \in D(\phi)$. Un piano di trasporto $\{\mu\} \in \mathscr{P}_2(X \times X)$ appartiene al sottodifferenziale di ϕ in μ, indicato con $\partial^- \phi(\mu)$, se*

(i) $\pi^1_\# \boldsymbol{\mu} = \mu$;

(ii) *per ogni $\sigma \in \mathscr{P}_2(X)$ esiste $\gamma \in \mathscr{P}_2(X \times X \times X)$ tale che $\pi^{1,2}_\# \gamma = \boldsymbol{\mu}$ e $\pi^{1,3}_\# \gamma \in \Gamma_0(\mu, \sigma)$, soddisfacente*

$$\phi(\sigma) \geq \phi(\mu) + \int_{X^3} \langle x_2, x_3 - x_1 \rangle \, d\gamma. \tag{3.20}$$

Si noti che anche se $\mu \in \mathscr{P}_2^r(X)$ la precedente definizione non si riduce a quella data in 3.19, dato che potrebbe succedere che una misura $\boldsymbol{\mu}$ nel sottodifferenziale non sia indotta da una mappa v; la formula (3.20) si riduce alla (3.19) se μ è regolare e $\boldsymbol{\mu} = (Id \times v)_\# \mu$.

Osservazione 3.15: *L'insieme $\partial^- \phi(\mu)$ è un sottoinsieme chiuso e convesso di $\mathscr{P}_2(X \times X)$.*

Esempio 3.16: (Energia interna). Supponiamo che $X = \mathbb{R}^d$ e che $F : [0, \infty) \to (-\infty, +\infty]$ sia tale che $s \mapsto s^d F(s^{-d})$ è convessa e non crescente in $(0, \infty)$ e $F(0) = 0$. Supponiamo anche che F abbia una crescita più che lineare all'infinito (si veda [4] per casi più generali). Definiamo il funzionale energia interna come segue:

$$\phi(\mu) = \int_{\mathbb{R}^n} F(\rho)\, dx \quad \text{per } \mu = \rho \mathscr{L}^d$$

e $\phi(\mu) = +\infty$ se μ non è assolutamente continua rispetto a $\mathscr{L}^d$. Allora ϕ è convesso lungo le geodetiche di $\mathscr{P}_p(X)$.

Come esempi di funzioni F si possono considerare:

(a) $F(t) = \dfrac{t^m}{m-1}$ con $m \geq 1 - \dfrac{1}{d}$ (questa restrizione su m viene dalla convessità di $s \mapsto s^d F(s^{-d})$)

(b) Come caso limite per $m \to 1$ dell'esempio precedente otteniamo $F(t) = t \log t$ (il funzionale entropia).

Indicata con $L_F(t) = tF'(t) - F(t)$ la trasformata di Legendre di F, sotto opportune ipotesi di crescita su F, soddisfatte nei casi (a) e (b), si può mostrare che, per $\mu = \rho \mathscr{L}^d$ nel dominio di ϕ, $\partial^- \phi(\mu)$ è non vuoto se e solo se $L_F(\rho) \in W^{1,1}(X)$, e in tal caso

$$\partial^- \phi(\mu) = (Id \times v)_\# \mu \qquad \text{con} \qquad v := \frac{\nabla L_F(\rho)}{\rho}. \tag{3.21}$$

Si può mostrare che il sottodifferenziale appena definito è un operatore chiuso, nel senso specificato dalla prossima proposizione.

Proposizione 3.17: (Chiusura dell'operatore sottodifferenziale). *Sia $\phi : \mathscr{P}_2(X) \to \to (-\infty, +\infty]$ un funzionale semicontinuo inferiormente e supponiamo che $(\mu_h) \subset D(\phi)$ converga a $\mu \in D(\phi)$ in $\mathscr{P}_2(X)$ e che $\boldsymbol{\mu}_h \in \partial^- \phi(\mu_h)$ convergano a $\boldsymbol{\mu}$, con $\int |x_2|^2 d\boldsymbol{\mu}_h$ limitata. Allora $\boldsymbol{\mu} \in \partial^- \phi(\mu)$.*

Possiamo ora introdurre il concetto di flusso gradiente, basato sulla struttura *differenziale* dello spazio di Wasserstein; vedremo tuttavia che per funzionali convessi lungo le geodetiche questo concetto coincide con quello puramente metrico di curva di massima pendenza.

Definizione 3.18: (Flusso gradiente). *Diremo che $\mu_t \in AC^2_{\mathrm{loc}}((0, +\infty); \mathscr{P}_2(X))$ è una soluzione del flusso gradiente*

$$\dot{\mu}_t \in -\partial^- \phi(\mu_t) \tag{3.22}$$

se per $\mathscr{L}^1$-q.o. $t > 0$ il suo campo tangente di velocità $v_t \in \mathrm{Tan}_{\mu_t} \mathscr{P}_2(X)$ soddisfa $(Id \times -v_t)_\# \mu \in \partial^- \phi(\mu_t)$.

Nel caso dell'Esempio 3.16 abbiamo, grazie alla (3.21), che il flusso gradiente corrisponde a imporre il campo di velocità $v = \nabla L_F(\rho)/\rho$; sostituendo quindi tale campo nell'equazione di continuità si trova l'equazione di diffusione non lineare (con $\mu_t = \rho_t \mathscr{L}^d$)

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla L_F(\rho)).$$

L'ingrediente fondamentale per l'equivalenza tra il punto di vista "metrico" e quello "differenziale" è il seguente lemma.

LEMMA 3.19: (Pendenza locale e sottodifferenziale) *Sia* $\phi : \mathscr{P}_2(X) \to (-\infty, +\infty]$ *un funzionale semicontinuo inferiormente e convesso lungo le geodetiche, soddisfacente*

$$\phi(\mu) \geq -aW_2^2(\mu, \bar{\mu}) + b \qquad \forall \mu \in \mathscr{P}_2(X) \tag{3.23}$$

for qualche $a > 0$, $b \in \mathbb{R}$ *e* $\bar{\mu} \in \mathscr{P}_2(X)$. *Allora*

$$|\partial \phi|^2(\mu) = \min\left\{ \int\limits_{X \times X} |x_2|^2 \, d\boldsymbol{\mu} : \boldsymbol{\mu} \in \partial^- \phi(\mu) \right\} \qquad \forall \mu \in D(\phi),$$

con la convenzione $\min \emptyset = +\infty$.

TEOREMA 3.20: (Curve di massima pendenza coincidono con flussi gradiente). *Sia* $\phi : \mathscr{P}_2(X) \to (-\infty, +\infty]$ *un funzionale semicontinuo inferiormente e convesso lungo le geodetiche, soddisfacente* (3.23). *Allora* $\mu_t : (0, +\infty) \to \mathscr{P}_2(X)$ *è una curva di massima pendenza in base alla Definizione 2.4 se e solo se è un flusso gradiente.*

Grazie a questa equivalenze, per mostrare l'esistenza di flussi gradiente possiamo applicare i Teoremi 2.8, 2.9, 2.13, 2.16 della sezione precedente; in questo caso è conveniente scegliere come topologia debole σ quella indotta dalla convergenza stretta (i.e. nella dualità con le funzioni continue e limitate in X). Ma si può anche fare uso della struttura differenziale dello spazio e passare al limite in una sorta di inclusione sottodifferenziale discreta, come in [16]. Tuttavia, gli argomenti variazionali impliciti nell'approccio di De Giorgi sembrano essere più generali, e consentono per esempio di trattare le metriche W_p o le misure non regolari.

Come abbiamo già osservato, l'unicità delle curve di massima pendenza è ancora un problema aperto sotto la sola ipotesi di convessità lungo le geodetiche. In questo caso, tuttavia, si può far uso della struttura differenziale dello spazio di Wasserstein, e in particolare le proprietà di differenziabilità di $W_2^2(\cdot, \sigma)$ lungo le curve assolutamente continue, per mostrare una proprietà di contrazione che implica l'unicità.

TEOREMA 3.21: (Unicità dei flussi gradiente). *Se* $\phi : \mathscr{P}_2(X) \to (-\infty, +\infty]$ *è un funzionale convesso lungo le geodetiche allora per ogni* $\mu_0 \in \mathscr{P}_2(X)$ *esiste al più un flusso gradiente* $\mu_t : (0, +\infty) \to \mathscr{P}_2(X)$ *che soddisfi* $\lim\limits_{t \downarrow 0} \mu_t = \mu_0$.

Infine, è importante osservare che, sorprendentemente, per una vasta classe di funzionali convessi lungo le geodetiche e studiati nella letteratura (e precisamente l'energia *interna*, l'energia *potenziale*

$$\mu \mapsto \int_X V \, d\mu \qquad \text{con } V \text{ convessa}$$

e l'energia di *interazione*

$$\mu \mapsto \int_{X\times X} W(x-y)\, d\mu(x)d\mu(y) \qquad \text{con } W \text{ convessa}$$

considerate in [20] e tutte le loro possibili combinazioni), anche l'Ipotesi 2.15 è soddisfatta e quindi si ha non solo unicità del flusso gradiente, ma anche stime esplicite dell'errore per lo schema di distretizzazione implicito, grazie al Teorema 2.16 (inoltre l'esistenza non richiede più alcuna impotesi di compattezza, neanche debole, dei sottolivelli di ϕ). Infatti, anche se le geodetiche dello spazio $\mathscr{P}_2(X)$ non soddisfano la disuguaglianza (2.25) (incdentalmente è stato osservato in [22] che vale sempre la disuguaglianza opposta!), è possibile interpolare con diverse curve lungo le quali il funzionale è ancora convesso e vale la (2.25). Date μ_0, $\mu_1 \in \mathscr{P}_2(X)$, le "geodetiche generalizzate" in questione sono

$$\mu_t := \left((1-t)T_\mu^{\mu_0} + tT_\mu^{\mu_1}\right)_{\#}\mu \tag{3.24}$$

nel caso in cui $\mu \in \mathscr{P}_2^r(X)$ (la definizione generale richiede piani di trasporto multipli, come nella definizione di sottodifferenziale).

REFERENCES

[1] M. Agueh, *Existence of solutions to degenerate parabolic equations via the monge-kantorovich theory*, Preprint, (2002).

[2] L. Ambrosio, *Minimizing movements*, Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5) **19** (1995), 191-246. MR 97c:49044

[3] L. Ambrosio - N. Gigli - G. Savaré, *Gradient flows with metric and differentiable structures, and applications to the Wasserstein space*, s. 9 **15** (2004), 327-343.

[4] L. Ambrosio - N. Gigli - G. Savaré, *Gradient flows in metric spaces and in the wasserstein space of probability measures*, ETH Lectures in Mathematics, Birkhäuser, 2005.

[5] L. Ambrosio - P. Tilli, *Topics on "analysis in metric spaces"*, Oxford University Press, 2003.

[6] C. Baiocchi, *Discretization of evolution variational inequalities*, Partial differential equations and the calculus of variations, Vol. I (F. Colombini, A. Marino, L. Modica, and S. Spagnolo, eds.), Birkhäuser Boston, Boston, MA, 1989, pp. 59-92. MR 90m:49002

[7] J.-D. Benamou - Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), no. 3, 375-393. MR 2000m:65111

[8] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, 1973, North-Holland Mathematics Studies, No. 5. Notas de Matemática (50). MR 50#1060

[9] E.A. Carlen - W. Gangbo, *Constrained steepest descent in the 2-wasserstein metric*, Ann. Math., 2 **157** (2003), 807-846.

[10] J. A. Carrillo - R. J. McCann - C. Villani, *Contraction in the 2-wasserstein metric length space*

and thermalization of granular media, Di prossima pubblicazione su Archive for Rational Mechanics and Analysis.

[11] E. De Giorgi, *New problems on minimizing movements*, Boundary Value Problems for PDE and Applications (Claudio Baiocchi and Jacques Louis Lions, eds.), Masson, 1993, pp. 81-98.

[12] E. De Giorgi, Antonio Marino and Mario Tosques, *Problems of evolution in metric spaces and maximal decreasing curve*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8) **68** (1980), no. 3, 180-187. MR 83m:49052

[13] D. Feyel - A. Süleyman Ustünel, *Measure transport on Wiener space and the Girsanov theorem*, C. R. Math. Acad. Sci. Paris **334** (2002), no. 11, 1025-1028. MR 1 913 729

[14] W. Gangbo - R. J. McCann, *The geometry of optimal transportation*, Acta Math. **177** (1996), no. 2, 113-161. MR 98e:49102

[15] J. Heinonen - P. Koskela, *Quasiconformal maps in metric spaces with controlled geometry*, Acta Math. **181** (1998), no. 1, 1-61. MR 99j:30025

[16] R. Jordan - D. Kinderlehrer - F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), no. 1, 1-17 (electronic). MR 2000b:35258

[17] Jürgen Jost, *Nonpositive curvature: geometric and analytic aspects*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1997. MR98g:53070

[18] A. Marino - C. Saccon - M. Tosques, *Curves of maximal slope and parabolic variational inequalities on nonconvex constraints*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **16** (1989), no. 2, 281-330. MR 91d:58035

[19] U. F. Mayer, *Gradient flows on nonpositively curved metric spaces and harmonic maps*, Comm. Anal. Geom. **6** (1998), no. 2, 199-253. MR 99m:58067

[20] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), no. 1, 153-179. MR98e:82003

[21] R. H. Nochetto - G. Savaré - C. Verdi, *A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations*, Comm. Pure Appl. Math. **53** (2000), no. 5, 525-589. MR1 737 503

[22] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26** (2001), no. 1-2, 101-174. MR 2002j:35180

[23] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics **58**, American Mathematical Society, 2003.

Direttore responsabile: Prof. A. BALLIO - Autorizz. Trib. di Roma n. 7269 dell'8-12-1959
«Monograf» - Via Collamarini, 5 - Bologna