Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE

Corso di Laurea in Matematica

ESISTENZA E UNICITÀ DI FUNZIONI "MEASURE-PRESERVING"

Relatore:
Chiar.mo Prof.
Alberto Parmeggiani

Presentata da:
Michele Motta

Sessione autunnale
Anno Accademico 2019-2020

Introduzione

Dati due insiemi X e Y, due σ-algebre $\mathfrak{M}$ ed $\mathfrak{N}$ su X e Y rispettivamente, una misura μ su $(X,\mathfrak{M})$ e una funzione misurabile $f : (X, \mathfrak{M}) \to (Y, \mathfrak{N})$, è possibile definire una misura ν indotta da μ su $(Y, \mathfrak{N})$ attraverso la funzione f ponendo:

$$\nu(N) = \mu(f^{-1}(N)).$$

In tal caso si scrive $\nu = f_{\#}\mu$.
Una questione che sorge quasi spontanea, che comunque è rilevante in diverse aree della Matematica, è quella di stabilire, assegnate due misure μ e ν su opportuni spazi misurabili, sotto quali condizioni esista una funzione f tale che $\nu = f_{\#}\mu$.
L'obiettivo di questa tesi è quello di ripercorrere l'articolo di Robert McCann *"Existence and uniqueness of monotone measure-preserving maps"*, nel quale si dimostra il seguente teorema:

Teorema. *Siano μ e ν due misure di probabilità di Borel su $\mathbb{R}^n$ e supponiamo che μ si annulli sui sottoinsiemi di $\mathbb{R}^n$ aventi dimensione di Hausdorff $n-1$. Allora esiste una funzione convessa $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ tale che $\nabla\psi_{\#}\mu = \nu$. Inoltre, $\nabla\psi$ è univocamente determinato μ-quasi-dappertutto.*

Nel primo capitolo sono presentati alcuni importanti risultati di Analisi Convessa e Analisi Funzionale che intervengono nel seguito. Nel secondo capitolo si mostrerà l'esistenza della funzione "measure-preserving" convessa cercata, nel terzo invece si discute l'unicità del gradiente.

Questo problema, quando μ e ν sono due misure finite, come nel caso che esamineremo, si può interpretare fisicamente come segue: date due distribuzioni di massa, una iniziale e una finale, si cerca un modo di ridistribuire la prima in modo da ottenere la seconda. Una condizione aggiuntiva, abbastanza ragionevole, è quella di richiedere che la funzione che si sta cercando non solo realizzi questa ridistribuzione, ma che in qualche modo minimizzi un certo funzionale, che si può interpretare come l'energia richiesta per ridistribuire la massa.
Non è difficile dedurre[1] che, qualora una funzione siffatta esista, il grafico di una funzione di questo tipo deve essere ciclicamente monotono[2]. Questa proprietà è ben nota in Analisi Convessa: per un Teorema, dovuto a Rockafellar, è noto che essa caratterizza completamente il sottodifferenziale, di una funzione convessa, che a sua volta determina, a meno di una costante additiva, la funzione stessa, in modo completamente analogo a quanto avviene per le funzioni C^1 di una variabile reale.
Alla luce di questi fatti, risulta quindi ragionevole la direzione indicata da McCann, che fa un ampio uso dell'Analisi Convessa, per affrontare tale problema.

[1]Si veda per esempio [10], Sezione 2.1

[2]Per questa e altre definizioni ed enunciati precisi si veda la sezione 1.3.

Indice

Capitolo 1

Premesse

1.1 Funzioni convesse: proprietà di base

In questa sezione preliminare richiamiamo alcune importanti proprietà delle funzioni convesse, che saranno fondamentali in seguito. Cominciamo con alcune definizioni di base:

Definizione 1.1.1 (Insieme convesso). *Sia $C \subseteq \mathbb{R}^n$. Diremo che C è un insieme convesso se per ogni $x, y \in C$ e per ogni $t \in [0,1]$ il punto $tx+(1-t)y$ appartiene a C.*

Dalla definizione segue che l'intersezione di una famiglia qualsiasi di insiemi convessi è ancora un insieme convesso. Si può quindi definire l'insieme convesso generato da un sottoinsieme di $\mathbb{R}^n$ qualsiasi nel modo seguente:

Definizione 1.1.2 (Inviluppo convesso). *Sia $S \subseteq \mathbb{R}^n$ qualsiasi e sia $(C_i)_{i\in I}$, dove I è un insieme di indici, la famiglia di tutti i sottoinsiemi di $\mathbb{R}^n$ convessi contenenti S. Allora l'insieme $\operatorname{conv}(S) := \bigcap_{i\in I} C_i$ contiene S, è convesso e si dirà inviluppo convesso di S.*

Osservazione 1.1.3. *La chiusura $\overline{C}$ di un insieme convesso C è ancora un insieme convesso. Infatti, siano $x, y \in \overline{C}$, $t \in [0,1]$ e consideriamo il punto $tx + (1-t)y$. Siccome $x, y \in \overline{C}$, per definizione esistono due successioni*

$(x_k)_{k\in\mathbb{N}}$ e $(y_k)_{k\in\mathbb{N}}$ in C tali che $\lim_{k\to\infty} x_k = x$ e $\lim_{k\to\infty} y_k = y$. Consideriamo ora la successione definita da $z_k := tx_k + (1-t)y_k$. Per la convessità di C, $z_k \in C$, per ogni $k \in \mathbb{N}$. Inoltre, $\lim_{k\to\infty} tx_k + (1-t)y_k = tx + (1-t)y$. Di conseguenza $tx + (1-t)y \in \overline{C}$.

Osservazione 1.1.4. *Sia $C \subseteq \mathbb{R}^n$ un insieme convesso e $L : \mathbb{R}^n \to \mathbb{R}^m$ una funzione lineare. Allora $L(C)$ è un sottoinsieme convesso di $\mathbb{R}^m$. Infatti, dati $a, b \in L(C)$, allora esistono $x, y \in C$ tali che $L(x) = a$ ed $L(y) = b$. Dunque, per ogni $t \in [0,1]$, si ha che $ta+(1-t)b = tL(x)+(1-t)L(y) = L(tx+(1-t)y)$, quindi, siccome $tx+(1-t)y \in C$, $ta+(1-t)b \in L(C)$, ossia $L(C)$ è convesso.*

Definizione 1.1.5 (Epigrafico)**.** *Sia $A \subseteq \mathbb{R}^n$ ed $f : A \to \mathbb{R} \cup \{\pm\infty\}$ una funzione. L'insieme*

$$\mathrm{epi}(f) = \{(x,\lambda) \in A \times \mathbb{R} \mid \lambda \geq f(x)\}$$

si dice epigrafico della funzione f.

Osservazione 1.1.6. *Dato $S \subseteq \mathbb{R}^{n+1}$ tale che $\{x\} \times [\lambda, +\infty) \subseteq S$ per ogni $(x, \lambda) \in S$, allora è possibile definire $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ in modo che f abbia S come epigrafico. Infatti, basta porre per ogni $x \in \mathbb{R}^n$*

$$f(x) := \inf\{\lambda \in \mathbb{R} \mid (x,\lambda) \in S\}.$$

Per convenzione, poniamo $\inf \emptyset = +\infty$.

Definizione 1.1.7 (Funzione convessa)**.** *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione. Diremo che f è una funzione convessa se il suo epigrafico è un insieme convesso di $\mathbb{R}^{n+1}$.*

Definizione 1.1.8 (Dominio effettivo)**.** *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa. Chiameremo dominio effettivo di f l'insieme*

$$\mathrm{dom}(f) := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}.$$

Osservazione 1.1.9 (Convessità del dominio effettivo)**.** *Il dominio effettivo di una funzione convessa $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ è un insieme convesso, infatti è l'immagine dell'epigrafico di f, che è un insieme convesso, attraverso la proiezione $\pi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, $\pi(x, \lambda) = x$, che è una funzione lineare.*

Definizione 1.1.10 (Funzione convessa propria)**.** *Sia* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *una funzione convessa e* $\mathrm{dom}(f)$ *il suo dominio effettivo. Allora* f *si dice propria se* $\mathrm{dom}(f)$ *non è vuoto e* $f(x) > -\infty$ *per ogni* $x \in \mathrm{dom}(f)$.

Da un punto di vista geometrico, la definizione precedente equivale a richiedere che l'epigrafico di f non contenga rette verticali.
Ricordiamo che una funzione $f : \mathbb{R}^n \to \mathbb{R}$ si dice inferiormente semicontinua se, per ogni successione $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ tale che $\lim_{k\to\infty} x_k = x$, si ha

$$f(x) \leq \liminf_{k\to\infty} f(x_k).$$

Proposizione 1.1.11. *Sia* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *una funzione. Allora* f *è inferiormente semicontinua se e solo se l'epigrafico di* f *è un sottoinsieme chiuso di* $\mathbb{R}^{n+1}$.

Dimostrazione. Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione inferiormente semicontinua. L'inferiore semicontinuità di f in un punto x può essere riformulata nel modo seguente: date due successioni $(\lambda_k)_{k\in\mathbb{N}}$ e $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}$ e $\mathbb{R}^n$ rispettivamente tali che $\lambda_k \xrightarrow[k\to\infty]{} \lambda$, $x_k \xrightarrow[k\to\infty]{} x$ e $f(x_k) \leq \lambda_k$ per ogni $k \in \mathbb{N}$, allora $f(x) \leq \lambda$. Ma questa condizione è equivalente alla chiusura dell'epigrafico di f. Infatti, sia $((x_k, \lambda_k))_{k\in\mathbb{N}}$ una successione nell'epigrafico di f tale che $(x_k, \lambda_k) \xrightarrow[k\to\infty]{} (x, \lambda)$. Si ha che $f(x_k) \leq \lambda_k$ per ogni $k \in \mathbb{N}$, dunque $f(x) \leq \lambda$ per la semicontinuità inferiore, cioè (x, λ) è ancora un punto dell'epigrafico di f, da cui la sua chiusura. Il viceversa è analogo. □

Definizione 1.1.12 (Chiusura di una funzione)**.** *Data* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *una funzione convessa propria, la funzione* $\bar{f}$ *che ha come epigrafico l'insieme* $\overline{\mathrm{epi}(f)}$*, che sarà convessa e propria, si dirà chiusura della funzione* f*. Una funzione si dice chiusa se* $f = \bar{f}$.

Proposizione 1.1.13. *Sia* $(f_i)_{i\in I}, f_i : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *una famiglia di funzioni convesse chiuse, dove* I *è un insieme qualsiasi di indici. Allora, la funzione* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *definita da*

$$f(x) := \sup\{f_i(x) \mid i \in I\}$$

è convessa e chiusa.

Dimostrazione. Basta osservare che

$$\mathrm{epi}(f) = \bigcap_{i \in I} \mathrm{epi}(f_i)$$

e che l'intersezione qualsiasi di insiemi chiusi e convessi è chiusa e convessa. □

Proposizione 1.1.14. *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa e sia $A \subseteq \mathbb{R}^n$ aperto. Se f assume valori finiti su tutto A, allora f è continua in A.*

Dimostrazione. Si veda [2], Teorema 10.1, pag. 82. □

Proposizione 1.1.15. *Sia $C \subseteq \mathbb{R}^n$ un sottoinsieme convesso. Allora, indicando con ∂C la frontiera di C e con $\mathcal{H}^{n-1}$ la misura di Hausdorff $(n-1)$-dimensionale, per ogni $x \in \partial C$ esiste un intorno U di x tale che si abbia $\mathcal{H}^{n-1}(\partial C \cap U) < +\infty$.*

Dimostrazione. Si veda [9]. □

1.2 Coniugata di una funzione convessa

In questa sezione introduciamo la nozione di funzione coniugata di una funzione convessa. Nel seguito vedremo che ciò sarà molto importante per dimostrare alcune proprietà fondamentali delle funzioni convesse.
Prima di proseguire oltre, è il caso di fissare alcune notazioni. Ogni iperpiano affine di $\mathbb{R}^{n+1}$ si può rappresentare come il luogo degli zeri di una funzione affine di $\mathbb{R}^{n+1}$, ossia, una funzione del tipo

$$\begin{aligned} h_{\mathbf{b}} : \mathbb{R}^n \times \mathbb{R} &\to \mathbb{R} \\ (x, \mu) &\mapsto \langle x, b \rangle - \beta\mu - \gamma \end{aligned}$$

dove $\mathbf{b} = (b, \beta, \gamma) \in \mathbb{R}^n \times \{0, 1\} \times \mathbb{R}$ è un vettore fissato. Chiameremo iperpiani affini verticali di $\mathbb{R}^{n+1}$ gli iperpiani con $\beta = 0$, non verticali quelli con $\beta = 1$. Ogni iperpiano affine non verticale divide $\mathbb{R}^{n+1}$ in due semispazi chiusi, ossia:

- $\{(x, \mu) \mid \mu \geq \langle x, b\rangle - \gamma\} = \mathrm{epi}(h_{\mathbf{b}})$;
- $\{(x, \mu) \mid \mu \leq \langle x, b\rangle - \gamma\}$.

Chiameremo il primo *semispazio superiore* dell'iperpiano e il secondo *semispazio inferiore* dell'iperpiano.
Si ha il seguente teorema:

Teorema 1.2.1. *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa, propria e chiusa. Sia $\mathscr{H}$ la collezione degli iperpiani affini non verticali i cui semispazi superiori contengono* $\mathrm{epi}(f)$*. Allora l'epigrafico di f è l'intersezione di tutti i semispazi superiori degli iperpiani di $\mathscr{H}$.*
Equivalentemente, $f(x) = \sup\{h_{\mathbf{b}}(x) \mid h_{\mathbf{b}} \leq f\}$.

Dimostrazione. Si veda [2], Teorema 12.1, pag. 102. □

Quindi, data una funzione convessa f, ha senso definire l'insieme F^* di tutte le coppie $(x^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}$ tali che la funzione affine definita da $h(x) = \langle x, x^*\rangle - \mu^*$ è maggiorata da f. Più precisamente, abbiamo che $(x^*, \mu^*) \in F^*$ se e solo se:

$$\mu^* \geq \sup\{\langle x, x^*\rangle - f(x) \mid x \in \mathbb{R}^n\}.$$

Dunque, per l'Osservazione 1.1.5, F^* risulta essere l'epigrafico della funzione f^* definita da

$$f^*(x^*) := \sup\{\langle x, x^*\rangle - f(x) \mid x \in \mathbb{R}^n\}.$$

Chiameremo f^* la funzione coniugata della funzione f. Osserviamo che, per come è definita, f^* è definita puntualmente come l'estremo superiore di una collezione di funzioni affini del tipo $g(x^*) = \langle x, x^*\rangle - \mu$ dove $(x, \mu) \in \mathrm{epi}(f)$. Di conseguenza, per la Proposizione 1.1.11, f^* è ancora una funzione convessa ed è chiusa se f è chiusa. Si verifica inoltre che $(f^*)^* = f$.

Osservazione 1.2.2 (Disuguaglianza di Fenchel)**.** *Dalla definizione di f^*, segue subito che*

$$\langle x, x^*\rangle \leq f(x) + f^*(x^*) \qquad \textit{per ogni } x, x^* \in \mathbb{R}^n.$$

Più in generale, se consideriamo l'insieme delle coppie di funzioni (f, g) *che soddisfano la disuguaglianza*

$$\langle x, x^* \rangle \leq f(x) + g(x^*) \qquad \textit{per ogni } x, x^* \in \mathbb{R}^n$$

e si suppone che la coppia (f, g) *minimizzi la quantità di destra, nel senso che se anche per* (f_1, g_1) *vale la stessa disuguaglianza e si ha che* $f_1 \leq f$ *e* $g_1 \leq g$*, allora necessariamente* $f_1 = f$ *e* $g_1 = g$*, allora deve essere*

$$g(x^*) \geq \sup\{\langle x, x^* \rangle - f(x) \mid x \in \mathbb{R}^n\} = f^*(x^*),$$

ossia, la coppia (f, f^*) *non solo soddisfa la disuguglianza, ma minimizza la quantità di destra.*

1.3 Sottodifferenziale e monotonia ciclica

In questa sezione richiamiamo alcuni importanti concetti relativi alle funzioni convesse.

Definizione 1.3.1 (Sottogradiente e sottodifferenziale). *Sia* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *una funzione convessa,* x *un punto di* $\mathbb{R}^n$ *e* $\langle \cdot, \cdot \rangle$ *il prodotto scalare standard su* $\mathbb{R}^n$*. Se per un certo* $y \in \mathbb{R}^n$ *vale*

$$f(z) \geq \langle y, z - x \rangle + f(x) \tag{1.1}$$

per ogni $z \in \mathbb{R}^n$*, allora* y *si dirà un sottogradiente per* f *in* x*. Chiameremo sottodifferenziale di* f *l'insieme*

$$\partial f := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid x, y \textit{ soddisfano la } (1.1)\ \forall z \in \mathbb{R}^n\}.$$

Inoltre, indicheremo con $\partial f(x)$ *l'insieme dei sottogradienti di* f *nel punto* x*. Diremo che la funzione* f *è sottodiferenziabile in* x *se* $\partial f(x)$ *è non vuoto.*

Si vede immediatamente che se $f(z) < +\infty$ per qualche z e $f(x) = +\infty$, allora deve essere $\partial f(x) = \emptyset$, poiché in tal caso la (1.1) non può essere soddisfatta quando $f(z)$ assume valori finiti. Vale anche il viceversa, ossia:

Proposizione 1.3.2. *Sia* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *una funzione convessa e* $x \in \mathbb{R}^n$. *Allora* $\partial f(x)$ *è un sottoinsieme di* $\mathbb{R}^n$ *non vuoto e limitato se e solo se* $x \in \operatorname{int}(\operatorname{dom}(f))$.

Dimostrazione. Si veda [2], Teorema 23.4, pag. 217. □

Quando x è un punto interno di $\operatorname{dom}(f)$ ed $f(x) \in \mathbb{R}$, il sottogradiente di f può essere interpretato geometricamente come segue: sia $y \in \partial f(x)$ e sia h la funzione definita da $h(z) = \langle y, z - x\rangle + f(x)$, $z \in \mathbb{R}^n$. Allora il grafico di h è un iperpiano affine di $\mathbb{R}^{n+1}$ che supporta (cioè "sta sotto") l'epigrafico di f nel punto $(x, f(x))$.
Vediamo un semplice esempio:

Esempio 1.3.3. *Sia* $f : \mathbb{R} \to \mathbb{R}, f(x) = |x|$. *La funzione* f *è convessa e derivabile ovunque tranne in* $x = 0$. *Tuttavia il sottogradiente di* f *in* $x = 0$ *è non vuoto: infatti, è dato dall'intervallo* $[-1, 1]$.

Notiamo che $\partial f(x)$ è un insieme convesso. Infatti, se $y_1, y_2 \in \partial f(x)$, dalla (1.1) segue che:

$$t\,(f(z) - f(x)) \geq t\langle y_1, z - x\rangle$$
$$(1 - t)\,(f(z) - f(x)) \geq (1 - t)\langle y_2, z - x\rangle$$

per ogni $t \in [0, 1]$ e per ogni $z \in \mathbb{R}^n$, da cui segue

$$\begin{aligned}\langle (1-t)y_1 + ty_2, z - x\rangle &= (1-t)\langle y_2, z - x\rangle + t\langle y_1, z - x\rangle \leq \\ &\leq t(f(z) - f(x)) + (1-t)(f(z) - f(x)) = f(z) - f(x)\end{aligned}$$

che implica la convessità di $\partial f(x)$. Inoltre, $\partial f(x)$ è un insieme chiuso. Infatti, sia $y \in \overline{\partial f(x)}$, allora esiste una successione $(y_k)_{k\in\mathbb{N}}$ in $\partial f(x)$ tale che $\lim_{k\to\infty} y_k = y$. Dunque, per ogni z, si ha

$$\langle y, x - z\rangle = \langle y - y_k, x - z\rangle + \langle y_k, x - z\rangle \leq |y - y_k||x - z| + f(z) - f(x).$$

Siccome $|y - y_k| \xrightarrow[k\to\infty]{} 0$ e $|x - z|$ non dipende da k, per ogni $\varepsilon > 0$ e per k maggiore di un certo k_ε dipendente da ε, si ha

$$\langle y, x - z\rangle \leq f(z) - f(x) + \varepsilon,$$

quindi, per l'arbitrarietà di ε, possiamo concludere che $y \in \partial f(x)$, cioè $\partial f(x)$ è chiuso.

Il sottodifferenziale di una funzione convessa è strettamente legato al differenziale classico. Ricordiamo qui alcuni risultati:

Proposizione 1.3.4. *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa. Se f è differenziabile in $x \in \mathbb{R}^n$, allora $\nabla f(x)$ è l'unico elemento di $\partial f(x)$. Viceversa, se $\partial f(x)$ ha un unico elemento, allora f è differenziabile in x e $\nabla f(x)$ coincide con l'unico elemento di $\partial f(x)$.*

Dimostrazione. Si veda [2], Teorema 25.1, pag. 242. □

Corollario 1.3.5. *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa. Se f è differenziabile in $x \in \mathbb{R}^n$, allora $x \in \mathrm{int}(\mathrm{dom}(f))$.*

Dimostrazione. Segue dalla Proposizione precedente e dalla Proposizione 1.3.2. □

Teorema 1.3.6. *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa e propria. Allora, indicando con D l'insieme dei punti in cui f è differenziabile, D è denso in $\mathrm{dom}(f)$ e $\mathrm{dom}(f) \setminus D$ ha al più dimensione di Hausdorff $n - 1$. Infine, la mappa $x \mapsto \nabla f(x), x \in D$, è continua.*

Dimostrazione. Si veda [9]. □

Vediamo ora qualche risultato che lega il sottodifferenziale di una funzione convessa f alla sua coniugata f^*.

Proposizione 1.3.7. *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa e propria e siano $x, x^*, z \in \mathbb{R}^n$. Allora sono equivalenti:*

(i) $x^* \in \partial f(x)$;

(ii) $\langle z, x^* \rangle - f(z)$ *ha massimo in* $z = x$;

(iii) $f(x) + f^*(x^*) \leq \langle x, x^* \rangle$;

(iv) $f(x) + f^*(x^*) = \langle x, x^* \rangle$.

Dimostrazione. Per definizione di sottogradiente, $x^* \in \partial f(x)$ se e solo se

$$f(z) \geq \langle x^*, z - x \rangle + f(x) \quad \forall z \in \mathbb{R}^n,$$

che equivale a

$$\langle x^*, x \rangle - f(x) \geq \langle x^*, z \rangle - f(z) \quad \forall z \in \mathbb{R}^n,$$

osservando che i due membri della disuguaglianza coincidono per $z = x$, si ha l'equivalenza tra (i) e (ii).
Ora, per definizione si ha $f^*(x^*) = \sup\{\langle z, x^* \rangle - f(z) \mid z \in \mathbb{R}^n\}$. Se $\langle z, x^* \rangle - f(z)$ ha massimo in $z = x$, allora dalla definizione segue che $f(x) + f^*(x^*) = \langle x, x^* \rangle$. Viceversa, se $f(x) + f^*(x^*) = \langle x, x^* \rangle$, allora $\langle z, x^* \rangle - f(z)$ ha massimo in $z = x$, quindi (ii) e (iv) sono equivalenti.
La (iii) e la (iv) sono equivalenti per la disuguaglianza di Fenchel (vedi Osservazione 1.2.2). □

Questa proposizione ha un'importantissima conseguenza.

Teorema 1.3.8. *Sia* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *una funzione convessa, propria e chiusa. Allora* ∂f *è un insieme chiuso di* $\mathbb{R}^n \times \mathbb{R}^n$.

Dimostrazione. Siano $(x_k)_{k \in \mathbb{N}}$ e $(x_k^*)_{k \in \mathbb{N}}$ due successioni in $\mathbb{R}^n$ tali che $x_k^* \in \partial f(x_k)$ per ogni k e $x_k \xrightarrow[k \to +\infty]{} x_0$ e $x_k^* \xrightarrow[k \to +\infty]{} x_0^*$. Vogliamo mostrare che $x_0^* \in \partial f(x_0)$. Come abbiamo visto nella Proposizione 1.1.9, la funzione f è inferiormente semicontinua. Combinando questo fatto con il punto (iii) della proposizione precedente, si ha

$$f(x_0) + f^*(x_0^*) \leq \liminf_{k \to +\infty} f(x_k) + \liminf f(x_k^*) \leq \liminf_{k \to +\infty} \langle x_k, x_k^* \rangle = \langle x_0, x_0^* \rangle,$$

che sempre per la proposizione precedente equivale a $x_0^* \in \partial f(x_0)$. □

Qui e nel seguito, diremo che una funzione $\rho : \mathbb{R}^n \to \mathbb{R}^n$ è *multivoca* intendendo che, per ogni $x \in \mathbb{R}^n$, $\rho(x)$ è un sottoinsieme non vuoto di $\mathbb{R}^n$. Il grafico di ρ sarà l'insieme $\Gamma_\rho = \{(x, y) \mid x \in \mathbb{R}^n \text{ e } y \in \rho(x)\}$. Osserviamo allora che il sottodifferenziale di una funzione convessa f può essere inteso anche come il grafico della funzione multivoca $x \mapsto \partial f(x)$.
Ricordiamo ora una proprietà molto importante che, come vedremo, caratterizza completamente il sottodifferenziale di una funzione convessa.

Definizione 1.3.9 (Monotonia ciclica)**.** *Un sottoinsieme $S \subset \mathbb{R}^n \times \mathbb{R}^n$ si dice ciclicamente monotono se ogni insieme finito di punti di $(x_i, y_i) \in S$, $i = 1, 2, \ldots, k$ verifica la seguente disuguaglianza:*

$$\langle y_1, x_2 - x_1\rangle + \langle y_2, x_3 - x_2\rangle + \cdots + \langle y_k, x_1 - x_k\rangle \leq 0, \tag{1.2}$$

dove $\langle \cdot, \cdot \rangle$, come prima, è il prodotto scalare standard di $\mathbb{R}^n$.
Una funzione multivoca $\rho : \mathbb{R}^n \to \mathbb{R}^n$ si dirà ciclicamente monotona se il suo grafico lo è.

Questa definizione, che a prima vista può sembrare un po' macchinosa, quando S è un insieme finito è equivalente a richiedere che le x e le y siano accoppiate in modo da massimizzare la somma $\sum_{i=1}^n \langle x_i, y_i\rangle$.

Teorema 1.3.10. *Sia $k \in \mathbb{N}$. Comunque scelti $2k$ punti arbitrari, non necessariamente distinti, $x_i, y_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, k$, allora esiste una permutazione $\sigma \in S_k$ tale che le coppie $(x_{\sigma(i)}, y_i)$ formino un sottoinsieme ciclicamente monotono di $\mathbb{R}^n \times \mathbb{R}^n$.*

Dimostrazione. Siano $x_i, y_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, k$ come nell'enunciato. Poiché l'insime S_k è finito, l'insieme

$$\left\{ \sum_{i=1}^{k} \langle y_i, x_{\sigma(i)}\rangle \mid \sigma \in S_k \right\}$$

è finito, quindi ha massimo. Senza perdita di generalità, possiamo assumere che la permutazione identica realizzi tale massimo.
Mostriamo che l'insieme così riordinato è ciclicamente monotono, ossia, che

ogni sottoinsieme T delle coppie (x_i, y_i), $i = 1, 2, \dots, k$ rispetta la disuguaglianza (1.2). Sia quindi $T = \{(a_1, b_1), \dots, (a_m, b_m)\}$, $m \leq k$ uno di questi sottoinsiemi. Per ogni $j = 1, 2, \dots, m$ esiste uno e un solo $i \in \{1, \dots, k\}$ tale che $(a_j, b_j) = (x_i, y_i)$. Poniamo $\tau(j) = i$. Possiamo completare τ in modo che sia una permutazione qualunque di S_k. Quindi, abbiamo che

$$\begin{aligned}
&\sum_{j=1}^{m-1} \langle b_j, a_{j+1} - a_j \rangle + \langle b_m, a_1 - a_m \rangle = \\
&= \sum_{j=1}^{m-1} \langle y_{\tau(j)}, x_{\tau(j+1)} - x_{\tau(j)} \rangle + \langle y_{\tau(m)}, x_{\tau(1)} - x_{\tau(m)} \rangle = \\
&= \sum_{j=1}^{m-1} \langle y_{\tau(j)}, x_{\tau(j+1)} \rangle + \langle y_{\tau(m)}, x_{\tau(1)} \rangle - \sum_{j=1}^{m} \langle y_{\tau(j)}, x_{\tau(j)} \rangle = \\
&= \sum_{j=1}^{m-1} \langle y_{\tau(j)}, x_{\tau(j+1)} \rangle + \langle y_{\tau(m)}, x_{\tau(1)} \rangle + \sum_{j=m+1}^{k} \langle y_{\tau(j)}, x_{\tau(j)} \rangle - \sum_{j=1}^{k} \langle y_{\tau(j)}, x_{\tau(j)} \rangle.
\end{aligned}$$

Ora, osservando che

$$\sum_{j=1}^{m-1} \langle y_{\tau(j)}, x_{\tau(j+1)} \rangle + \langle y_{\tau(m)}, x_{\tau(1)} \rangle + \sum_{j=m+1}^{k} \langle y_{\tau(j)}, x_{\tau(j)} \rangle$$

è della forma $\sum_{i=1}^{k} \langle y_i, x_{\sigma(i)} \rangle$, con $\sigma \in S_k$ opportuna, e che $\sum_{j=1}^{k} \langle y_j, x_j \rangle = \sum_{j=1}^{k} \langle y_{\tau(j)}, x_{\tau(j)} \rangle$ e ricordando come avevamo riordinato le coppie (x_i, y_i), possiamo concludere che

$$\sum_{j=1}^{m-1} \langle b_j, a_{j+1} - a_j \rangle + \langle b_m, a_1 - a_m \rangle \leq 0,$$

da cui la monotonia ciclica. □

Osservazione 1.3.11. *Sia f una funzione convessa, ∂f il suo sottodifferenziale e siano $(x_i, y_i) \in \partial f$, $i = 1, 2, \dots, n$. Allora, sostituendo nella (1.1) $x = x_i$, $y = y_i$ e $z = x_{i+1}$, si ottiene che*

$$\begin{aligned}
&\langle y_1, x_2 - x_1 \rangle + \langle y_2, x_3 - x_2 \rangle + \cdots + \langle y_n, x_1 - x_n \rangle \leq \\
\leq &[f(x_2) - f(x_1)] + [f(x_3) - f(x_2)] + \cdots + [f(x_1) - f(x_n)] = 0,
\end{aligned}$$

ovvero, ∂f è un insieme ciclicamente monotono.

Vediamo ora che è vero anche il viceversa dell'osservazione precedente, ossia, ogni insieme ciclicamente monotono è contenuto nel sottodifferenziale di una funzione convessa.

Teorema 1.3.12 (Rockafellar). *Sia $\rho : \mathbb{R}^n \to \mathbb{R}^n$ una funzione multivoca. Allora, il grafico di ρ è un sottoinsieme di $\mathbb{R}^n \times \mathbb{R}^n$ ciclicamente monotono se e solo se è contenuto nel sottodifferenziale di una funzione convessa, propria e chiusa.*

Dimostrazione. Se il grafico di ρ è contenuto nel sottodifferenziale di una funzione convessa f, chiaramente deve essere ciclicamente monotono poiché, come abbiamo visto, il sottodifferenziale di f è ciclicamente monotono, quindi, a maggior ragione, lo sono i suoi sottoinsiemi.
Viceversa, sia ρ una funzione multivoca e supponiamo che il suo grafico Γ sia ciclicamente monotono. Allora, per ogni coppia $(x_0, y_0) \in \Gamma$ (supponiamo che Γ sia non vuoto) possiamo definire f come segue:

$$\begin{aligned} f(x) = \sup\{ &\langle x - x_m, y_m\rangle + \langle x_m - x_{m-1}, y_{m-1}\rangle + \cdots + \langle x_1 - x_0, y_0\rangle \mid \\ &\mid (x_i, y_i) \in \Gamma \text{ per ogni } i = 1, 2, \dots, m,\ m \in \mathbb{N} \cup \{0\}\}. \end{aligned}$$

Poiché f è l'estremo superiore di una famiglia di applicazioni affini, una per ogni m-upla di (x_i, y_i), per la Proposizione 1.1.10 f è una funzione convessa e chiusa. Inoltre, poiché Γ è ciclicamente monotona, $f(x_0) \leq 0$. In particolare, prendendo come unico punto (x_0, y_0) nella definizione di $f(x_0)$, si ottiene che deve essere $f(x_0) \geq 0$, da cui $f(x_0) = 0$, quindi il dominio effettivo di f non è vuoto. Inoltre, se $f(x) < +\infty$, poiché $\langle x - x_0, y_0\rangle > -\infty$ per ogni $x \in \mathbb{R}^n$, f assume solo valori finiti sul suo dominio effettivo. Questo ci dice che f è una funzione propria.
Ora, siano x e y tali che $y \in \rho(x)$. Vogliamo mostrare che $y \in \partial f(x)$, da cui segue che il grafico di ρ è contenuto nel sottodifferenziale di f. Sarà sufficiente verificare che per ogni $\lambda < f(x)$ e $z \in \mathbb{R}^n$ si ha

$$f(z) > \lambda + \langle z - x, y\rangle.$$

Sia dunque $\lambda < f(x)$. Per definizione di f, esistono m punti del grafico di ρ tali che

$$\lambda < \langle x - x_m, y_m \rangle + \langle x_m - x_{m-1}, y_{m-1} \rangle + \cdots + \langle x_1 - x_0, y_0 \rangle \tag{1.3}$$

Sempre per definizione di f, per ogni $z \in \mathbb{R}^n$ si ha

$$f(z) \geq \langle z - x, y \rangle + \langle x - x_m, y_m \rangle + \langle x_m - x_{m-1}, y_{m-1} \rangle + \cdots + \langle x_1 - x_0, y_0 \rangle.$$

Per la (1.3), segue che

$$f(z) > \langle z - x, y \rangle + \lambda,$$

che è esattamente quanto voluto. $\square$

Osservazione 1.3.13. *Ogni sottoinsieme S di $\mathbb{R}^n \times \mathbb{R}^n$ si può ottenere come grafico di una funzione multivoca. Infatti, indicando con π_1 e π_2 la proiezione sul primo e sul secondo $\mathbb{R}^n$ rispettivamente, basta porre:*

$$\begin{aligned} \rho : \ & \pi_1(S) \to \mathbb{R}^n \\ & x \mapsto \pi_2((\{x\} \times \mathbb{R}^n) \cap S). \end{aligned}$$

Teorema 1.3.14. *Il sottodifferenziale di una funzione convessa, propria e chiusa è una funzione ciclicamente monotona massimale, ossia, il suo grafico non è contenuto propriamente nel grafico di alcun'altra funzione ciclicamente monotona. Inoltre, ogni funzione convessa, propria e chiusa è completamente determinata a meno di una costante additiva dal proprio sottodifferenziale.*

Dimostrazione. Si veda [2], Teorema 24.9, pag. 239. $\square$

1.4 La topologia debole*

In questa sezione richiamiamo alcuni importanti risultati di Analisi funzionale che interverranno nel seguito.

Sia X spazio vettoriale reale e $\|\cdot\|$ una norma su X. Se X è uno spazio metrico completo rispetto alla metrica $d(x, y) := \|x - y\|$, $x, y \in X$, diremo

che X è uno *spazio di Banach*. Se X, Y sono spazi vettoriali normati, una funzione $F : X \to Y$ si dice *limitata* se esiste $C > 0$ tale che $\|F(x)\| \leq C\|x\|$ per ogni $x \in X$. In particolare se X, Y sono spazi normati, una funzione $F : X \to Y$ lineare è continua se e solo se è limitata (si veda [3], Teorema 5.4, pag. 96).
Chiameremo *spazio duale di X* l'insieme

$$X^* := \{f : X \to \mathbb{R} \mid f \text{ lineare e continua}\}.$$

Su X^* si può definire la norma

$$\|f\| := \inf\{\mu \in \mathbb{R} \mid |f(x)| \leq \mu\|x\|\} = \sup\{|f(x)| \mid x \in X, \|x\| \leq 1\}.$$

Infine indichiamo con $X^{**} := (X^*)^*$. Se X ha dimensione infinita, si può dimostrare che la palla chiusa di raggio unitario $\overline{B} = \{x \in X \mid \|x\| \leq 1\}$ **non** è compatta (si veda [5], Teorema 6.5). Questo ci dice che, in generale, gli insiemi compatti in uno spazio vettoriale di dimensione qualsiasi, come lo sono i principali spazi di funzioni, sono difficili da caratterizzare rispetto a quanto avviene, per esempio, in $\mathbb{R}^n$. Per aggirare questo ostacolo, si costruiscono delle topologie strettamente contenute nella topologia di partenza, ossia con meno aperti, con l'obiettivo di avere più insiemi compatti.
A tal fine, richiamiamo il seguente risultato di topologia.

Proposizione 1.4.1. *Sia X un insieme e sia $\mathcal{B} \subseteq \mathcal{P}(X)$. Allora $\mathcal{B}$ è una base per una topologia su X se e solo se:*

- $\displaystyle\bigcup_{B \in \mathcal{B}} B = X$;
- *per ogni $B_1, B_2 \in \mathcal{B}$ e per ogni $x \in B_1 \cap B_2$ esiste $B_3 \in \mathcal{B}$ tale che $x \in B_3 \subseteq B_1 \cap B_2$.*

In tal caso, gli aperti della topologia di cui $\mathcal{B}$ è base sono tutti e soli gli insiemi che si ottengono con unioni qualsiasi di elementi di $\mathcal{B}$.

Sfruttando questo fatto, per definire una nuova topologia su X si può ricorrere alla seguente costruzione: si individua la famiglia di insiemi $\mathcal{A} =$

$(A_i)_{i\in I}$ che si vuole che siano aperti nella nuova topologia e si considera la famiglia di insiemi

$$\mathcal{B} = \{\cap_{j\in J} A_j \mid A_j \in \mathcal{A} \text{ per ogni } J \subset I, J \text{ insieme finito}\}, \tag{1.4}$$

formata da tutte le intersezioni finite di insiemi di $\mathcal{A}$. Questa famiglia, per definizione, rispetta i criteri della Proposizione 1.3.1, quindi definisce una topologia per X. Vediamo subito due esempi concreti.

Definizione 1.4.2 (Topologia debole)**.** *Sia X uno spazio di Banach e X^* il suo duale. La topologia debole su X è la topologia meno fine che rende continui tutti i funzionali $f \in X^*$. Più precisamente, per ogni $f \in X^*$ consideriamo la famiglia di sottoinsiemi $\mathcal{A}_f = \{f^{-1}(\omega) \mid \omega \text{ aperto di } \mathbb{R}\}$ di X e sia $\mathcal{A}_X = \cup_{f\in X^*}\mathcal{A}_f$. Sia $\mathcal{B}_X$ definito come in (1.4) con $\mathcal{A} = \mathcal{A}_X$. Per la Proposizione 1.3.1 $\mathcal{B}_X$ è una base di una topologia su X, che diremo topologia debole di X.*

Definizione 1.4.3 (Topologia debole*)**.** *Sia X uno spazio di Banach e X^* il suo duale. Per ogni $x \in X$, sia $\varphi_x : X^* \to \mathbb{R}$ definita da $\varphi_x(f) = f(x)$, per ogni $f \in X^*$.*
La topologia debole su X^* è la topologia meno fine che rende continui tutti i funzionali φ_x. Più precisamente, per ogni $x \in X$ consideriamo la famiglia di sottoinsiemi $\mathcal{A}_x = \{\varphi_x^{-1}(\omega) \mid \omega \text{ aperto di } \mathbb{R}\}$ di X^* e sia $\mathcal{A}_{X^*} = \cup_{x\in X}\mathcal{A}_x$. Sia $\mathcal{B}_{X^*}$ definito come in (1.4) con $\mathcal{A} = \mathcal{A}_{X^*}$. Per la Proposizione 1.3.1, $\mathcal{B}_{X^*}$ è una base di una topologia su X^*, che diremo topologia debole* di X^*.*

Queste due topologie hanno molte importanti proprietà, tra le quali: sono Hausdorff, sono strettamente meno fini nella topologia indotta dalla norma quando X ha dimensione infinita. Qui ci limiteremo a richiamare solo quelle che saranno rilevanti per il seguito. Per un approfondimento di questo tema di veda [5], capitolo 3.

Proposizione 1.4.4. *Sia X uno spazio di Banach. Allora X^* con la topologia debole* è uno spazio di Hausdorff. In particolare, ogni successione in X^*, se ha limite, allora ha un unico limite.*

Dimostrazione. Siano $f_1, f_2 \in X^*$ distinti. Allora esite $x \in X$ tale che $f_1(x) \neq f_2(x)$. Senza perdita di generalità, possiamo supporre che $f_1(x) < f_2(x)$. Quindi esiste $\alpha \in \mathbb{R}$ tale che $f_1(x) < \alpha < f_2(x)$. Poniamo:

$$A_1 = \varphi_x^{-1}((-\infty, \alpha)) \quad A_2 = \varphi_x^{-1}((\alpha, +\infty)).$$

Per come è definita la topologia debole*, A_1 e A_2 sono aperti e, per come sono definiti, sono chiaramente disgiunti e $f_1 \in A_1$, $f_2 \in A_2$.
L'unicità del limite discende dalla proprietà di Hausdorff. $\square$

Proposizione 1.4.5. *Sia X uno spazio di Banach e sia X^* il suo duale. Una successione $(x_n)_{n\in\mathbb{N}}$ in X converge debolmente a $x \in X$ se e solo se per ogni $f \in X^*$ si ha $\lim_{n\to\infty} f(x_n) = f(x)$.*

Dimostrazione. Per definizione, la topologia debole è la topologia meno fine che rende continue tutte le $f \in X^*$, quindi se $x_n \xrightarrow[n\to\infty]{} x$ debolmente, allora $f(x_n) \xrightarrow[n\to\infty]{} f(x)$.
Viceversa, supponiamo che $(x_n)_{n\in\mathbb{N}}$ sia una successione in X e che $f(x_n) \xrightarrow[n\to\infty]{} f(x)$ per ogni $f \in X^*$. Sia U un intorno di x in X nella topologia debole. Non è restrittivo supporre che sia un aperto di base, ossia, che sia della forma $U = \bigcap_{i=1}^m f_i^{-1}(\omega_i)$. Siccome $f_i(x_n) \xrightarrow[n\to\infty]{} f_i(x)$, per ogni $i = 1, 2, \ldots, m$, esiste un $N_i \in \mathbb{N}$ tale che $f_i(x_n) \in \omega_i$ per ogni $n > N_i$. Quindi, se $n > \max\{N_i \mid i = 1, 2, \ldots, m\}$, $x_n \in \bigcap_{i=1}^m f_i^{-1}(\omega_i) = U$. Per l'arbitrarietà di U, segue che $x_n \xrightarrow[n\to\infty]{} x$ debolmente. $\square$

Proposizione 1.4.6. *Sia X uno spazio di Banach e sia X^* il suo duale. Una successione $(f_n)_{n\in\mathbb{N}}$ in X^* converge debolmente* a $f \in X$ se e solo se per ogni $x \in X$ si ha $\lim_{n\to\infty} \varphi_x(f_n) = \varphi_x(f)$, ossia, se le f_n convergono puntualmente a f.*

Dimostrazione. La dimostrazione è uguale alla precedente sostituendo le φ_x a f_i. $\square$

Vediamo ora un importante teorema il quale ci dice che, come volevamo, nella topologia debole* di X^* abbiamo speranze di avere degli insiemi compatti abbastanza buoni.

Teorema 1.4.7 (Banach-Alaoglu)**.** *La palla chiusa di raggio unitario* $\overline{B}_{X^*} = \{f \in X^* \mid \|f\| \leq 1\}$ *di* X^* *è compatta nella topologia debole*.*

Dimostrazione. Si veda [5], Teorema 3.16, pag. 66. □

1.5 Misure a valori complessi

In questa sezione richiamiamo alcuni concetti di teoria della misura che interverrano nel seguito.
Per cominciare, dato X un insieme qualsiasi, diremo che una collezione non vuota $\mathfrak{M}$ di sottoinsiemi di X è una *σ-algebra* su X se è chiusa per unioni numerabili e passaggio al complementare, ossia se, data una succesione $\{E_k\}_{k\in\mathbb{N}}$ in $\mathfrak{M}$, allora $\bigcup_{k=1}^{\infty} E_k \in \mathfrak{M}$ e per ogni $E \in \mathfrak{M}$ si ha $E^c \in \mathfrak{M}$, dove $E^c := X \setminus E$. Dato X un insieme qualsiasi e $\mathfrak{M}$ una σ-algebra su X, la coppia $(X, \mathfrak{M})$ si dirà uno spazio misurabile.
Dato uno spazio misurabile $(X, \mathfrak{M})$, chiameremo misura su $(X, \mathfrak{M})$ una funzione $\mu : \mathfrak{M} \to \mathbb{C}$ che sia σ-additiva, cioè, per ogni successione $\{E_k\}_{k\in\mathbb{N}}$ in $\mathfrak{M}$ di insiemi a due a due disgiunti, si ha $\mu(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k)$, con $\sum_{k=1}^{\infty} \mu(E_k)$ serie convergente in $\mathbb{C}$. Osserviamo che, per ogni permutazione $\tau : \mathbb{N} \xrightarrow[su]{1-1} \mathbb{N}$, si ha $\bigcup_{k=1}^{\infty} E_k = \bigcup_{k=1}^{\infty} E_{\tau(k)}$, e quindi $\sum_{k=1}^{\infty} \mu(E_k) = \sum_{k=1}^{\infty} \mu(E_{\tau(k)})$, da cui deduciamo che la serie $\sum_{k=1}^{\infty} \mu(E_k)$ converge assolutamente.
Diremo che una misura μ su $(X, \mathfrak{M})$ è positiva se $\mu(E) \in [0, +\infty]$, per ogni $E \in \mathfrak{M}$ (in questo caso è ammesso che la serie $\sum_{k=1}^{\infty} \mu(E_k)$ della definizione sopra diverga).

Definizione 1.5.1. *Data* $\mathfrak{M}$ *una* σ*-algebra sull'insieme* X *e* μ *una misura su* $\mathfrak{M}$*, chiameremo la funzione d'insieme*

$$\begin{aligned} |\mu| : &\mathfrak{M} \to \mathbb{R} \\ &E \mapsto \sup\left\{ \sum_{i=1}^{\infty} |\mu(E_i)| \;\middle|\; \{E_i\}_{i\in\mathbb{N}} \text{ partizione di } E \right\} \end{aligned}$$

variazione totale di μ.

È immediato verificare che, se μ è una misura positiva, allora $|\mu| = \mu$. Inoltre, per la variazione totale di una misura μ valgono i seguenti risultati:

Teorema 1.5.2. *La variazione totale* $|\mu|$ *di una misura* μ *a valori complessi, definita su una* σ*-algebra* $\mathfrak{M}$, *è una misura positiva su* $\mathfrak{M}$.

Dimostrazione. Si veda [3], Teorema 6.2, pag. 117. □

Teorema 1.5.3. *Se* μ *è una misura a valori complessi sulla* σ*-algebra* $\mathfrak{M}$ *definita su* X, *allora* $|\mu|(X) < +\infty$.

Dimostrazione. Si veda [3], Teorema 6.4, pag.118. □

Osservazione 1.5.4. *Se* λ *e* μ *sono due misure a valori complessi definite sulla stessa* σ*-algebra* $\mathfrak{M}$ *su* X, *definendo:*

$$(\lambda + \mu)(E) = \lambda(E) + \mu(E)$$
$$(c\mu)(E) = c\mu(E)$$

per ogni $E \in \mathfrak{M}$ *e per ogni* $c \in \mathbb{C}$, *si verifica che* $\lambda + \mu$ *e* $c\mu$ *sono ancora misure su* $\mathfrak{M}$. *Quindi l'insieme delle misure complesse su* $\mathfrak{M}$ *è uno spazio vettoriale. Inoltre, si può verificare che* $\|\mu\| := |\mu|(X)$ *è una norma per tale spazio vettoriale.*

1.6 Il teorema di rappresentazione di Riesz

In questa sezione, X indicherà sempre uno spazio topologico di Hausdorff localmente compatto.

Chiameremo σ-algebra di Borel su X la σ-algebra generata dagli aperti di X e chiameremo boreliani i suoi elementi. Diremo che una misura μ su un insieme X è di Borel se è definita su una σ-algebra contenente la σ-algebra di Borel.

Diremo poi che una misura μ di Borel definita sulla σ-algebra $\mathfrak{M}$ è regolare se per $|\mu|$ valgono:

- $|\mu|(E) = \inf\{|\mu|(A) \mid E \subseteq A, A \text{ aperto}\}$ per ogni $E \in \mathfrak{M}$;

- se μ è una misura a valori complessi, chiederemo che sia
$|\mu|(E) = \sup\{|\mu|(K) \mid K \subseteq E, K \text{ compatto}\}$ per ogni $E \in \mathfrak{M}$;

- se μ è una misura positiva, chiederemo che sia
$\mu(E) = \sup\{|\mu|(K) \mid K \subseteq E, K \text{ compatto}\}$ per ogni aperto E e per ogni $E \in \mathfrak{M}$ tale che $\mu(E) < +\infty$.

Una misura μ Borel regolare che assume valori finiti sui compatti si dirà misura di Radon.

Indicheremo con $C_c(X)$ lo spazio delle funzioni continue da X in $\mathbb{R}$ a supporto compatto contenuto in X. $C_c(X)$ si intenderà sempre equipaggiato della norma

$$\|f\| := \sup_{x\in X} |f(x)| \tag{1.5}$$

e della topologia indotta da tale norma. Analogamente si può definire $C_c(X, \mathbb{C})$ lo spazio delle funzioni continue da X in $\mathbb{C}$ a supporto compatto. Cambiando il valore assoluto in (1.5) con il modulo di un numero complesso si ottiene una norma per $C_c(X, \mathbb{C})$.
Diremo inoltre che una funzione continua da X a valori in $\mathbb{R}$ o in $\mathbb{C}$ tende a zero all'infinito se per ogni $\varepsilon > 0$ esiste un compatto $K \subseteq X$ tale che $|f(x)| < \varepsilon$ per ogni $x \in X \setminus K$. Indicheremo con $C_0(X)$ e con $C_0(X, \mathbb{C})$ l'insieme delle funzioni continue da X a valori reali e complessi rispettivamente che tendono a zero all'infinito. La norma definita in (1.5) è una norma anche per $C_0(X)$ e $C_0(X, \mathbb{C})$ (sostituendo, per quest'ultimo, il valore assoluto con il modulo di un numero complesso), rispetto alla quale sono spazi di Banach.
Ricordiamo infine che $\overline{C_c(X)} = C_0(X)$ e $\overline{C_c(X, \mathbb{C})} = C_0(X, \mathbb{C})$ (si veda [3], Teorema 3.17, pag 70).
Si può verificare senza difficoltà che, se μ è una misura Borel regolare a valori

complessi, allora il funzionale

$$\begin{aligned}\Lambda : C_0(X, \mathbb{C}) &\to \mathbb{C} \\ f &\mapsto \int_X f \mathrm{d}\mu\end{aligned}$$

è lineare e limitato con norma al più $|\mu|(X)$, dunque anche continuo per quanto osservato in precedenza. Il seguente fondamentale teorema ci dice che in realtà questi sono tutti e soli i funzionali lineari definiti su $C_0(X, \mathbb{C})$.

Teorema 1.6.1 (di rappresentazione di Riesz)**.** *Sia X uno spazio topologico di Hausdorff e localmente compatto e $\Lambda \in C_0(X, \mathbb{C})^*$. Allora esiste una σ-algebra $\mathfrak{M}$ su X e un'unica misura di Radon μ a valori complessi definita su $\mathfrak{M}$ tali che $\Lambda(f) = \int_X f(x)\mathrm{d}\mu$, per ogni $f \in C_0(X, \mathbb{C})$. Inoltre, $\|\Lambda\| = |\mu|(X)$.*

Dimostrazione. Si veda [6], Teorema 7.3.6, pag. 201. □

1.7 Il teorema di Krein-Milman

In uno spazio vettoriale di dimensione finita, non è difficile mostrare che, se un punto x appartiene all'inviluppo convesso di un certo insieme S, allora x si può scrivere come combinazione lineare convessa di un numero finito di elementi di S. Si pensi, per esempio, a un poliedro convesso: ogni punto del poliedro si può ottenere come combinazione lineare convessa dei vertici. Di seguito riportiamo una formulazione del teorema di Krein-Milman, che costituisce un analogo del risultato appena citato, ma in dimensione qualsiasi.

Definizione 1.7.1. *Sia X uno spazio vettoriale normato e sia K un suo sottoinsieme convesso. Un punto $z \in K$ si dice estremale se non può essere scritto come combinazione lineare convessa di altri punti distinti di K, ossia se*

$$z = tx + (1-t)y \text{ per } t \in [0,1] \text{ e } x, y \in K \implies z = x = y \text{ oppure } t \in \{0,1\}.$$

Teorema 1.7.2 (di Krein-Milman)**.** *Sia X uno spazio vettoriale normato e K un suo sottoinsieme convesso e compatto. Allora K è la chiusura dell'inviluppo convesso dei suoi punti estremali.*

Dimostrazione. Si veda [4], Teorema 3.23, pag. 75. □

Capitolo 2

Esistenza della funzione “measure-preserving”

Prima di procedere oltre, ricordiamo alcune definizioni di teoria della misura. Sia μ una misura positiva su $(X, \mathfrak{M})$, dove $\mathfrak{M}$ è una σ-algebra che contiene i boreliani, si definisce *supporto* di μ l'insieme

$$\operatorname{supp}(\mu) := \{x \in X \mid \text{per ogni intorno aperto } N_x \text{ di } x \text{ si ha } \mu(N_x) > 0\}.$$

Diremo che μ positiva è di probabilità su X se $\mu(X) = 1$. Denoteremo con $\mathcal{P}(X)$ l'insieme delle misure positive di probabilità di Borel su X. Ricordiamo che, se X è uno spazio di Hausdorff localmente compatto e a base numerabile, allora ogni misura di probabilità di Borel è anche regolare, quindi di Radon (si veda [6], Proposizione 7.2.3, pag. 190).
Infine, data una misura di Borel γ su $\mathbb{R}^n \times \mathbb{R}^n$, le misure $\mu(M) = \gamma(M \times \mathbb{R}^n)$ e $\nu(M) = \gamma(\mathbb{R}^n \times M)$, per ogni boreliano $M \subseteq \mathbb{R}^n$, si dicono rispettivamente *marginale sinistra* e *marginale destra* di γ. Denoteremo l'insieme delle misure di Borel con marginale sinistra μ e marginale destra ν con $\Gamma(\mu, \nu)$.

L'obiettivo di questo capitolo è quello di dimostrare il seguente risultato:

Teorema 2.0.1. *Siano μ e ν due misure di probabilità positive su $\mathbb{R}^n$ e supponiamo che $\mu(M) = 0$ per ogni insieme di Borel M che abbia dimensione di Hausdorff al più $n-1$. Allora esiste una funzione convessa, propria e chiusa $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ il cui gradiente $\nabla\psi$ sia tale che $\nabla\psi_{\#}\mu = \nu$.*

L'idea che sta alla base della dimostrazione è quella di costruire una misura di probabilità γ su $\mathbb{R}^n \times \mathbb{R}^n$ che abbia supporto ciclicamente monotono e le cui misure marginali siano μ e ν. A questo punto, il teorema di Rockafellar ci dice che il supporto di γ è contenuto nel sottodifferenziale di una funzione convessa ψ. Mostreremo infine che la ψ così costruita è la funzione cercata.

2.1 Esistenza della misura congiunta

Iniziamo con tre lemmi che ci permetteranno di dimostrare l'esistenza della misura γ con marginali μ e ν.
Come prima cosa, ci serve che $\mathcal{P}(\mathbb{R}^n)$ sia un sottoinsieme compatto di $C_0(\mathbb{R}^n)^*$ rispetto alla topologia debole*. Di conseguenza, siccome la topologia debole* è Hausdorff (Proposizione 1.4.4), sarà anche chiuso.
Per mostrare ciò, osserviamo innanzitutto che $\mathcal{P}(\mathbb{R}^n)$, per il Teorema di rappresentazione di Riesz (vedi Teorema 1.6.1), è contenuto nella palla chiusa $B = \{\mu \in C_0(\mathbb{R}^n)^* \mid \|\mu\| \leq 1\}$, che per il Teorema di Banach-Alaoglu (si veda Teorema 1.4.7) è compatta. Siccome $C_0(\mathbb{R}^n)$ è separabile (ossia, contenente un sottoinsieme denso e numerabile), possiamo applicare il seguente risultato:

Teorema 2.1.1. *Sia X uno spazio di Banach separabile. Allora esiste una metrica che induce la topologia debole* sull'insieme*

$$B_{X^*} = \{f \in X^* \mid \|f\| \leq 1\}.$$

Dimostrazione. Si veda [5], Teorema 3.28, pag. 74. □

Questo ci dice che compattezza e compattezza per successioni in B sono equivalenti. Ma si può dimostrare che $\mathcal{P}(\mathbb{R}^n)$ è compatto per successioni:

Proposizione 2.1.2. *Sia $\{\mu_k\}_{k\in\mathbb{N}}$ una successione in $\mathcal{P}(\mathbb{R}^n)$. Allora esiste una sottosuccessione $\{\mu_{k_j}\}_{j\in\mathbb{N}}$ di $\{\mu_k\}_{k\in\mathbb{N}}$ e una misura $\mu \in \mathcal{P}(\mathbb{R}^n)$ tali che $\mu_{k_j} \xrightarrow{j\to+\infty} \mu$ debolmente*.*

Dimostrazione. Si veda [8], Teorema 1.41, pag. 66. □

Dunque, riassumendo:

Osservazione 2.1.3. *Il sottoinsieme $\mathcal{P}(\mathbb{R}^n) \subseteq C_0(\mathbb{R}^n)^*$ è compatto e chiuso rispetto alla topologia debole*.*

Nel seguito indicheremo con δ_x, $x \in \mathbb{R}^n$, la misura di Dirac centrata in x, ossia, per ogni sottoinsieme $A \subseteq \mathbb{R}^n$, $\delta_x(A) = 1$ se $x \in A$, $\delta_x(A) = 0$ se $x \notin A$.

Proposizione 2.1.4. *L'insieme $\mathcal{P}(\mathbb{R}^n)$ è un insieme convesso e i suoi punti estremali (si veda la Definizione 1.7.1) sono la misura identicamente nulla e le misure del tipo δ_x, $x \in \mathbb{R}^n$.*

Dimostrazione. Vediamo prima la convessità di $\mathcal{P}(\mathbb{R}^n)$. Siano $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^n)$ e sia $t \in [0,1]$. Definiamo $\nu(M) = t\mu_1(M) + (1-t)\mu_2(M)$, per ogni boreliano M. Mostriamo che ν è una misura positiva e che $\nu(\mathbb{R}^n) = 1$. Siccome il vuoto è un boreliano, $\nu(\emptyset) = t\mu_1(\emptyset) + (1-t)\mu_2(\emptyset) = 0$. Inoltre, ν assume solo valori non negativi, poiché, dalla definizione, si ha $\nu(M) = t\mu_1(M) + (1-t)\mu_2(M) \geq 0$. La σ-additività segue direttamente dalla σ-additività di μ_1 e μ_2: sia $(M_k)_{k\in\mathbb{N}}$ una successione di boreliani a due a due disgiunti, si ha

$$\nu\left(\bigcup_{k=1}^{\infty} M_k\right) = t\mu_1\left(\bigcup_{k=1}^{\infty} M_k\right) + (1-t)\mu_2\left(\bigcup_{k=1}^{\infty} M_k\right) =$$
$$= t\left(\sum_{k=1}^{\infty} \mu_1(M_k)\right) + (1-t)\left(\sum_{k=1}^{\infty} \mu_2(M_k)\right) =$$
$$\sum_{k=1}^{\infty}[t\mu_1(M_k) + (1-t)\mu_2(M_k)] = \sum_{k=1}^{\infty} \nu(M_k).$$

Infine, $\nu(\mathbb{R}^n) = 1$ segue da $\nu(\mathbb{R}^n) = t\mu_1(\mathbb{R}^n) + (1-t)\mu_2(\mathbb{R}^n) = t + 1 - t = 1$.

Vediamo ora i punti estremali di $\mathcal{P}(\mathbb{R}^n)$. La misura identicamente nulla ζ è sicuramente uno di questi. Infatti, se $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^n)$ e $t \in [0,1]$ sono tali che $\zeta = t\mu_1 + (1-t)\mu_2$, se μ_1 non è identicamente nulla, per la regolarità, allora esiste un boreliano M tale che $\mu_1(M) > 0$ e $t\mu_1(M) + (1-t)\mu_2(M) = 0$. Ora, se $0 < t \leq 1$, allora $t\mu_1(M) > 0$, quindi, essendo $(1-t)\mu_2(M) \geq 0$, si ha che $t\mu_1(M) + (1-t)\mu_2(M) > 0$, che è assurdo. Quindi deve essere $t = 0$, ossia $\zeta = \mu_2$. Allo stesso modo, si vede che in alternativa solo l'altro caso $\zeta = \mu_1$ è possibile.

Verifichiamo che anche le misure che sono punti estremali non nulle di $\mathcal{P}(\mathbb{R}^n)$ sono della forma δ_x, $x \in \mathbb{R}^n$.
Sia μ una misura estremale non identicamente nulla di $\mathcal{P}(\mathbb{R}^n)$. Esiste quindi un insieme di Borel A tale che $\mu(A) > 0$. Vogliamo mostrare che $\mu(A^c) = 0$, dove $A^c := \mathbb{R}^n \setminus A$. Se per assurdo $\mu(A^c) > 0$, allora possiamo definite le misure

$$\mu_1(M) = \frac{\mu(M \cap A)}{\mu(A)}, \quad \mu_2(M) = \frac{\mu(M \cap A^c)}{\mu(A^c)}.$$

Ricordando che $\mu(A^c) = 1 - \mu(A)$ e che $0 < \mu(A) \leq 1$, possiamo concludere che

$$\begin{aligned} \mu(A)\mu_1(M) + (1-\mu(A))\mu_2(M) &= \\ = \mu(A)\frac{\mu(M \cap A)}{\mu(A)} + \mu(A^c)\frac{\mu(M \cap A^c)}{\mu(A^c)} &= \\ = \mu(M \cap A) + \mu(M \cap A^c) = \mu(M). \end{aligned}$$

Poiché avevamo supposto che μ fosse estremale, deve necessariamente essere $\mu(A) = 1$ e di conseguenza $\mu(A^c) = 0$.
Questo ci dice anche che se $\mu(C), \mu(A) > 0$, allora $\mu(A \cap C) = 1$, poiché $\mu((A \cap C)^c) = \mu(A^c \cup C^c) \leq \mu(A^c) + \mu(C^c) = 0$. Inoltre, se C è un sottoinsieme proprio di A con $\mu(C), \mu(A) > 0$, allora $\mu(C) = 1 = \mu(A)$. Ora, sia A un boreliano tale che $\mu(A) = 1$ e supponiamo che per ogni suo sottoinsieme proprio C si abbia $\mu(C) = 0$. Se A contiene almeno due punti, allora è possibile costruirne una partizione $(E_i)_{i \in I}$ non banale, con I un insieme di

indici finito. Ma allora l'additività di μ ci dice che

$$\mu(A) = \sum_{i \in I} \mu(E_i) = 0,$$

che è chiaramente assurdo.
Quindi, riassumendo, se A è un insieme contenente almeno due punti e tale che $\mu(A) = 1$, allora contiene almeno un sottoinsieme proprio C tale che $\mu(C) = 1$. Ma allora, definendo sui boreliani di misura 1 la relazione d'ordine parziale data da

$$N \leq M \iff M \subseteq N, \quad M, N \text{ boreliani tali che } \mu(M) = 1 = \mu(N),$$

per quanto osservato in precedenza, il Lemma di Zorn ci dice che esiste un elemento massimale, che non può che essere un singoletto $\{x\}$, per un opportuno $x \in \mathbb{R}^n$, altrimenti conterrebbe un sottoinsieme proprio di misura 1 per quanto già osservato. Infine x appartiene a tutti gli insiemi di misura 1, poiché l'intersezione di due insiemi di misura 1 è non vuota. Dunque μ è una delta di Dirac. □

Veniamo finalmente al primo dei tre lemmi:

Lemma 2.1.5 (Densità delle misure discrete)**.** *Sia $\mathcal{P}(\mathbb{R}^n) \subset C_0(\mathbb{R}^n)^*$ l'insieme delle misure di probabilità di Borel su $\mathbb{R}^n$. Allora l'insime*

$$\mathcal{D} = \left\{ \frac{1}{m} \sum_{i=1}^{m} \delta_{x_i} \mid x_1, \ldots, x_m \in \mathbb{R}^n, m \in \mathbb{N} \right\}$$

è denso in $\mathcal{P}(\mathbb{R}^n)$ rispetto alla topologia debole.*

Dimostrazione. Per quanto già detto all'inizio del capitolo (si veda in particolare l'Osservazione 2.1.3) $\mathcal{P}(\mathbb{R}^n)$ è un sottoinsieme compatto della palla metrica chiusa di raggio unitario di $C_0(\mathbb{R}^n)^*$. Per la Proposizione 2.1.4 è anche un insieme convesso i cui punti estremali sono le misure di Dirac δ_x. Questo ci dice che, per il teorema di Krein-Milman (si veda Teorema 1.7.2) le combinazioni convesse

$$\sum_{i=1}^{m} t_i \delta_{x_i}, \quad \sum_{i=1}^{m} t_i = 1, \quad t_i > 0 \text{ per } i = 1, 2, \ldots, m, \quad m \in \mathbb{N}$$

sono dense in $\mathcal{P}(\mathbb{R}^n)$. Siccome $\mathbb{Q}$ è denso in $\mathbb{R}$, possiamo restringerci al caso in cui i t_i sono razionali. Indicando con m il minimo comun denominatore dei t_i e ammettendo le ripetizioni di δ_{x_i} nella sommatoria, otteniamo la densità di $\mathcal{D}$. □

Lemma 2.1.6. *Sia $(\gamma_k)_{k\in\mathbb{N}}$ una successione in $\mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ convergente debolmente* a $\gamma \in C_0(\mathbb{R}^n \times \mathbb{R}^n)^*$. Se γ_k ha supporto ciclicamente monotono per ogni $k \in \mathbb{N}$, allora anche γ ha supporto ciclicamente monotono.*

Dimostrazione. Supponiamo per assurdo che $\operatorname{supp}(\gamma)$ non sia ciclicamente monotono. Allora esistono m punti $(x_i, y_i) \in \operatorname{supp}(\gamma)$, $i = 1, 2, \ldots, m$, tali che

$$\langle y_1, x_2 - x_1\rangle + \langle y_2, x_3 - x_2\rangle + \cdots + \langle y_m, x_1 - x_m\rangle > 0. \tag{2.1}$$

Inoltre, poiché la (2.1) è una disuguaglianza stretta e poiché il prodotto scalare è continuo nei due argomenti, esistono degli aperti U_i tali che $(x_i, y_i) \in U_i$ e sostituendo $(u_i, v_i) \in U_i$ al posto di (x_i, y_i) nella (2.1) il maggiore stretto continua a valere. Siccome gli (x_i, y_i) sono dei punti del supporto di γ, $\gamma(U_i) > 0$, per ogni $i = 1, 2, \ldots, m$. Ma allora per ogni $i = 1, 2, \ldots, m$ esiste k_i tale che anche $\gamma_k(U_i) > 0$ per $k \geq k_i$. Infatti, se per un certo i fosse $\gamma_k(U_i) = 0$ per ogni $k \in \mathbb{N}$, allora per ogni $f \in C_0(\mathbb{R}^n \times \mathbb{R}^n)$ si avrebbe

$$\int_{U_i} f \mathrm{d}\gamma_k = 0 \quad \text{per ogni } k \in \mathbb{N},$$

quindi, passando al limite per $k \to \infty$, si ottiene

$$\int_{U_i} f \mathrm{d}\gamma = 0$$

qualsiasi sia $f \in C_0(\mathbb{R}^n \times \mathbb{R}^n)$, che è chiaramente assurdo dal momento che $\gamma(U_i) > 0$. Quindi, poiché lo stesso ragionamento si può ripetere per qualunque intorno aperto di (x_i, y_i) contenuto in U_i, per ogni $i = 1, 2, \ldots, m$, i punti (x_i, y_i) sono contenuti nel supporto delle γ_k, per k abbastanza grande. Ma questo è assurdo perché avevamo supposto che il loro supporto fosse ciclicamente monotono.

□

Prima di passare all'ultimo lemma, è opportuno fare un paio di osservazioni preliminari.

Osservazione 2.1.7. *Ogni funzione $f : X \to \mathbb{R}$ si può vedere come una funzione di due variabili $f : X \times Y \to \mathbb{R}$ indipendente dal secondo argomento, ossia $f(x, y_1) = f(x, y_2)$ per ogni $y_1, y_2 \in Y$. Dove non risulti ambiguo, indicheremo indifferentemente con f sia la funzione di una sola variabile, sia quella in due variabili.*

Osservazione 2.1.8. *Siano X e Y due spazi topologici e γ una misura di Borel su $X \times Y$. Allora μ è la marginale sinistra di γ se e solo se, per ogni funzione $f : X \times Y \to \mathbb{R}$ che sia γ-misurabile, μ-misurabile rispetto al primo argomento e indipendente dal secondo argomento si ha*

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} f \mathrm{d}\gamma = \int_{\mathbb{R}^n} f \mathrm{d}\mu. \tag{2.2}$$

Infatti, se nella (2.2) si sostituisce $\mathbb{1}_M$ a f si ottiene $\gamma(M \times \mathbb{R}^n) = \mu(M)$, per ogni boreliano $M \subseteq \mathbb{R}^n$. Se invece è vero che $\gamma(M \times \mathbb{R}^n) = \mu(M)$, allora si ha

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \mathbb{1}_M \mathrm{d}\gamma = \gamma(M \times \mathbb{R}^n) = \mu(M) = \int_{\mathbb{R}^n} \mathbb{1}_M \mathrm{d}\mu.$$

Per linearità dell'integrale si estende il risultato alle funzioni semplici. Per il teorema di Beppo Levi e per densità delle funzioni semplici si estende a tutte le funzioni positive e infine, di nuovo per linearità dell'integrale, a una funzione qualsiasi separando parte positiva e parte negativa.

Lemma 2.1.9. *Sia $(\gamma_k)_{k\in\mathbb{N}}$ una successione in $\mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ convergente debolmente* a $\gamma \in C_0(\mathbb{R}^n \times \mathbb{R}^n)^*$. Indicando con μ_k e ν_k le misure marginali di γ_k, se le μ_k tendono debolmente* a μ e le ν_k tendono debolmente* a ν, allora le marginali di γ sono μ e ν.*

Dimostrazione. Se fosse vero che ogni $f \in C_0(\mathbb{R}^n)$ può essere estesa a una funzione $f \in C_0(\mathbb{R}^n \times \mathbb{R}^n)$ indipendente dalla seconda variabile, per quanto detto nell'Osservazione 2.1.8, avremmo concluso, perché, per ipotesi, per ogni $k \in \mathbb{N}$ avremmo

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} f \mathrm{d}\gamma_k = \int_{\mathbb{R}^n} f \mathrm{d}\mu_k,$$

e passando al limite per $k \to \infty$ otterremmo la (2.2). Questo però è vero solo per $f \equiv 0$. Infatti, per ogni $x \in X$ si avrebbe $f(x,y) = \lim_{\|y\| \to \infty} f(x,y) = \lim_{\|(x,y)\| \to \infty} f(x,y) = 0$.
Tuttavia, si può aggirare questo ostacolo. Indicando con $\overline{\mathbb{R}^n}$ la compattificazione di Alexandrov di $\mathbb{R}^n$ (si veda pagg. 89 e 90 in [7]), è sempre possibile estendere una $f \in C_0(\mathbb{R}^n)$ a una funzione $f \in C(\overline{\mathbb{R}^n} \times \overline{\mathbb{R}^n})$ che non dipende dalla seconda variabile. Infatti basta porre $f(x,y) = f(x)$ e $f(\infty, y) = \lim_{\|x\| \to \infty} f(x) = 0$. Ora, siccome $(\gamma_k)_{k \in \mathbb{N}}$ è una successione nella palla di raggio unitario di $C(\overline{\mathbb{R}^n} \times \overline{\mathbb{R}^n})^*$, che per il teorema di Banach-Alaoglu è compatta nella topologia debole*, a meno di passare a una sottosuccessione, possiamo supporre che $\gamma_k \overset{k \to \infty}{\longrightarrow} \gamma'$ debolmente*. Siccome i compatti negli spazi di Hausdorff sono chiusi, anche γ' appartiene alla palla unitaria ed è una misura positiva, quindi $\gamma(\overline{\mathbb{R}^n} \times \overline{\mathbb{R}^n}) = 1$.
Per concludere, ci basterà mostrare che γ' è concentrata in $\mathbb{R}^n \times \mathbb{R}^n$, ossia che per ogni boreliano $M \subseteq \overline{\mathbb{R}^n} \times \overline{\mathbb{R}^n}$ si ha $\gamma'(M) = \gamma'(M \cap \mathbb{R}^n \times \mathbb{R}^n)$. Infatti, per quanto abbiamo osservato in precedenza, la sua misura marginale sinistra sarà μ e, con un argomento analogo, la sua marginale destra sarà ν. Inoltre, per l'unicità del limite debole*, γ' deve coincidere con γ.
Per il teorema di Riesz si ha:

$$\gamma'(\mathbb{R}^n \times \overline{\mathbb{R}^n}) = \sup_{\substack{f \in C(\mathbb{R}^n \times \overline{\mathbb{R}^n}) \\ \|f\| \leq 1}} \int_{\mathbb{R}^n \times \overline{\mathbb{R}^n}} f \mathrm{d}\gamma' \geq \sup_{\substack{f \in C(\mathbb{R}^n) \\ \|f\| \leq 1}} \int_{\mathbb{R}^n} f \mathrm{d}\mu = \mu(\mathbb{R}^n) = 1.$$

Quindi $\gamma'(\{\infty\} \times \overline{\mathbb{R}^n}) = 0$. Sostituendo μ con ν si ottiene, in modo analogo, $\gamma'(\overline{\mathbb{R}^n} \times \{\infty\}) = 0$. Quindi il supporto di γ' è contenuto in $\mathbb{R}^n \times \mathbb{R}^n$, come volevamo. $\square$

Siamo finalmente pronti per dimostrare l'esistenza di una misura $\gamma \in \Gamma(\mu, \nu)$ con supporto ciclicamente monotono.

Teorema 2.1.10 (Esistenza della misura congiunta). *Siano $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$. Allora esiste una misura di probabilità di Borel γ su $\mathbb{R}^n \times \mathbb{R}^n$ che abbia supporto ciclicamente monotono e le sue marginali siano μ e ν.*

Dimostrazione. Per il Lemma 2.1.5 possiamo scegliere due successioni $(\mu_k)_{k\in\mathbb{N}}$ e $(\nu_k)_{k\in\mathbb{N}}$ della forma

$$\mu_k = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, \quad \nu_k = \frac{1}{M}\sum_{i=1}^{M}\delta_{y_i},$$

convergenti debolmente* a μ e ν rispettivamente. Per ogni k, possiamo assumere $N = M$: infatti, possiamo sostituire N ed M con MN e aumentare la molteplicità di ogni δ_{x_i} e δ_{y_i} di un fattore M ed N rispettivamente. Inoltre, per ogni k, possiamo riordinare gli x_i in modo che i punti (x_i, y_i), $i = 1, 2, \ldots, N$ formino un insieme ciclicamente monotono (si veda il Teorema 1.3.10). Possiamo ora definire la successione di misure $(\gamma_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n \times \mathbb{R}^n$ data da

$$\gamma_k = \frac{1}{N}\sum_{i=1}^{N}\delta_{(x_i,y_i)}.$$

Per come sono costruite, per ogni k, γ_k ha supporto ciclicamente monotono, ha come marginali μ_k e ν_k e sta nella palla unitaria di $C_0(\mathbb{R}^n \times \mathbb{R}^n)^*$, quindi converge, a meno di passare a una sottosuccessione, a una certa misura γ, che sarà una misura di probabilità di Borel positiva, per il Lemma 2.1.6 avrà supporto ciclicamente monotono e per il Lemma 2.1.9 le sue marginali saranno μ e ν. □

2.2 Esistenza della funzione cercata

Nella sezione precedente abbiamo mostrato che, date due misure μ e ν di probabilità di Borel su $\mathbb{R}^n$, allora esiste una misura γ di probabilità di Borel su $\mathbb{R}^n \times \mathbb{R}^n$ con supporto ciclicamente monotono. Per il Teorema di Rockafellar (si veda 1.3.12) esiste una funzione convessa, propria e chiusa ψ tale che il suo sottodifferenziale $\partial\psi$ contenga il supporto di γ. Inoltre, dal Teorema

1.3.14 sappiamo che $\partial\psi$ non è contenuto nel sottodifferenziale di alcun'altra funzione convessa, propria e chiusa.
In questa sezione, come anticipato a inizio capitolo, mostreremo che la misura indotta da μ tramite $\nabla\psi$ è proprio ν, ossia $\nabla\psi_{\#}\mu = \nu$, esattamente come volevamo.

Teorema 2.2.1. *Sia γ una misura di probabilità di Borel supportata nel sottogradiente di una funzione convessa, propria e chiusa $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Siano μ e ν le marginali di γ. Se μ si annulla sugli insiemi di Borel di dimensione di Hausdorff di dimensione $n-1$, allora $\gamma = (\mathrm{Id} \times \nabla\psi)_{\#}\mu$, in particolare $\nabla\psi_{\#}\mu = \nu$.*

Dimostrazione. Prima di entrare nel vivo della dimostrazione, bisogna fare qualche osservazione preliminare. In primo luogo, per il Teorema 1.3.8, $\partial\psi$ è un sottoinsieme chiuso di $\mathbb{R}^n \times \mathbb{R}^n$. ricordiamo poi che l'insieme $\mathrm{dom}(\nabla\psi) := \{x \in \mathbb{R}^n \mid \text{ esiste } \nabla\psi(x)\}$ è contenuto nell'interno di $\mathrm{dom}(\psi)$ (si veda Corollario 1.3.5). Inoltre, siccome $\operatorname{supp}\gamma \subseteq \partial\psi \subseteq \mathrm{dom}(\psi) \times \mathbb{R}^n$, allora il supporto di μ è contenuto nella chiusura $\mathrm{dom}(\psi)$.
In secondo luogo ci occorrerà che $\nabla\psi$ sia definito μ-quasi-dappertutto. Questo è vero, poiché ψ è differenziabile tranne al più in un insieme che ha dimensione di Hausdorff $n-1$ contenuto in $\mathrm{dom}(\psi)$ (si ricordi il Teorema 1.3.6). Siccome poi $\mathrm{dom}(\psi)$ è un insieme convesso (si veda l'Osservazione 1.1.9), la sua frontiera ha anch'essa dimensione di Hausdorff $n-1$ (si veda la Proposizione 1.1.15), quindi $\mu(\mathrm{dom}(\nabla\psi)) = \mu(\mathrm{dom}(\psi)) = 1$.
Inoltre, $\nabla\psi$ è una funzione Borel misurabile, poiché, per ogni $x, z \in \mathbb{R}^n$, $\langle\nabla\psi, z\rangle$ è il limite di una successione di funzioni continue, quindi Borel misurabili. Infatti si ha

$$\langle\nabla\psi(x), z\rangle = \lim_{k\to\infty} k\left(\psi\left(x + \frac{z}{k}\right) - \psi(x)\right).$$

Sostituendo a z i vettori della base canonica si ottiene la misurabilità ciascuna componente di $\nabla\psi$. Di conseguenza $\nabla\psi$ è Borel misurabile.
Ora, se $(\mathrm{Id} \times \nabla\psi)_{\#}\mu = \gamma$, per definizione di misura marginale, si ha immediatamente che $\nabla\psi_{\#}\mu = \nu$. Quindi, per completare la dimostrazione sarà

sufficiente mostrare che $\mathrm{Id} \times \nabla\psi_{\#}\mu$ e γ coincidono sugli insiemi del tipo $M \times N$, dove $M, N \subseteq \mathbb{R}^n$ sono due insiemi di Borel. A tal fine, definiamo $S := \{(x, \nabla\psi(x)) \mid x \in \mathrm{dom}(\nabla\psi)\}$. Si ha che $S = (\mathrm{dom}(\nabla\psi) \times \mathbb{R}^n) \cap \partial\psi$, poiché se $(x, \nabla\psi) \in S$, allora $x \in \mathrm{dom}(\nabla\psi)$, quindi in x la ψ è sottodifferenziabile e $\partial\psi(x) = \{\nabla\psi(x)\}$. Quindi S, essendo intersezione di insiemi di Borel, è un insieme di Borel. Inoltre S contiene $(\mathrm{dom}(\nabla\psi) \times \mathbb{R}^n) \cap \mathrm{supp}\,\gamma$ ed essendo $\mu(\mathrm{dom}(\nabla\psi)) = 1$, per ogni insieme G γ-misurabile si ha $\gamma(G) = \gamma(G \cap S)$.
Osservando che $(M \times N) \cap S$ è l'insieme delle coppie $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ tali che $x \in \mathrm{dom}(\nabla\psi) \cap M$ e $y = \nabla\psi(x)$, possiamo scrivere

$$(M \times N) \cap S = ((M \cap (\nabla\psi)^{-1}(N)) \times \mathbb{R}^n) \cap S. \tag{2.3}$$

Segue che

$$\begin{aligned}
&\gamma(M \times N) = \gamma((M \times N) \cap S) = \\
&= \gamma(((M \cap (\nabla\psi)^{-1}(N)) \times \mathbb{R}^n) \cap S) = \\
&= \gamma((M \cap (\nabla\psi)^{-1}(N)) \times \mathbb{R}^n) = \\
&= \mu(M \cap (\nabla\psi)^{-1}(N)) = (\mathrm{Id} \times \nabla\psi)_{\#}\mu(\mathrm{M} \times \mathrm{N})
\end{aligned}$$

Questo conclude la dimostrazione. □

Dimostrazione del Teorema 2.0.1. Il Teorema 2.1.10 garantisce l'esistenza di una misura congiunta γ, avente supporto ciclicamente monotono, che ha come marginale sinistra μ e marginale destra ν. Per il Teorema di Rockafellar (si veda 1.3.12) esiste una funzione ψ convessa, propria e chiusa che contiene il supporto di γ. Infine il Teorema 2.2.1 ci garantisce che $\psi_{\#}\mu = \nu$. □

Capitolo 3

Unicità della funzione "measure-preserving"

Nel capitolo precedente abbiamo dimostrato che, date due misure $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$, con μ tale che si annulli sugli insiemi aventi dimensione di Hausdorff $n-1$, allora esiste una funzione ψ convessa, propria e chiusa tale che $\nabla\psi_{\#}\mu = \nu$. L'obiettivo di questo capitolo è quello di dimostrare che $\nabla\psi$ è unico μ-quasi-dappertutto. Più precisamente:

Teorema 3.0.1. *Siano $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$, con μ tale che si annulli sugli insiemi di Borel aventi dimensione di Hausdorff $n-1$. Se $\psi, \varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ sono due funzioni convesse, proprie e chiuse tali che*

$$\nabla\psi_{\#}\mu = \nu, \quad \nabla\varphi_{\#}\mu = \nu,$$

allora $\nabla\psi = \nabla\varphi$ μ-quasi-dappertutto.

3.1 Il teorema della funzione implicita nel caso convesso

In questa sezione presentiamo una variante del teorema della funzione implicita che non richiede la differenziabilità della funzione di cui si vuole esplicitare una variabile in funzione delle altre, ma soltanto la convessità. Prima di vedere questo teorema è necessario un Lemma, che costituisce un analogo del teorema del valor medio per le funzioni convesse.
Nel seguito useremo la notazione f'_+ e f'_- per indicare la derivata destra e sinistra rispettivamente, qualora esistano, di una funzione f di una variabile reale a valori in $\mathbb{R}$.
Ricordiamo un paio di risultati sulle derivate parziali delle funzioni convesse che interverrano nella dimostrazione del Lemma.

Proposizione 3.1.1. *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa e sia $x \in \mathbb{R}^n$ un punto tale che $f(x) \in \mathbb{R}$. Allora, per ogni $y \in \mathbb{R}^n$, la funzione*

$$\lambda \mapsto \frac{f(x+\lambda y) - f(x)}{\lambda}, \quad \lambda > 0,$$

è crescente. In particolare, la derivata direzionale in x di f lungo la direzione y esiste e si ha:

$$\frac{\partial f}{\partial y}(x) = \inf_{\lambda>0} \frac{f(x+\lambda y) - f(x)}{\lambda}$$

Dimostrazione. Siano $\lambda, \mu \in \mathbb{R}$, $\mu > \lambda > 0$. Vogliamo mostrare che

$$\frac{f(x+\lambda y) - f(x)}{\lambda} \leq \frac{f(x+\mu y) - f(x)}{\mu}.$$

La disuguaglianza precedente è equivalente a

$$f(x+\lambda y) \leq \frac{\lambda}{\mu} f(x+\mu y) + \left(1 - \frac{\lambda}{\mu}\right) f(x).$$

Osservando che

$$x + \lambda y = \frac{\lambda}{\mu}(x+\mu y) + \left(1 - \frac{\lambda}{\mu}\right) x$$

e che $\frac{\lambda}{\mu} < 1$, dalla convessità di f segue che

$$f(x+\lambda y) = f\left(\frac{\lambda}{\mu}(x+\mu y) + \left(1-\frac{\lambda}{\mu}\right)x\right) \leq \frac{\lambda}{\mu} f(x+\mu y) + \left(1-\frac{\lambda}{\mu}\right) f(x),$$

che è precisamente la disuguaglianza che volevamo ottenere. □

Proposizione 3.1.2. *Sia $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ una funzione convessa propria e siano $x, y \in \mathbb{R}^n$. Allora*

$$\frac{\partial f}{\partial y}(x) = \sup\{\langle x^*, y\rangle \mid x^* \in \partial f(x)\}.$$

Dimostrazione. Si veda [2], Teorema 23.4, pag. 217. □

Lemma 3.1.3 (Teorema del valor medio)**.** *Siano $\psi, \varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ due funzioni convesse e $p, q \in \operatorname{int}(\operatorname{dom}(\psi)) \cap \operatorname{int}(\operatorname{dom}(\varphi))$ tali che $\psi(p) - \varphi(p) = \psi(q) - \varphi(q)$. Allora, per qualche $x = tp + (1-t)q$, $t \in (0,1)$, esistono $u \in \partial\psi(x)$ e $v \in \partial\varphi(x)$ tali che $\langle u - v, q - p\rangle = 0$.*

Dimostrazione. Per la convessità, ψ e φ assumono valori finiti in un intorno del segmento di estremi p, q. Poiché ψ e φ sono continue su tale intorno (si veda Proposizione 1.1.14), la funzione $f(t) := \psi(tp+(1-t)q) - \varphi(tp+(1-t)q)$ è continua per ogni $t \in [0,1]$. Siccome $[0,1]$ è compatto e $f(0) = f(1)$, deve esistere un $\bar{t} \in (0,1)$ punto di massimo o di minimo per f. Sia $x = \bar{t}p + (1-\bar{t})q$. Siccome x appartiene all'interno del dominio effettivo di ψ, $\partial\psi(x)$ è chiuso e limitato, quindi è compatto. Dunque esistono u_+ e u_- punti di massimo e minimo rispettivamente per la funzione continua $g(u) := \langle u, q-p\rangle$, $u \in \partial\psi(x)$. Analogamente esistono v_+ e v_- punti di massimo e di minimo per la funzione continua $h(v) := \langle v, q-p\rangle$, $v \in \partial\varphi(x)$. Definiamo $\Psi(t) = \psi(tp + (1-t)q)$ e $\Phi(t) = \varphi(tp+(1-t)q), t \in \mathbb{R}$, in modo che $f = \Psi - \Phi$. Quindi, per la Proposizione 3.1.2:

$$\begin{aligned}
\Psi'_+(\bar{t}) &= \langle u_+, q-p\rangle, \quad \Psi'_-(\bar{t}) = \langle u_-, q-p\rangle,\\
\Phi'_+(\bar{t}) &= \langle v_+, q-p\rangle, \quad \Phi'_-(\bar{t}) = \langle v_-, q-p\rangle,\\
f'_+(\bar{t}) &= \Psi'_+(\bar{t}) - \Phi'_+(\bar{t}) = \langle u_+ - v_+, q-p\rangle,\\
f'_-(\bar{t}) &= \Psi'_-(\bar{t}) - \Phi'_-(\bar{t}) = \langle u_- - v_-, q-p\rangle,
\end{aligned}$$

Se x è un massimo, deve essere

$$f'_+(\bar{t}) \leq 0 \leq f'_-(\bar{t})$$
$$\langle u_+ - v_+, q - p\rangle \leq 0 \leq \langle u_- - v_-, q - p\rangle,$$

se invece x è un minimo le disuguaglianze sono rovesciate.
Possiamo quindi scegliere $\lambda \in [0, 1]$ tale che:

$$(1-\lambda)\langle u_+ - v_+, q - p\rangle + \lambda\langle u_- - v_-, q - p\rangle = 0$$
$$\Big\langle [(1-\lambda)u_+ + \lambda u_-] - [(1-\lambda)v_+ + \lambda v_-]\ ,\ q - p\Big\rangle = 0.$$

Siccome $\partial\psi(x)$ e $\partial\varphi(x)$ sono convessi, $u := (1-\lambda)u_+ + \lambda u_- \in \partial\psi(x)$ e $v := (1-\lambda)v_+ + \lambda v_- \in \partial\varphi(x)$. Questo conclude la dimostrazione.

□

Ricordiamo questo risultato sull'estensione delle funzioni Lipschitziane, che ci servirà nella dimostrazione del teorema della funione implicita.

Proposizione 3.1.4. *Sia $A \subseteq \mathbb{R}^n$ ed $f : A \to \mathbb{R}$ una funzione Lipschitziana di costante $L > 0$. Allora esiste una funzione $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ tale che $\bar{f}_{|A} = f$ e $\bar{f}$ è Lipschitziana di costante L su tutto $\mathbb{R}^n$.*

Dimostrazione. Definiamo

$$\bar{f}(x) := \inf_{a\in A}\{f(a) + L|x - a|\}.$$

Chiaramente, $-\infty < \inf_{a\in A}\{f(a) + L|x - a|\} < +\infty$ per ogni $x \in \mathbb{R}^n$. Mostriamo che $\bar{f}$ ha le proprietà desiderate. Innanzitutto, vediamo che coincide con f se ristretta ad A. Infatti, se $b \in A$, allora, prendendo $a = b$ nella definizione si ottiene che $\bar{f}(b) \leq f(b)$. D'altra parte, poiché f è Lipschitziana su A, si ha che:

$$f(b) - f(a) \leq |f(b) - f(a)| \leq L|b - a|,$$

da cui segue che $f(b) \leq \inf_{a\in A}\{f(a) + L|b - a|\} = \bar{f}(b)$, quindi deve essere $\bar{f}(b) = f(b)$. Vediamo ora che $\bar{f}$ è Lipschitziana di costante L. Siano $x, y \in \mathbb{R}^n$. Allora

$$\bar{f}(x) = \inf_{a\in A}\{f(a)+L|x-a|\} \leq \inf_{a\in A}\{f(a)+L(|x-y|+|y-a|)\} = \bar{f}(y)+L|x-y|,$$

ossia $\bar{f}(x)-\bar{f}(y) \leq L|x-y|$. Scambiando x con y, si mostra che $\bar{f}(y)-\bar{f}(x) \leq L|y-x|$, da cui la Lipschitzianità. $\square$

Teorema 3.1.5. *Siano $\psi, \varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ due funzioni convesse differenziabili in $p \in \mathbb{R}^n$ tali che $\psi(p) = \varphi(p)$ e $\nabla\psi(p) \neq \nabla\varphi(p)$. Supponiamo che $\nabla\psi(p) - \nabla\varphi(p)$ sia parallelo al primo vettore della base canonica e_1 e orientato nel verso positivo (possiamo assumere che sia sempre così a meno di ruotare il sistema di riferimento). Allora esiste una funzione $f : \mathbb{R}^{n-1} \to \mathbb{R}$ Lipschitziana con costante 1 e un intorno U di p dentro il quale valga $\psi(x) = \varphi(x)$ se e solo se $x_1 = f(x_2, \ldots, x_n)$.*

Dimostrazione. Senza perdita di generalità, possiamo supporre $p = 0$. Poiché ψ è una funzione convessa, per ogni $((x_k, y_k))_{k\in\mathbb{N}}$ successione in $\partial\psi$ tale che $x_n \xrightarrow[k\to+\infty]{} 0$, si ha che $y_k \xrightarrow[k\to+\infty]{} \nabla\psi(0)$ (si veda Teorema 1.3.8). Le stesse considerazioni valgono per φ. Siccome $\nabla\psi(0) - \nabla\varphi(0) = \lambda e_1$, con $\lambda > 0$, esisterà un intorno U dell'origine tale che, indicando con $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ la funzione $(u_1, u_2, \ldots, u_n) \mapsto (u_2, \ldots, u_n)$, per ogni $x \in U$ e $(u_1, u_2, \ldots, u_n) \in \partial\psi(x), (v_1, v_2, \ldots, v_n) \in \partial\varphi(x)$ si abbia

$$u_1 - v_1 > |\pi(u) - \pi(v)|. \tag{3.1}$$

Non è restrittivo supporre che U sia convesso. Infatti, se non lo fosse, essendo U un intorno dell'origine, esso contiene una palla aperta centrata nell'origine, che è convessa, quindi ci si può sempre restringere a essa. Inoltre, poiché ψ e φ sono differenziabili nell'origine, questa è un punto interno al loro dominio effettivo, di conseguenza sono continue in un intorno V di 0, che, a meno di restringerci a $V \cap U$, possiamo supporre coincidere con U.
Osserviamo che $u_1 - v_1 > 0$. Questo ci dice che la funzione $h = \psi - \varphi$ deve essere strettamente monotona in U rispetto alla prima variabile.
Infatti, siccome h è continua, se non fosse strettamente monotona nella prima variabile, allora esiterebbe $x \in U$ per cui la funzione $h_x(\lambda) := h(x + \lambda e_1)$, $\lambda \in \mathbb{R}$, non è strettamente monotona. Quindi, essendo h_x continua, non può essere iniettiva. Possiamo supporre allora che esista $\alpha > 0$ tale che $h(x) = h(x + \alpha e_1)$, con $x + \alpha e_1 \in U$. A questo punto, per il Lemma 3.1.3,

esisterebbe β, $0 < \beta < \alpha$, tale che per opportuni $u \in \partial\psi(x + \beta e_1)$ e $v \in \partial\varphi(x + \beta e_1)$ si abbia $\langle u - v, (\alpha - \beta)e_1 \rangle = 0$. Ma ciò è assurdo, poiché $\langle u - v, (\alpha - \beta)e_1 \rangle = (u_1 - v_1)(\alpha - \beta) > 0$ (osservare che, siccome U era stato scelto in modo che fosse convesso, $x + \beta e_1 \in U$ e quindi vale la (3.1)).
Prendiamo $p^+, p^- \in U$ tali che $p^+ = (p_1^+, 0, \dots, 0)$, $p^- = (p_1^-, 0, \dots, 0)$, con $p_1^- < 0 < p_1^+$ per i quali si abbia $h(p_-) < 0 < h(p^+)$ (ricordiamo che $h(0) = 0$). Per la continuità di h, esistono due palle aperte B_+, B_- di centro p^+, p^- rispettivamente ed entrambe di raggio r tali che $h(x) > 0$ per ogni $x \in B_+$ e $h(x) < 0$ per ogni $x \in B_-$. Possiamo scegliere r in modo che $B_+, B_- \subseteq U$.
Sia $w \in B(0, r) \subseteq \mathbb{R}^{n-1}$. La funzione $h_w(x_1) = h(x_1, w)$, $x_1 \in \mathbb{R}$ assume sia valori negativi che positivi in U. Infatti, siccome $|w| < r$, $(p^+, w) \in B_+$ e $(p^-, w) \in B_-$. Quindi, essendo h_w continua e strettamente monotona, esiste un unico x_1 tale che $h_w(x_1) = 0$. Poniamo $f(w) = x_1$.
In $\mathbb{R} \times B(0, r) \subseteq \mathbb{R}^n$, la f così definita soddisfa chiaramente la condizione $f(x_2, \dots, x_n) = x_1$ se e solo se $\psi(x_1, x_2, \dots, x_n) = \varphi(x_1, x_2, \dots, x_n)$. Resta da mostrare la lipschitzianità di f su $B(0, r) \subseteq \mathbb{R}^{n-1}$. Una volta fatto ciò, si può estendere f su $\mathbb{R}^{n-1}$ per la Proposizione 3.1.4, senza modificarne la costante di Lipschitz.
Siano quindi $w, z \in B(0, r) \subseteq \mathbb{R}^{n-1}$. Poiché h si annulla sia su $(f(w), w)$ sia su $(f(z), z)$, per il Lemma 3.1.3 esiste un certo x sul segmento di estremi $(f(w), w)$ e $(f(z), z)$ tale che, per opportuni $u \in \partial\psi(x)$ e $v \in \partial\varphi(x)$, si abbia

$$\left\langle \begin{pmatrix} f(w) - f(z) \\ w - z \end{pmatrix}, u - v \right\rangle = 0,$$

di conseguenza

$$|f(w) - f(z)||u_1 - v_1| = |\langle w - z, \pi(u) - \pi(v) \rangle|,$$

da cui segue

$$|f(w) - f(z)||u_1 - v_1| \leq |\pi(u) - \pi(v)||w - z|,$$

infine, ricordando che $|u_1 - v_1| > |\pi(u) - \pi(v)|$, otteniamo la tesi. □

Ricordiamo ora una disuguaglianza sulla misura di Hausdorff, che insieme al risultato precedente ci fornirà un Corollario che interverrà nella dimostrazione del Teorema 3.0.1.

Proposizione 3.1.6. *Sia $f : \mathbb{R}^n \to \mathbb{R}^m$ una funzione lipschitziana di costante L, $A \subseteq \mathbb{R}^n$ e $0 \leq s < +\infty$. Allora*

$$\mathcal{H}^s(f(A)) \leq L^s \mathcal{H}^s(A) \tag{3.2}$$

Dimostrazione. Si veda [8], Teorema 2.8, pag 97. □

Corollario 3.1.7. *Siano $\psi, \varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ due funzioni convesse differenziabili in $p \in \mathbb{R}^n$ tali che $\psi(p) = \varphi(p)$ e $\nabla\psi(p) \neq \nabla\varphi(p)$. Allora esiste un intorno U di p tale che l'insieme $X = \{x \in U \mid \psi(x) = \varphi(x)\}$ ha misura di Hausdorff $(n-1)$-dimensionale finita.*

Dimostrazione. Per il Teorema 3.1.5 esistono un intorno U di p, che possiamo supporre limitato, e una funzione f come nell'enunciato di tale Teorema. Definiamo

$$\begin{aligned} g :&\mathbb{R}^{n-1} \to \mathbb{R}^n \\ &w \mapsto (f(w), w). \end{aligned}$$

Osserviamo che g è lipschitziana, infatti

$$|g(w) - g(z)| = |(f(w), w) - (f(z), z)| \leq |f(w) - f(z)| + |w - z| \leq 2|w - z|,$$

per ogni $w, z \in \mathbb{R}^{n-1}$.
Sia $A = g^{-1}(U)$. A è limitato. Infatti, per ogni $w, z \in A$ si ha

$$|w - z| \leq |(f(w), w) - (f(z), z)| = |g(w) - g(z)|,$$

e, per definizione di A, $g(w), g(z) \in U$. Poiché U è limitato, $|g(w) - g(z)|$ è limitato superiormente dal diametro di U, da cui segue la limitatezza di A. Per la Proposizione 3.1.6 si ha quindi che

$$\mathcal{H}^{n-1}(U) = \mathcal{H}^{n-1}(g(A)) \leq 2^{n-1}\mathcal{H}^{n-1}(A).$$

Siccome A è un sottoinsieme limitato di $\mathbb{R}^{n-1}$, la sua misura di Hausdorff $(n-1)$-dimensionale è finita. Inoltre, siccome X è chiaramente contenuto in U, anche X ha misura di Hausdorff $(n-1)$-dimensionale finita.

$\square$

3.2 Dimostrazione dell'unicità

Ricordiamo qui un risultato che ci servirà nella dimostrazione dell'unicità.

Definizione 3.2.1 (Insieme universalmente misurabile)**.** *Un sottoinsieme $A \subseteq \mathbb{R}^n$ si dice universalmente misurabile se è μ-misurabile per ogni μ misura di probabilità di Borel su $\mathbb{R}^n$.*

Proposizione 3.2.2. *Siano $1 \leq k \leq n$ e $C \subseteq \mathbb{R}^n$ un insieme di Borel. Allora, indicando con $\pi : \mathbb{R}^n \to \mathbb{R}^k$ la proiezione $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_k)$, $\pi(C)$ è universalmente misurabile in $\mathbb{R}^k$.*

Dimostrazione. Si veda [6], Proposizione 8.4.4, pag. 264. $\square$

Prima di passare alla dimostrazione dell'unicità, ci serve un ultimo lemma, derivato da un argomento usato da Alexandrov per dimostrare l'esistenza di una superficie convessa con una certa curvatura gaussiana assegnata. Nel seguito, data φ una funzione convessa su $\mathbb{R}^n$ e $M \subseteq \mathbb{R}^n$, indicheremo con $\partial\varphi(M) := \bigcup_{m \in M} \partial\varphi(m)$.

Lemma 3.2.3. *Siano $\psi, \varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ due funzioni convesse e differenziabili in $p \in \mathbb{R}^n$ per cui valga $\psi(p) = \varphi(p) = 0$ e $\nabla\varphi(p) \neq \nabla\psi(p) = 0$. Siano inoltre $M := \{x \in \mathrm{dom}(\psi) \mid \varphi(x) > \psi(x)\}$ e $X := \nabla\psi^{-1}(\partial\varphi(M))$. Allora $X \subseteq M$ e la distanza di p da X è strettamente positiva.*

Dimostrazione. Sia $x \in X$. Allora, per come è definito X, esiste un $m \in M$ tale che $\nabla\psi(x) \in \partial\varphi(m)$. Dalla definizione di sottogradiente, per ogni $z \in \mathbb{R}^n$ si ha:

$$\varphi(z) \geq \langle \nabla\psi(x), z - m \rangle + \varphi(m)$$
$$\psi(m) \geq \langle \nabla\psi(x), m - x \rangle + \psi(x).$$

Per definizione di M, sappiamo che $\psi(m) < \varphi(m)$. Combinando questa con le due disuguaglianze precedenti otteniamo

$$\varphi(z) \geq \langle \nabla\psi(x), z - m \rangle + \varphi(m) > \langle \nabla\psi(x), z - m \rangle + \psi(m) \geq$$
$$\geq \langle \nabla\psi(x), z - m \rangle + \langle \nabla\psi, m - x \rangle + \psi(x) = \langle \nabla\psi(x), z - x \rangle + \psi(x),$$

ossia, riassumendo

$$\varphi(z) > \langle \nabla\psi(x), z - x \rangle + \psi(x) \quad \text{per ogni } z \in \mathbb{R}^n. \tag{3.3}$$

Prendendo $z = x$ otteniamo $\varphi(x) > \psi(x)$, ossia $X \subseteq M$.
Vediamo la seconda parte. Supponiamo per assurdo che la distanza di p da X sia nulla, ossia $p \in \overline{X}$. Allora esiste una successione $(x_k)_{k\in\mathbb{N}}$ in X tale che $x_k \xrightarrow[k\to+\infty]{} p$. Per ogni $k \in \mathbb{N}$, esistono $\nabla\psi(x_k)$ un sottogradiente di ψ in x_k e $m_k \in M$ tale che $\nabla\psi(x_k) \in \partial\varphi(m_k)$. Siccome $\nabla\psi(p) = 0$ e $\psi(p) = 0$, si ha che $\psi \geq 0$ e per la continuità del sottogradiente di una funzione convessa $\nabla\psi(x_k) \xrightarrow[k\to+\infty]{} 0$. D'altra parte $\nabla\varphi(p) \neq 0$ e $\varphi(p) = 0$, quindi deve esistere $z \in \mathbb{R}^n$ tale che $\varphi(z) < 0$. Per tale valore di z, ricordando la (3.3) si ha

$$0 > \varphi(z) > \langle \nabla\psi(x_k), z - x_k \rangle + \psi(x_k) \geq -|\nabla\psi(x_k)||z - x_k|. \tag{3.4}$$

Ma poiché $|\nabla\psi(x_k)|$ tende a zero e $|z - x_k|$ è limitata, per k abbastanza grande otteniamo $-|\nabla\psi(x_k)||z - x_k| > \varphi(z)$. Siamo quindi giunti a un assurdo, perciò la distanza di p da $\overline{X}$ deve essere strettamente positiva. $\square$

A questo punto, abbiamo tutto quello che ci serve per dimostrare il Teorema 3.0.1.

Dimostrazione del Teorema 3.0.1. Siano ψ e φ due funzioni convesse, proprie e chiuse su $\mathbb{R}^n$ tali che $\nabla\psi_{\#}\mu = \nu$ e $\nabla\varphi_{\#}\mu = \nu$, μ e ν come nell'enunciato del Teorema. Supponiamo, per assurdo, che non valga $\nabla\psi = \nabla\varphi$ μ-quasi dappertutto, ossia, supponiamo che esista $p \in \operatorname{supp}\mu$ tale che ψ e φ siano differenziabili in p e $\nabla\psi(p) \neq \nabla\varphi(p)$. Ricordiamo che, essendo ψ e φ chiuse, sono inferiormente semicontinue e $\partial\varphi$ è un insieme chiuso. A meno di sottrarre una costante a ψ e φ, modifica che lascia inalterati i sottogradienti

delle due funzioni, possiamo supporre che $\psi(p) = \varphi(p) = 0$.
Siccome $p \in \operatorname{supp}\mu$, per ogni intorno V di p si ha $\mu(V) > 0$. Per il Corollario 3.1.7 esiste un intorno di p dove ψ e φ coincidono al più su un insieme che ha misura di Hausdorff $(n-1)$-dimensionale finita. Tale insieme ha misura nulla per μ.
Ogni intorno V di p si può scomporre in tre insiemi:

- $V_{\psi<\varphi} := \{x \in V \mid \psi(x) < \varphi(x)\}$;
- $V_{\psi=\varphi} := \{x \in V \mid \psi(x) = \varphi(x)\}$;
- $V_{\varphi<\psi} := \{x \in V \mid \varphi(x) < \psi(x)\}$.

Abbiamo visto che $V_{\psi=\varphi}$ ha misura nulla. Essendo $\mu(V) > 0$, almeno uno tra $V_{\psi<\varphi}$ e $V_{\varphi<\psi}$ deve avere misura strettamente positiva. Possiamo quindi supporre che, a meno di scambiare ψ e φ, sia $\mu(V_{\psi<\varphi}) > 0$.
Sottraiamo a ψ una trasformazione lineare in modo che si abbia $\nabla\psi(p) = 0$. Tale modifica corrisponde a una traslazione di ν.
Definiamo

$$M := \{x \in \operatorname{int}(\operatorname{dom}(\psi)) \mid \psi(x) < \varphi(x)\} = [\varphi - \psi]^{-1}((0, +\infty)) \cap \operatorname{int}(\operatorname{dom}(\psi)).$$

Osserviamo che φ è inferiormente semicontinua, quindi anche φ ristretta a $\operatorname{int}(\operatorname{dom}(\psi))$ lo è. Inoltre ψ ristretta a $\operatorname{int}(\operatorname{dom}(\psi))$ è continua, quindi anche $-\psi$ è continua e in particolare inferiormente semicontinua. Siccome la somma di funzioni inferiormente semicontinue è ancora inferiormente semicontinua, $\varphi - \psi$ ristretta a $\operatorname{int}(\operatorname{dom}(\psi))$ è inferiormente semicontinua. Di conseguenza M è aperto, dunque un boreliano. Siccome, come avevamo già detto, $\partial\varphi$ è chiuso, quindi anche lui è un boreliano, segue che $(M \times \mathbb{R}^n) \cap \partial\varphi$ è un boreliano, dunque, per la Proposizione 3.2.2, la sua proiezione sul secondo $\mathbb{R}^n$, ossia $Y := \partial\varphi(M)$, è universalmente misurabile, in particolare μ-misurabile.
Mostriamo che

$$\mu(\nabla\psi^{-1}(Y)) < \mu(\nabla\varphi^{-1}(Y))$$

e che quindi $\psi_{\#}\mu$ e $\varphi_{\#}\mu$ non possono coincidere entrambe con ν. La disuguaglianza precedente ha senso in quanto $\nabla\psi$ e $\nabla\varphi$ sono due funzioni

μ-misurabili (si mostra in modo analogo a quanto fatto nella dimostrazione del Teorema 2.2.1).
In primo luogo osserviamo che $1 = \nu(\mathbb{R}^n) = \mu(\nabla\varphi^{-1}(\mathbb{R}^n)) = \mu(\mathrm{dom}(\nabla\varphi))$. Dunque $\mu(A) = \mu(A \cap \mathrm{dom}(\nabla\varphi))$, per ogni boreliano A. Inoltre, $M \cap \mathrm{dom}(\nabla\varphi)$ è il sottoinsieme dei punti di M in cui φ è differenziabile, quindi è contenuto in $\nabla\varphi^{-1}(Y) = \nabla\varphi^{-1}(\partial\varphi(M))$ (questo poiché se una funzione è differenziabile allora è anche sottodifferenziabile). Dunque, mettendo insieme queste due osservazioni, abbiamo che:

$$\mu(M) = \mu(M \cap \mathrm{dom}(\nabla\varphi)) \leq \mu(\nabla\varphi^{-1}(\partial\varphi(M))) = \mu(\nabla\varphi^{-1}(Y)).$$

D'altra parte, per il Lemma 3.2.3, $\nabla\psi^{-1}(\partial\varphi(M)) \subseteq M$. Sempre per il Lemma 3.2.3, p ha distanza strettamente positiva da $\nabla\psi^{-1}(\partial\varphi(M))$, quindi avrà un intorno disgiunto da $\nabla\psi^{-1}(\partial\varphi(M))$ che, per quanto osservato nella prima parte della dimostrazione, interseca M in un insieme di misura strettamente positiva. Quindi deve essere

$$\mu(\nabla\psi^{-1}(Y)) = \mu(\nabla\psi^{-1}(\partial\varphi(M))) < \mu(M) \leq \mu(\nabla\varphi^{-1}(Y)).$$

Questo conclude la dimostrazione.

$\square$

Bibliografia

[1] Robert McCann. *Existence and uniqueness of monotone measure-preserving maps*. Duke Mathematical Journal, 1995.

[2] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[3] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 1987 (International edition).

[4] W. Rudin. *Functional Analysis*. McGraw-Hill, 1991 (seconda edizione).

[5] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2011.

[6] Donald L. Cohn, *Measure Theory*, Birkhäuser, 2013 (seconda edizione).

[7] M. Manetti, *Topologia*, Springer, 2014 (seconda edizione).

[8] Lawrence C. Evans, Ronald F. Gariepy, *Measure Theory and fine properties of function*, CRC Press, 2015 (revised edition).

[9] R. D. Anderson e V. L. Klee, Jr., *Convex functions and upper semi-continuous collections*, Duke Mathematical Journal, 1952.

[10] Lawrence C. Evans, *Partial Differential Equations and Monge–Kantorovich Mass Transfer*, (disponibile online) 2001.