

Indice

Introduzione	3
1 Elementi di Calcolo delle Variazioni	7
1.1 L'equazione di Eulero	7
1.2 Altre condizioni necessarie per i minimi	10
1.3 Calibrazioni	12
1.4 Campi di estremali	13
2 Un teorema di regolarità parziale di Tonelli	23
2.1 Alcuni lemmi preliminari	24
2.2 Il teorema di Tonelli	28
3 Esistenza dei minimi: un teorema di semicontinuità	33
3.1 Alcuni lemmi preliminari	34
3.2 La dimostrazione del teorema di semicontinuità	37
4 Spazi di Carnot-Carathéodory	43
4.1 La distanza di Carnot-Carathéodory	43
4.1.1 Una definizione equivalente della distanza d_{CC}	46
4.2 Il teorema di Chow	51
5 Geodetiche negli spazi di Carnot-Carathéodory	57
5.1 Esistenza	57
5.2 Regolarità	59
5.2.1 Il problema del controllo ottimo e il Principio di Pontryagin	59
5.2.2 Geodetiche normali e singolari	61
5.3 Un esempio concreto: il gruppo di Heisenberg	64
6 Gruppi di Carnot	69
6.1 Gruppi di Lie e algebre relative	69
6.2 Gruppi di Carnot	71

6.3	Coordinate graduate	72
6.3.1	Il gruppo di Heisenberg	74
6.3.2	Campi invarianti a sinistra	75
6.4	I gruppi di Carnot come spazi di Carnot-Carathéodory	79
6.5	Il teorema di Pansu	81
7	Regolarità delle geodetiche nei Gruppi di Carnot	85
7.1	Regolarità per curve orizzontali	85
7.2	Una proprietà dei gruppi di Carnot	88
7.3	Regolarità delle geodetiche: un'estensione del teorema di Tonelli ai Gruppi di Carnot	91
7.4	Regolarità globale delle geodetiche	92
A	Funzioni assolutamente continue	95
	Bibliografia	99

Introduzione

Storicamente le prime questioni su strutture di tipo sub-riemanniano appaiono in un lavoro del 1909, [5], di C. Carathéodory. Qui un processo termodinamico viene rappresentato da una opportuna curva in \mathbb{R}^n , e il calore scambiato nel processo dall'integrale lungo la curva stessa di una opportuna 1-forma θ . I lavori del fisico Jules Carnot dimostrarono l'esistenza di due stati che non potevano essere collegati da processi adiabatici, quelle curve cioè lungo cui θ si annulla: utilizzando un linguaggio più moderno si usa chiamare *orizzontali* queste curve. I lavori di Carnot e di Carathéodory dimostravano che se esistono due punti non connettibili orizzontalmente, allora θ è integrabile: in particolare esistono due funzioni T ed S (che vengono interpretate come temperatura ed entropia rispettivamente) tali che $\theta = T dS$.

Il passo successivo fu naturalmente quello di generalizzare questa struttura: l'idea fu quella di considerare non più una 1-forma e prendere come orizzontale la distribuzione di sottospazi di codimensione 1 su cui essa si annulla, bensì di considerare direttamente una distribuzione di sottospazi di codimensione qualsiasi. La condizione chiave per ottenere la connessione fra due punti qualsiasi diventa allora il fatto che l'algebra di Lie della distribuzione genera in ogni punto l'intero spazio tangente: tale condizione è ben nota in varie branche della Matematica e assume di volta in volta i nomi di “totale nonolonomicità”, “condizione di Hörmander”, “bracket-generating condition” e “condizione di Chow”. Questo risultato fu dimostrato in maniera indipendente nel 1938-39 da Chow [7] e da Rashevsky.

Successivamente la teoria si è sviluppata in varie direzioni, affrontando questioni che sono tipiche delle più diverse branche della Matematica: dall'Analisi Armonica alla Teoria della Regolarità per PDE, fino alla Teoria Geometrica della Misura. L'Analisi e la Geometria degli Spazi Metrici trovano negli spazi di Carnot-Carathéodory ottimi campi di ricerca e innumerevoli spunti.

In particolare è emerso il problema della geodetiche, ovvero delle curve orizzontali più brevi che congiungono due punti fissati. Il primo metodo generale per trovare le geodetiche negli spazi di Carnot-Carathéodory si deve a Hermann nel suo lavoro [13] del 1973. Per quanto riguarda la regolarità si sfruttano solitamente tecniche di controllo ottimale: negli spazi di Carnot-Carathéodory si hanno fenomeni che non si

verificano nella Geometria Riemanniana, come la non unicità (neppure locale!) delle geodetiche e il fatto che non tutte le curve di minima lunghezza verificano l'equazione delle geodetiche. Per capire quanto sono sottili queste questioni, si consideri che il primo esempio di quest'ultimo fatto risale al 1994 (Montgomery [17]).

Il problema della regolarità delle geodetiche sub-riemanniane è a tutt'oggi ancora aperto perfino nella classe più ristretta dei gruppi di Carnot, che sono particolari gruppi di Lie in cui vale una forma molto forte della bracket-generating condition: essa fa sì che siano valide proprietà notevoli che legano l'operazione di gruppo ed alcuni suoi automorfismi alla distanza di Carnot-Carathéodory e che creano fenomeni tipici della geometria frattale.

La tesi affronta il problema della regolarità delle geodetiche nei gruppi di Carnot attraverso alcuni classici strumenti del Calcolo delle Variazioni: essenzialmente si tratterà di considerare Lagrangiane che possono assumere anche il valore $+\infty$ in corrispondenza di quelle direzioni che non sono orizzontali.

Nei primi tre capitoli affrontiamo alcuni fra gli argomenti più classici del Calcolo delle Variazioni. Il primo è dedicato allo studio delle proprietà dei minimi di funzionali del tipo

$$\mathcal{L}(x) = \int_I L(t, x(t), \dot{x}(t)) dt$$

al variare di x all'interno di una classe \mathcal{C} di funzioni reali definite sull'intervallo I . Tale classe verrà indicata di volta in volta: tipicamente saranno le funzioni di una certa regolarità con valori agli estremi fissati. Inizieremo ad elencare le proprietà di tali minimi a partire dalle più semplici, come l'equazione di Eulero, per poi passare ad altre più sofisticate attraverso la teoria dei campi di estremali: dimostreremo in particolare che sotto certe ipotesi sulla Lagrangiana L (essenzialmente l'ellitticità della sua matrice Hessiana) l'equazione di Eulero è condizione sufficiente a garantire la minimalità quando l'intervallo I è sufficientemente breve.

Questi risultati verranno largamente utilizzati nel capitolo successivo, dedicato alla presentazione di un risultato di regolarità parziale dovuto a Tonelli: esso garantisce, ancora sotto l'ipotesi di ellitticità della matrice Hessiana della Lagrangiana, che un eventuale minimo di \mathcal{L} deve essere di classe \mathcal{C}^∞ su un sottoinsieme aperto e di misura piena dell'intervallo I . Inoltre, nei punti che non appartengono a questo aperto, esiste il limite dei rapporti incrementali, e questo limite vale $\pm\infty$.

Il terzo capitolo è invece dedicato allo studio del problema dell'esistenza dei minimi: il risultato che presentiamo garantisce la semicontinuità inferiore del funzionale \mathcal{L} rispetto alla convergenza debole nello spazio delle funzioni assolutamente continue su I , quando la Lagrangiana L è semicontinua inferiormente nella prima variabile e convessa nella terza (fissate le prime due). Si noti che questa ipotesi di convessità

ben si sposa con la condizione di ellitticità della matrice Hessiana della Lagrangiana di cui abbiamo già accennato.

Col capitolo 4 iniziamo lo studio del nostro ambiente di lavoro, gli spazi di Carnot-Carathéodory. Lo studio di questi spazi è motivato, intuitivamente, dall'idea di volersi spostare solo lungo certe direzioni che assegnamo punto per punto: essenzialmente i nostri dati saranno una varietà differenziale (anche se per comodità ci ridurremo al caso di aperti di \mathbb{R}^n) e una famiglia di campi di vettori X_1, \dots, X_m localmente Lipschitziani. Tali campi individuano punto per punto un sottospazio lineare dello spazio tangente alla varietà che costituisce l'insieme delle direzioni consentite, che chiameremo orizzontali; chiameremo orizzontali anche le curve che costituiranno l'oggetto dei nostri studi, ovvero quelle che in quasi ogni punto hanno velocità orizzontale. Ammettendo a competere queste sole curve, saremo in grado di definire la distanza di Carnot-Carathéodory come l'estremo inferiore delle lunghezze delle curve assolutamente continue e orizzontali che hanno come estremi tali punti. Il primo problema che si pone è quello di capire se tale distanza è sempre finita, ovvero se due punti qualsiasi sono connessi da una curva orizzontale: un risultato ottenuto da Chow nel 1939 dà una risposta affermativa alla questione nel caso in cui i commutatori (eventualmente iterati) dei campi X_1, \dots, X_m generino punto per punto l'intero spazio tangente (quando questa ipotesi è verificata si parla anche di spazi sub-riemanniani). L'idea che sta alla base della dimostrazione è quella che si può approssimare, tramite la formula di Baker-Campbell-Hausdorff, la direzione del campo $[X, Y]$ seguendo prima il campo X , poi quello Y , poi $-X$ ed infine il campo $-Y$: in questo modo tutte le direzioni (anche quelle proibite) possono essere approssimate combinando opportunamente i campi orizzontali.

Nel capitolo successivo ci occupiamo dello studio delle geodetiche fra due punti di uno spazio di Carnot-Carathéodory, ovvero delle curve orizzontali di lunghezza minima. Cominceremo col dimostrarne l'esistenza sotto ipotesi piuttosto generali (essenzialmente la non degenerazione all'infinito dei campi X_1, \dots, X_m): qui useremo il risultato di semicontinuità del capitolo 3. Successivamente passeremo allo studio della regolarità delle geodetiche attraverso strumenti classici come il problema del controllo ottimo. In particolare il cosiddetto Principio di Pontryagin ci consentirà di dividere le geodetiche in due classi, normali e singolari: le prime risulteranno avere la stessa regolarità dei campi X_1, \dots, X_m . Infine studieremo le geodetiche nel caso particolare del gruppo di Heisenberg, che ci fornisce un esempio concreto di come le geodetiche negli spazi sub-riemanniani non siano uniche neppure localmente, a differenza di quanto succede nella geometria riemanniana (in cui peraltro le geodetiche, essendo tutte normali, sono automaticamente regolari).

Nel capitolo 6 forniamo una breve trattazione della teoria dei gruppi di Carnot; dopo una breve introduzione sui gruppi di Lie, enunceremo la definizione ed esporremo alcune proprietà dei primi, con particolare risalto alla struttura delle di-

latazioni. In questo contesto rileggeremo il gruppo di Heisenberg come un particolare gruppo di Carnot, per poi passare allo studio dei campi vettoriali invarianti a sinistra. Infine, considerando dei particolari campi X_1, \dots, X_m invarianti a sinistra, interpreteremo i gruppi di Carnot come varietà sub-riemanniane, studiando in particolare il legame fra la distanza di Carnot-Carathéodory e le dilatazioni. Infine enunceremo (e dimostreremo nel caso di mappe definite su \mathbb{R}) il teorema di Pansu, che altro non è che la generalizzazione del classico risultato di Rademacher al contesto dei gruppi di Carnot.

L'ultimo capitolo contiene alcuni risultati parziali sul problema della regolarità delle geodetiche in questi spazi; dopo aver mostrato una proprietà di alcune di esse, ispirati dal risultato di regolarità parziale di Tonelli discusso nel secondo capitolo andremo ad enunciare una proprietà (essenzialmente chiediamo l'unicità e la regolarità delle geodetiche che congiungono l'identità del gruppo ed alcuni altri suoi punti) che garantisce la regolarità delle geodetiche all'infuori di un chiuso di misura nulla. Una seconda proprietà di alcune geodetiche ci consentirà poi di ottenere la regolarità globale. Entrambe queste proprietà possono essere verificate direttamente nel gruppo di Heisenberg, grazie al fatto che in questo gruppo esiste una classificazione completa delle geodetiche; non è nota al momento la loro validità per gruppi di Carnot generali.

Infine nell'appendice esponiamo, brevemente e senza dare alcuna dimostrazione, la teoria delle funzioni assolutamente continue.

Capitolo 1

Elementi di Calcolo delle Variazioni

Il Calcolo delle Variazioni è una delle branche più importanti e vitali della Matematica. Lungi dal volerne dare una trattazione esauriente, esponiamo brevemente alcune idee che ci serviranno nel seguito, seguendo l'impostazione di [4], al quale rimandiamo per una trattazione più generale delle questioni.

Iniziamo col derivare nella prima sezione l'equazione di Eulero $L_x(t, x(t), \dot{x}(t)) = \frac{d}{dt} L_v(t, x(t), \dot{x}(t))$ per minimi x di classe \mathcal{C}^2 di funzionali $\mathcal{L}(x) = \int_I L(t, x(t), \dot{x}(t)) dt$. Nelle due sezioni successive dimostriamo l'equazione di Eulero in forma integrale per minimi più generali, per poi passare alla definizione ed alle prime proprietà delle calibrazioni. Nella quarta sezione infine, dopo aver introdotto la teoria dei campi di estremali (con particolare riferimento a quelli ottimali, di Weierstrass e di Mayer), la sfrutteremo per dimostrare che, in presenza di una Lagrangiana $L(t, x, v)$ di classe \mathcal{C}^3 e tale che $L_{vv} > 0$, allora ogni estremale (ovvero ogni soluzione dell'equazione di Eulero) è, in un intervallo sufficientemente piccolo, anche un minimo locale. Sfrutteremo ampiamente questo risultato nel capitolo successivo, quando dimostreremo un teorema di regolarità per questi minimi dovuto a Tonelli.

1.1 L'equazione di Eulero

Consideriamo l'integrale variazionale

$$\mathcal{L}(x) = \int_I L(t, x(t), \dot{x}(t)) dt$$

dove $I =]a, b[$ è un intervallo limitato di \mathbb{R} . Per semplicità la funzione $L(t, x, v) : \bar{I} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, detta **Lagrangiana**, verrà supposta di classe \mathcal{C}^1 . Ne segue che \mathcal{L} è ben definita per ogni funzione $x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$.

Osservazione 1.1. Per i nostri scopi in realtà sarà quasi sempre sufficiente che \mathcal{L} sia definita in un qualche intorno del grafico $\{(t, x(t), \dot{x}(t)) : t \in \bar{I}\} \subset \bar{I} \times \mathbb{R}^n \times \mathbb{R}^n$ di una certa funzione $x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ che vogliamo studiare.

In ogni caso, date una $x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ su cui \mathcal{L} è definita e una $\varphi \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$, siamo in grado di definire la funzione

$$\Phi :]-\epsilon_0, \epsilon_0[\ni \epsilon \longmapsto \Phi(\epsilon) := \mathcal{L}(x + \epsilon\varphi) \in \mathbb{R}$$

per un certo ϵ_0 sufficientemente piccolo. D'altra parte $\Phi \in \mathcal{C}^1(-\epsilon_0, \epsilon_0)$ ed un semplice calcolo dà

$$\Phi'(0) = \int_I [L_x(t, x, \dot{x})\varphi + L_v(t, x, \dot{x})\dot{\varphi}] dt.$$

Definizione 1.1. Chiamiamo **variazione prima di \mathcal{L} in x lungo la direzione φ** la quantità

$$\delta\mathcal{L}(x, \varphi) := \Phi'(0). \quad (1.1.1)$$

Notiamo che $\delta\mathcal{L}(x, \varphi) = \int_I [L_x(t, x, \dot{x})\varphi + L_v(t, x, \dot{x})\dot{\varphi}] dt$ è un funzionale lineare di $\varphi \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$.

Definizione 1.2. Una funzione $x \in \mathcal{C}^1(I, \mathbb{R}^n)$ è un **estremale debole di \mathcal{L}** se soddisfa

$$\int_I [L_x(t, x, \dot{x})\varphi + L_v(t, x, \dot{x})\dot{\varphi}] dt = 0 \quad (1.1.2)$$

per ogni $\varphi \in \mathcal{C}_c^\infty(I, \mathbb{R}^n)$.

Osserviamo immediatamente che se $x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ allora essere un estremale debole equivale ad essere un punto critico di \mathcal{L} , ovvero al fatto che

$$\delta\mathcal{L}(x, \varphi) = 0 \text{ per ogni } \varphi \in \mathcal{C}_c^\infty(I, \mathbb{R}^n).$$

Inoltre è chiaro che un minimo (anche non stretto) $x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ di \mathcal{L} è anche un estremale debole di \mathcal{L} .

Lemma 1.1 (Lemma fondamentale del Calcolo delle Variazioni). Sia $f : I \rightarrow \mathbb{R}$ una funzione continua e tale che

$$\int_I f(t)\psi(t) dt = 0 \text{ per ogni } \psi \in \mathcal{C}_c^\infty(I).$$

Allora $f \equiv 0$ su I .

Dimostrazione. Fissiamo $t_0 \in I$ e sia δ tale che $I_0 :=]t_0 - \delta, t_0 + \delta[\subseteq I$. Poichè $\mathcal{C}_c^\infty(I)$ è denso in $L^2(I)$ possiamo approssimare la funzione caratteristica di I_0 , che chiamiamo χ_0 , in norma $L^2(I)$ con funzioni $\mathcal{C}_c^\infty(I)$, ottenendo così che

$$\int_{I_0} f(t) dt = \int_I f(t) \chi_0(t) dt = 0$$

dal momento che f è continua su $\overline{I_0}$ e dunque $f|_{I_0} \in L^2(I_0)$. Concludiamo notando che, poichè f è continua,

$$f(t_0) = \lim_{\delta \rightarrow 0^+} \frac{1}{2\delta} \int_{t_0-\delta}^{t_0+\delta} f(t) dt = 0.$$

□

Siamo ora in grado di derivare l'equazione di Eulero.

Teorema 1.1. *Sia $x : I \rightarrow \mathbb{R}^n$ un estremo debole di \mathcal{L} di classe \mathcal{C}^2 . Supponiamo inoltre che $L \in \mathcal{C}^2(\mathcal{U})$ per un certo intorno \mathcal{U} del grafico di x . Allora x soddisfa l'equazione*

$$\frac{d}{dt} L_v(t, x(t), \dot{x}(t)) - L_x(t, x(t), \dot{x}(t)) = 0 \quad \text{per ogni } t \in I. \quad (1.1.3)$$

Dimostrazione. Dal momento che x è un estremo debole deve valere la (1.1.2), e integrando per parti si ottiene che

$$\int_I \left[-\frac{d}{dt} L_v(t, x(t), \dot{x}(t)) + L_x(t, x(t), \dot{x}(t)) \right] \varphi(t) dt = 0$$

per ogni $\varphi \in \mathcal{C}_c^\infty(I, \mathbb{R}^n)$. Scegliamo $\varphi = (\varphi^1, \dots, \varphi^n)$ in modo che siano costantemente nulle tutte le componenti φ^j per $j \neq i$ (i fissato). A questo punto, posto $\psi := \varphi^i \in \mathcal{C}_c^\infty(I)$, otteniamo

$$\int_I \left[-\frac{d}{dt} L_{v^i}(t, x(t), \dot{x}(t)) + L_{x^i}(t, x(t), \dot{x}(t)) \right] \psi(t) dt = 0$$

per ogni $\psi \in \mathcal{C}_c^\infty(I)$. Dal lemma fondamentale del Calcolo delle Variazioni 1.1 segue che

$$-\frac{d}{dt} L_{v^i}(t, x(t), \dot{x}(t)) + L_{x^i}(t, x(t), \dot{x}(t)) = 0$$

per ogni $i = 1, \dots, n$, e questo non è altro che la (1.1.3). □

Teorema 1.2 (Equazione di Eulero). *Supponiamo che $\mathcal{L}(x) \leq \mathcal{L}(y)$ per ogni $y \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ tale che $y(a) = x(a)$, $y(b) = x(b)$ e $\|x - y\|_{\mathcal{C}^1(I)} < \delta$ per un certo $\delta > 0$ sufficientemente piccolo. Supponiamo inoltre che $x \in \mathcal{C}^2(I, \mathbb{R}^n) \cap \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ e che $L \in \mathcal{C}^2$. Allora x soddisfa l'equazione di Eulero (1.1.3).*

Dimostrazione. Per ogni $\varphi \in \mathcal{C}_c^\infty(I, \mathbb{R}^n)$ tale che $\|\varphi\|_{\mathcal{C}^1(I)} < \delta$, e per $\epsilon \in]-1, 1[$ definiamo $\Phi(\epsilon) := \mathcal{L}(x + \epsilon\varphi)$. Per ipotesi Φ è derivabile in 0 e $\Phi(0) \leq \Phi(\epsilon)$ per ogni ϵ : ne segue che $\delta\mathcal{L}(x, \varphi) = \Phi'(0) = 0$ e quindi x è un estremales debole di \mathcal{L} . La tesi segue allora dal teorema 1.1. \square

1.2 Altre condizioni necessarie per i minimi

L'equazione di Eulero (1.1.3) è solo la prima e più famosa delle condizioni necessarie che deve soddisfare un minimo di \mathcal{L} di classe $\mathcal{C}^2(I, \mathbb{R}^n) \cap \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$. Vogliamo in questa sezione presentarne alcune altre cui faremo riferimento più avanti. Incominciamo con il seguente

Lemma 1.2 (Lemma di DuBoys-Reymond). *Supponiamo che $f \in L^1(I)$ sia tale che*

$$\int_I f(t)\dot{\varphi}(t) dt = 0 \quad (1.2.1)$$

per ogni $\varphi \in \mathcal{C}_c^\infty(I)$. Allora esiste $\alpha \in \mathbb{R}$ tale che $f(t) = \alpha$ per q.o. $t \in I$.

Dimostrazione. Siano $t_1, t_2 \in I$ fissati e supponiamo $a < t_1 < t_2 < b$. Preso $\delta > 0$ sufficientemente piccolo (in modo da garantire che $a < t_1 - \delta$ e $t_2 + \delta < b$) consideriamo la funzione φ lineare a tratti definita da $\varphi \equiv 0$ in $[a, t_1 - \delta] \cup [t_2 + \delta, b]$, $\varphi \equiv \delta$ in $[t_1, t_2]$ e lineare nei due rimanenti intervalli $[t_1 - \delta, t_1]$ e $[t_2, t_2 + \delta]$. Notiamo che $\dot{\varphi}$ vale 0 su $[a, t_1 - \delta] \cup [t_1, t_2] \cup [t_2 + \delta, b]$ mentre su $[t_1 - \delta, t_1]$ e $[t_2, t_2 + \delta]$ vale 1 e -1 rispettivamente. Approssimando φ in norma \mathcal{C}^1 con funzioni \mathcal{C}_c^∞ e passando al limite nella (1.2.1) otteniamo che

$$\int_I f(t)\dot{\varphi}(t) dt = 0$$

e quindi

$$\int_{t_1-\epsilon}^{t_1} f(t)dt - \int_{t_2}^{t_2+\epsilon} f(t)dt = 0.$$

Dividendo per δ la precedente relazione, scegliendo t_1 e t_2 fra i punti di Lebesgue di f e passando al limite per $\delta \rightarrow 0^+$ otteniamo che $f(t_1) = f(t_2)$. La validità del lemma è dunque dimostrata in quanto i punti di Lebesgue di una funzione L^1 sono quasi tutti. \square

Proposizione 1.1. Sia $x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ un estremaie debole di \mathcal{L} come nella definizione 1.2. Allora esiste $c \in \mathbb{R}^n$ tale che

$$L_v(t, x(t), \dot{x}(t)) = c + \int_a^t L_x(s, x(s), \dot{x}(s)) ds \quad (1.2.2)$$

per ogni $t \in I$.

Dimostrazione. Riscriviamo l'equazione di estremaie debole:

$$\int_a^b [L_x(t, x, \dot{x})\varphi + L_v(t, x, \dot{x})\dot{\varphi}] dt = 0$$

per ogni $\varphi \in \mathcal{C}_c^\infty(I)$. Integrando per parti il primo membro dell'integrale otteniamo

$$\int_a^b \left[- \int_a^t L_x(s, x(s), \dot{x}(s)) ds + L_v(t, x(t), \dot{x}(t)) \right] \dot{\varphi}(t) dt = 0$$

per ogni $\varphi \in \mathcal{C}_c^\infty(I)$; la tesi segue allora dal lemma di DuBoys-Reymond 1.2. \square

Corollario 1.1. Ogni estremaie debole $x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ di \mathcal{L} soddisfa l'equazione di Eulero

$$\frac{d}{dt} L_v(t, x(t), \dot{x}(t)) - L_x(t, x(t), \dot{x}(t)) = 0 \quad \text{per ogni } t \in I.$$

Dimostrazione. Notiamo che la funzione al secondo membro della (1.2.2) è di classe $\mathcal{C}^1(\bar{I}, \mathbb{R}^n)$, dunque lo è anche il primo membro. L'equazione di Eulero si ricava allora direttamente differenziando la (1.2.2). \square

Chiameremo la (1.2.2) **equazione di DuBoys-Reymond** o **equazione di Eulero in forma integrale**.

Ragionando come nella proposizione 1.1 e nel corollario 1.1 otteniamo la seguente

Proposizione 1.2. Sia $x \in \text{Lip}(I, \mathbb{R}^n)$ un estremaie debole per \mathcal{L} , cioè tale che $\delta\mathcal{L}(x, \varphi) = 0$ per ogni $\varphi \in \mathcal{C}_c^\infty(I, \mathbb{R}^n)$. Allora esiste un vettore $c \in \mathbb{R}^n$ tale che

$$L_v(t, x(t), \dot{x}(t)) = c + \int_a^t L_x(s, x(s), \dot{x}(s)) ds \quad \text{q.o. } t \in I. \quad (1.2.3)$$

Vale inoltre l'equazione di Eulero

$$\frac{d}{dt} L_v(t, x(t), \dot{x}(t)) - L_x(t, x(t), \dot{x}(t)) = 0 \quad \text{q.o. } t \in I. \quad (1.2.4)$$

1.3 Calibrazioni

Sia L una Lagrangiana di classe $\mathcal{C}^2(\mathbb{R}, \mathbb{R}^n, \mathbb{R}^n)$ e consideriamo l'integrale variazionale

$$\mathcal{L}(x) = \int_I L(t, x(t), \dot{x}(t)) dt$$

definito su una certa sottoclasse \mathcal{C} di $\mathcal{C}^1(\bar{I}, \mathbb{R}^n)$: ad esempio \mathcal{C} può essere $\mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ stessa, oppure le funzioni di tale classe con valori agli estremi prefissati. Infine sia $x_0 \in \mathcal{C}$ una funzione fissata.

Definizione 1.3. Una *calibrazione* per la tripla (L, x_0, \mathcal{C}) è una Lagrangiana $M(t, x, v)$ che soddisfa le seguenti condizioni:

1. per ogni $t \in I$ e ogni $x \in \mathcal{C}$ abbiamo

$$M(t, x(t), \dot{x}(t)) \leq L(t, x(t), \dot{x}(t))$$

mentre vale il segno di uguale per $x = x_0$.

2. l'integrale variazionale associato ad M

$$\mathcal{M}(x) := \int_I M(t, x(t), \dot{x}(t)) dt$$

è costante su \mathcal{C} , ovvero $\mathcal{M}(x_1) = \mathcal{M}(x_2)$ per ogni $x_1, x_2 \in \mathcal{C}$.

Proposizione 1.3. Supponiamo che la tripla (L, x_0, \mathcal{C}) ammetta una calibrazione M . Allora x_0 è un minimo di \mathcal{L} .

Dimostrazione. Per ogni $x \in \mathcal{C}$ abbiamo

$$\begin{aligned} \mathcal{L}(x) &= \int_I L(t, x(t), \dot{x}(t)) dt \geq \int_I M(t, x(t), \dot{x}(t)) dt \\ &= \int_I M(t, x_0(t), \dot{x}_0(t)) dt = \int_I L(t, x_0(t), \dot{x}_0(t)) dt \\ &= \mathcal{L}(x_0). \end{aligned}$$

□

Sia M una calibrazione per (L, x_0, \mathcal{C}) , e supponiamo che per ogni $x \in \mathcal{C}$ e ogni $\varphi \in \mathcal{C}_c^\infty(I, \mathbb{R}^n)$ si abbia $x + \epsilon\varphi \in \mathcal{C}$ per ogni ϵ sufficientemente piccolo. Allora $\mathcal{M}(x + \epsilon\varphi) = \mathcal{M}(x)$ e quindi

$$\delta\mathcal{M}(x, \varphi) = 0 \text{ per ogni } \varphi \in \mathcal{C}_c^\infty(I, \mathbb{R}^n).$$

Dal lemma fondamentale deduciamo allora che M è una **Lagrangiana nulla**, ovvero

$$M_x(t, x(t), \dot{x}(t)) - \frac{d}{dt}M_v(t, x(t), \dot{x}(t)) = 0$$

per ogni $x \in \mathcal{C}_c^2(I, \mathbb{R}^n) \cap \mathcal{C}$ e ogni $t \in I$.

Una eventuale calibrazione deve quindi essere cercata nella classe delle Lagrangiane nulle. Fra esse vi sono tutte quelle della forma

$$M(t, x, v) = S_t(t, x) + S_{x^i}(x, z)v^i$$

per una certa funzione S di classe \mathcal{C}^2 , cioè

$$\mathcal{M}(x) = \int_a^b \frac{d}{dt}S(t, x(t))dt = S(b, x(b)) - S(a, x(a))$$

per ogni $x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n)$ il cui grafico è contenuto nel dominio di S .

1.4 Campi di estremali

Sia $\mathcal{C}(a, b) := \{x \in \mathcal{C}^1(\bar{I}, \mathbb{R}^n) : x(a) = x_a, x(b) = x_b\}$ dove $x_a, x_b \in \mathbb{R}^n$ sono due valori al bordo fissati, e consideriamo il seguente problema: dato un estremoale $x_0 \in \mathcal{C}(a, b) \cap \mathcal{C}^2(I, \mathbb{R}^n)$ per il funzionale \mathcal{L} , sotto quali condizioni esiste una calibrazione M per $(L, x_0, \mathcal{C}_\delta(a, b))$, in modo cioè che x_0 sia un minimo di L nella classe $\mathcal{C}_\delta(a, b) := \mathcal{C}(a, b) \cap \{y : \|y - x_0\|_{L^\infty(I)} < \delta\}$?

Una prima condizione che richiediamo è che x_0 sia immerso in un campo di estremali come nella seguente

Definizione 1.4. Diciamo che x_0 è **immerso in un campo di estremali** di \mathcal{L} se

1. esiste un dominio semplicemente connesso $\Gamma = \{(t, c) : c \in Q, t \in I(c)\}$ dove $Q \subset \mathbb{R}^n$ è un insieme non vuoto di parametri e $I(c) \subset \mathbb{R}$ è un intervallo
2. esiste un diffeomorfismo $f : \Gamma \rightarrow G \subset \mathbb{R} \times \mathbb{R}^n$ di classe \mathcal{C}^1 e della forma

$$f(t, c) = (t, \varphi(t, c))$$

con $\dot{\varphi} = \frac{\partial \varphi}{\partial t} \in \mathcal{C}^1(\Gamma, \mathbb{R}^n)$ e tale che

$$x_0(t) = \varphi(t, c_0) \text{ per ogni } t \in \bar{I}$$

per un certo $c_0 \in Q$ con $[a, b] \subset I(c_0)$.

3. per ogni $c \in Q$ la funzione $\varphi(\cdot, c)$ è un estremo di \mathcal{L} , ovvero

$$L_x(\cdot, \varphi(\cdot, c), \dot{\varphi}(\cdot, c)) - \frac{d}{dt}L_v(\cdot, \varphi(\cdot, c), \dot{\varphi}(\cdot, c)) = 0$$

Tale f viene chiamato **campo di estremali** su G .

Chiameremo **funzione pendenza** del campo f la funzione $\mathcal{P} : G \rightarrow \mathbb{R}^n$ definita da $\mathcal{P} := \dot{\varphi} \circ f^{-1} \in \mathcal{C}^1(G, \mathbb{R}^n)$. Si ha $\dot{\varphi} = \mathcal{P} \circ f$ su tutto Γ , e in particolare

$$\dot{x}_0(t) = \mathcal{P}(t, x_0(t)) \text{ per ogni } t \in [a, b].$$

Chiameremo **campo di pendenza** di f il campo

$$\wp : G \ni (t, x) \mapsto (t, x, \mathcal{P}(t, x)) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$$

Definizione 1.5. Un campo di estremali f su G si dice **campo ottimale** su G se esiste una funzione $S \in \mathcal{C}^2(G, \mathbb{R})$ tale che, posto

$$\tilde{L}(t, x, v) := L(t, x, v) - S_t(t, x) - S_x(t, x) \cdot v, \quad (1.4.1)$$

si ha

$$\tilde{L}(t, x, v) \geq 0 \text{ per ogni } (t, x, v) \in G \times \mathbb{R}^n \text{ e} \quad (1.4.2)$$

$$\tilde{L} \circ \wp(t, x) = \tilde{L}(t, x, \mathcal{P}(t, x)) = 0 \text{ per ogni } (t, x) \in G. \quad (1.4.3)$$

Tale S viene chiamata **funzione eikonale** di f .

Se poi $\tilde{L}(t, x, v) > 0$ per ogni $(t, x) \in G$ e ogni $v \neq \mathcal{P}(t, x)$, f viene detto **campo di Weierstrass** su G .

Definizione 1.6. Diciamo che x è **adattata** a f se

$$\dot{x}(t) = \mathcal{P}(t, x(t)).$$

Definiamo la classe $\mathcal{C} := \mathcal{C}(a, b) \cap \{x : \text{il grafico di } x \text{ è contenuto in } G\}$ e supponiamo che $x_0 \in \mathcal{C}$ sia adattata ad un campo ottimale f con eikonale S . Definiamo una Lagrangiana nulla M su $G \times \mathbb{R}^n$ tramite

$$M(t, x, v) := S_t(t, x) + S_x(t, x) \cdot v.$$

Allora la Lagrangiana \tilde{L} definita nella (1.4.1) si può scrivere come $\tilde{L} = L - M$, e quindi la (1.4.2) implica che M soddisfa la condizione 1. richiesta dalla definizione 1.3 per essere una calibrazione per (L, x_0, \mathcal{C}) . La condizione 2. è sicuramente soddisfatta, in quanto

$$\begin{aligned}
\mathcal{M}(x) &= \int_I M(t, x(t), \dot{x}(t)) dt \\
&= \int_I S_t(t, x(t)) + S_x(t, x(t)) \cdot \dot{x}(t) dt \\
&= \int_I \frac{d}{dt} S(t, x(t)) dt \\
&\equiv S(b, x_b) - S(a, x_a)
\end{aligned}$$

Dunque M è una calibrazione per (L, x_0, \mathcal{C}) : ne segue che x_0 è un minimo di \mathcal{L} in \mathcal{C} , da cui

$$\mathcal{L}(x) \geq \mathcal{L}(x_0)$$

per ogni $x \in \mathcal{C}$. Ripercorrendo la dimostrazione della proposizione 1.3, notiamo immediatamente che se f è di Weierstrass vale la disuguaglianza stretta per $x \neq x_0$. Poichè x_0 è adattata al campo ottimale f vale

$$\begin{aligned}
&L(\cdot, x_0(\cdot), \dot{x}_0(\cdot)) - S_t(\cdot, x_0(\cdot)) - S_x(\cdot, x_0(\cdot)) = \\
&= \tilde{L}(\cdot, x_0(\cdot), \dot{x}_0(\cdot)) \\
&= \tilde{L}(\cdot, x_0(\cdot), \mathcal{P}(\cdot, x_0(\cdot))) \\
&= \tilde{L} \circ \wp(\cdot, x_0(\cdot)) \\
&= 0
\end{aligned}$$

ovvero $L(\cdot, x_0(\cdot), \dot{x}_0(\cdot)) = M(\cdot, x_0(\cdot), \dot{x}_0(\cdot))$. Quindi

$$\mathcal{L}(x_0) = \mathcal{M}(x_0) = \mathcal{M}(x) = S(b, x_b) - S(a, x_a)$$

per ogni $x \in \mathcal{C}$, che assieme al fatto che x_0 è un minimo ci dà

$$\mathcal{L}(x) \geq S(b, x_b) - S(a, x_a).$$

Se poi f è di Weierstrass vale la disuguaglianza stretta per ogni $x \in \mathcal{C}$ che non è adattata a f .

Riassumiamo quanto appena detto nel seguente

Teorema 1.3. *Posto*

$$\mathcal{C} := \mathcal{C}(a, b) \cap \mathcal{C}^2(I, \mathbb{R}^n) \cap \{x : \text{il grafico di } x \text{ è contenuto in } G\},$$

se esistono un campo ottimale f su G con eikonale S ed una $x_0 \in \mathcal{C}$ adattata a f , allora

$$\mathcal{L}(x) \geq \mathcal{L}(x_0) \tag{1.4.4}$$

per ogni $x \in \mathcal{C}$, con disuguaglianza stretta se f è di Weierstrass e x non è adattata. Si ha $\mathcal{L}(x_0) = S(b, x_b) - S(a, x_a)$ e quindi

$$\mathcal{L}(x) \geq S(b, x_b) - S(a, x_a) \quad (1.4.5)$$

per ogni $x \in \mathcal{C}$. Se poi x è adattata al campo f , allora vale

$$\mathcal{L}(x) = S(b, x_b) - S(a, x_a) \quad (1.4.6)$$

Se f è un campo di Weierstrass su G , allora nella (1.4.5) vale la disuguaglianza stretta per ogni $x \in \mathcal{C}$ che non è adattata a f (per cui cioè esiste un $t_0 \in \bar{I}$ tale che $\dot{x}(t_0) \neq \mathcal{P}(t_0, x(t_0))$).

Una condizione sufficiente affinché un estremo x_0 sia un minimo di \mathcal{L} nella classe \mathcal{C} è dunque data dal poter immergere x_0 in un campo ottimale, in modo che x_0 sia adattata ad esso. Il problema successivo che ci poniamo è naturalmente quello di capire quando questo si può fare; la prima cosa che faremo sarà quella di ricavare condizioni necessarie che deve soddisfare un campo ottimale.

Usando le stesse notazioni utilizzate finora osserviamo che $\tilde{L}(t, x, v) \geq 0$ e che vale il segno di uguale per $v = \mathcal{P}(t, x)$: perciò $\tilde{L}_v(t, x, \mathcal{P}(t, x)) = 0$. Inoltre dalla definizione di \tilde{L} abbiamo $\tilde{L}_v(t, x, v) = L_v(t, x, v) - S_x(t, x)$ da cui

$$S_x(t, x) = L_v(t, x, \mathcal{P}(t, x)). \quad (1.4.7)$$

Ricordando che $\tilde{L} \circ \wp = 0$, e sfruttando la (1.4.7) otteniamo

$$\begin{aligned} 0 &= \tilde{L}(t, x, \mathcal{P}(t, x)) \\ &= L(t, x, \mathcal{P}(t, x)) - S_t(t, x) - S_x(t, x) \cdot \mathcal{P}(t, x) \\ &= L(t, x, \mathcal{P}(t, x)) - S_t(t, x) - L_v(t, x, \mathcal{P}(t, x)) \cdot \mathcal{P}(t, x) \end{aligned}$$

ovvero

$$S_t(t, x) = L(t, x, \mathcal{P}(t, x)) - L_v(t, x, \mathcal{P}(t, x)) \cdot \mathcal{P}(t, x). \quad (1.4.8)$$

La (1.4.7) e la (1.4.8) sono dette **equazioni di Carathéodory**.

Definizione 1.7. Un campo di estremali f su G con campo di pendenza $\wp(t, x) = (t, x, \mathcal{P}(t, x))$ è detto **campo di Mayer** se esiste una funzione $S \in \mathcal{C}^2(G)$ che soddisfa le equazioni di Carathéodory (1.4.7) e (1.4.8), ovvero

$$S_t = L \circ \wp - (L_v \circ \wp) \cdot \mathcal{P} \text{ su } G$$

$$S_x = L_v \circ \wp \text{ su } G.$$

Tale S viene detta **eikonale** del campo f .

Osservazione 1.2. Sia γ la 1-forma su $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ definita da

$$\gamma := (L - v \cdot L_v)dt + L_{v^i}dx^i.$$

Essa viene anche chiamata **forma di Beltrami** associata a L .

Sia $\varphi^*\gamma$ il pull-back di γ rispetto a φ , ovvero

$$\varphi^*\gamma = (L(\varphi) - \mathcal{P} \cdot L_v(\varphi)) dt + L_{v^i}(\varphi) dx^i.$$

Le equazioni di Carathéodory sono dunque equivalenti a $dS = \varphi^*\gamma$, ovvero al fatto che $\varphi^*\gamma$ è una 1-forma esatta. Poichè G è semplicemente connesso questo è equivalente a dire che $\varphi^*\gamma$ è chiusa, cioè $d(\varphi^*\gamma) = 0$. Dunque un campo è di Mayer se e solo se $d(\varphi^*\gamma) = 0$.

Osserviamo inoltre che, poichè $dS = \varphi^*\gamma$, l'eikonale S di un campo di Mayer è univocamente determinata a meno di costanti.

Le considerazioni precedenti ci dicono che un campo ottimale su G con eikonale S e campo di pendenza φ è un campo di Mayer su G con eikonale S . Cerchiamo dunque condizioni affinché sia vera l'implicazione inversa: ricordando la definizione 1.5, quello che ci serve è avere $\tilde{L}(t, x, v) \geq 0$ su $G \times \mathbb{R}^n$, con segno di uguale se $v = \mathcal{P}(t, x)$. Per un campo di Mayer abbiamo

$$\begin{aligned} \tilde{L}(t, x, v) &= L(t, x, v) - S_t(t, x) - S_x(t, x) \cdot v \\ &= L(t, x, v) - L \circ \varphi(t, x) + (L_v \circ \varphi(t, x)) \cdot \mathcal{P}(t, x) - L_v \circ \varphi(t, x) \cdot v \\ &= L(t, x, v) - L(t, x, \mathcal{P}(t, x)) - L_v(t, x, \mathcal{P}(t, x)) \cdot (v - \mathcal{P}(t, x)). \end{aligned}$$

Definiamo pertanto la funzione

$$\begin{aligned} \mathcal{E} &: G \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \\ (t, x, v, w) &\mapsto L(t, x, v) - L(t, x, w) - L_v(t, x, w) \cdot (v - w). \end{aligned}$$

Notiamo che $\mathcal{E}(t, x, v, v) = 0$ per ogni $(t, x) \in G$ e ogni $v \in \mathbb{R}^n$. Se $L(t, x, v)$ è una funzione convessa in v per ogni $(t, x) \in G$ otteniamo immediatamente che $\mathcal{E}(t, x, v, w) \geq 0$. In particolare, poichè per costruzione $\tilde{L}(t, x, v) = \mathcal{E}(t, x, v, \mathcal{P}(t, x))$, ricaviamo esattamente quello che ci serve, ovvero $\tilde{L}(t, x, v) \geq 0$ su $G \times \mathbb{R}^n$, con segno di uguale se $v = \mathcal{P}(t, x)$.

Se poi L_{vv} è definita positiva per ogni (t, x, v) , ovvero se esiste un $\epsilon > 0$ tale che

$$L_{v^i v^j}(t, x, v) \xi^i \xi^j \geq \epsilon |\xi|^2 \text{ per ogni } \xi \in \mathbb{R}^n \quad (1.4.9)$$

allora L è addirittura strettamente convessa in v , dunque $\mathcal{E}(t, x, v, w) > 0$ se $v \neq w$, e quindi $\tilde{L}(t, x, v) > 0$ a meno che $v = \mathcal{P}(t, x)$. Queste considerazioni ci portano alla

Proposizione 1.4. *Se la Lagrangiana $L(t, x, v)$ è una funzione convessa di v per ogni $(t, x) \in G$, allora un campo di Mayer su G è anche un campo ottimale. Se poi $L_{vv} > 0$ il campo di Mayer è anche di Weierstrass.*

Ricordando il teorema 1.3 arriviamo infine al seguente

Teorema 1.4. *Sia x_0 un estremo dell'integrale variazionale \mathcal{L} con Lagrangiana $L(t, x, v)$ relativa convessa nella terza variabile, e tale che x_0 è adattata ad un campo di Mayer. Allora*

$$\mathcal{L}(x) \geq \mathcal{L}(x_0)$$

per ogni $x \in \mathcal{C}$. Se poi $L_{vv} > 0$ vale la disuguaglianza stretta per $x \in \mathcal{C}, x \neq x_0$.

Cerchiamo dunque di capire quando possiamo immergere in un campo di Mayer un estremo di un funzionale come nel teorema 1.4. Vedremo che la condizione chiave sarà un'ipotesi di ellitticità su L_{vv} come la (1.4.9). Prima però facciamo alcune considerazioni preliminari.

Definizione 1.8. *L'operatore di Eulero E_L è definito da*

$$E_L(x) := L_x(\cdot, x, \dot{x}) - \frac{d}{dt} L_v(\cdot, x, \dot{x}).$$

Definizione 1.9. *Dato un campo $f : \Gamma \rightarrow G$ della forma $f(t, c) = (t, \varphi(t, c))$ definiamo il **flusso di fase** $e : \Gamma \rightarrow G \times \mathbb{R}^n$ dato da*

$$e(t, c) := (t, \varphi(t, c), \dot{\varphi}(t, c)).$$

Siano $\eta_i : \Gamma \rightarrow \mathbb{R}$ i **momenti canonici** dati da $\eta_i(t, c) = L_{v^i}(e(t, c))$, e poniamo

$$\begin{aligned} \eta & : \Gamma \rightarrow \mathbb{R}^n \\ (t, c) & \rightarrow L_v(e) = (\eta_1(t, c), \dots, \eta_n(t, c)) \end{aligned}$$

Definiamo per $1 \leq i, j \leq n$ le **parentesi di Lagrange** $[c^i, c^j] : \Gamma \rightarrow \mathbb{R}$ date da

$$[c^i, c^j] := \frac{\partial \eta}{\partial c^i} \cdot \frac{\partial \varphi}{\partial c^j} - \frac{\partial \eta}{\partial c^j} \cdot \frac{\partial \varphi}{\partial c^i}.$$

Proposizione 1.5. *Un campo $f(t, c) = (t, \varphi(t, c))$ su un dominio G semplicemente connesso è di Mayer se e solo se valgono le condizioni*

$$E_L(\varphi) \cdot \varphi_{c^i} = 0 \quad \text{per ogni } 1 \leq i \leq n \quad (1.4.10)$$

$$[c^i, c^j] = 0 \quad \text{per ogni } 1 \leq i, j \leq n \quad (1.4.11)$$

Dimostrazione. Ricordando l'osservazione 1.2 sappiamo che f è di Mayer se e solo se $d(\varphi^*\gamma) = 0$, che è equivalente a $f^*d(\varphi^*\gamma) = 0$ in quanto f è un diffeomorfismo. Poichè

$$f^*d(\varphi^*\gamma) = d(f^*(\varphi^*\gamma)) = d((\varphi \circ f)^*\gamma) = d(e^*\gamma)$$

otteniamo che f è di Mayer se e solo se $d(e^*\gamma) = 0$. Ricordando le definizioni di e e di γ abbiamo

$$\begin{aligned} d(e^*\gamma) &= d[(L(e) - \dot{\varphi} \cdot L_v(e))dt + L_{v^k}(e)d\varphi^k] \\ &= d[L(e)dt + L_{v^k}(e)(d\varphi^k - \dot{\varphi}^k dt)] \\ &= L_{x^k}(e)d\varphi^k \wedge dt + L_{v^k}(e)d\dot{\varphi}^k \wedge dt + \\ &\quad + dL_{v^k}(e)(d\varphi^k - \dot{\varphi}^k dt) - L_{v^k}(e)d\dot{\varphi}^k \wedge dt \\ &= L_{x^k}(e)d\varphi^k \wedge dt + dL_{v^k}(e)(d\varphi^k - \dot{\varphi}^k dt) \end{aligned}$$

Notiamo che $d\varphi^k = \dot{\varphi}^k dt + \varphi_{c^i}^k dc^i$ e dunque

$$\begin{aligned} d(e^*\gamma) &= L_{x^k}(e)\varphi_{c^i}^k dc^i \wedge dt + dL_{v^k}(e)\varphi_{c^i}^k dc^i \\ &= \left(L_{x^k}(e) - \frac{\partial L_{v^k}}{\partial t}(e) \right) \varphi_{c^i}^k dc^i \wedge dt + \frac{\partial L_{v^k}}{\partial c^j}(e) \frac{\partial \varphi^k}{\partial c^j}(e) dc^j \wedge dc^i \\ &= E_L(\varphi) \cdot \varphi_{c^i}^k dc^i \wedge dt + \frac{\partial \eta}{\partial c^j} \cdot \frac{\partial \varphi}{\partial c^i} dc^j \wedge dc^i \end{aligned} \quad (1.4.12)$$

Dunque $d(e^*\gamma) = 0$ se e solo se valgono la (1.4.10) e la (1.4.11). La proposizione è dunque dimostrata. \square

Siamo ora in grado di enunciare condizioni sufficienti affinché un estremo possa essere immerso in un campo di Mayer. Tali condizioni garantiscono che il campo è addirittura ottimale, quindi l'estremo risulta un minimo locale.

Teorema 1.5. *Sia x un estremo di una Lagrangiana L di classe \mathcal{C}^3 che soddisfa la condizione di ellitticità*

$$L_{v^i v^j}(t, x(t), v) \xi^i \xi^j \geq \epsilon |\xi|^2 \text{ per ogni } \xi \in \mathbb{R}^n \quad (1.4.13)$$

per un certo $\epsilon > 0$ e per ogni $t \in [a, b]$, $v \in \mathbb{R}^n$ e $\xi \in \mathbb{R}^n$. Allora ogni tratto sufficientemente corto di x si può immergere in un campo di Mayer, ovvero ogni $t_0 \in [a, b]$ ha un intorno I tale che $x|_I$ si può immergere in un campo di Mayer.

Dimostrazione. Supponiamo $t_0 \in (a, b)$ (i casi $t_0 = a$ e $t_0 = b$ si trattano allo stesso modo). Scegliamo $t^* < t_0$ sufficientemente vicino a t_0 (in un senso che preciseremo), e definiamo $x^* := x(t^*)$. Per la (1.4.13) $\det L_{vv} \neq 0$ su un insieme del tipo $U \times \mathbb{R}^n$

per un certo intorno $U \subset \mathbb{R} \times \mathbb{R}^n$ del grafico di x . Questo ci consente di affermare che risolvere l'equazione di Eulero

$$E_L(\varphi) = L_x(\cdot, \varphi, \dot{\varphi}) - \frac{d}{dt}L_v(\cdot, \varphi, \dot{\varphi})$$

è equivalente a risolvere un'equazione ordinaria del tipo $\ddot{\varphi} = F(t, \varphi, \dot{\varphi})$, con $F \in \mathcal{C}^1$. Per risultati standard sulle equazioni differenziali ordinarie abbiamo che il problema

$$\begin{cases} E_L(\varphi(\cdot, c)) = 0 \\ \varphi(t^*, c) = x^* \\ \dot{\varphi}(t^*, c) = c \end{cases} \quad (1.4.14)$$

ammette un'unica soluzione $\varphi(t, c)$ in un intorno di (t^*, \bar{c}) che supponiamo della forma $\Gamma_\delta^* := (t^* - \delta, t^* + \delta) \times (\bar{c} - \delta, \bar{c} + \delta)$, dove abbiamo posto $\bar{c} := \dot{\varphi}(t^*)$. Inoltre $\dot{\varphi}$ e $\dot{\varphi}_c$ sono continui su Γ_δ^* . Espandendo in serie di Taylor otteniamo $\varphi(t, c) = x^* + c(t - t^*) + O(|t - t^*|^2)$, da cui $\varphi_c(t, c) = (t - t^*)\text{Id} + O(|t - t^*|^2)$ e quindi $\det \varphi_c(t, c) = (t - t^*)^n + O(|t - t^*|^2)$. Ponendo $f(t, c) := (t, \varphi(t, c))$ abbiamo

$$\begin{aligned} \det Df(t, c) &= \det \varphi_c(t, c) \\ &= (t - t^*)^n + O(|t - t^*|^{n+1}) \quad \text{per ogni } (t, c) \in \Gamma_\delta^*. \end{aligned} \quad (1.4.15)$$

Notiamo che la φ , e quindi anche la f , dipende dal parametro t^* . Tuttavia un ragionamento di continuità basato sulla (1.4.15) ci dice che esistono $\mu, \nu > 0$ tali che per ogni $t^* \in (t_0 - \nu, t_0)$ si ha

$$\det Df(t, c^*) \neq 0 \quad \text{per ogni } t \in (t^*, t^* + \mu) \text{ e per } c^* := x(t^*).$$

Dunque ci premuriamo di scegliere t^* in maniera che $t^* < t_0 < t^* + \mu$, in modo che risolvendo la (1.4.14) con questo t^* e definendo la f relativa a tale φ si ottenga $\det Df(t_0, c^*) \neq 0$, dove come prima $c^* = x(t^*)$. In questo modo troviamo un intorno Γ di (t_0, c^*) , che supponiamo della forma $\Gamma := (t_0 - \rho, t_0 + \rho) \times (c^* - \rho, c^* + \rho)$, tale che $f|_\Gamma$ è un diffeomorfismo con l'immagine $G := f(\Gamma)$. Inoltre $x(t) = \varphi(t, c^*)$ per $t \in (t^* - \mu, t^* + \mu)$ e quindi per un certo $0 < \epsilon < \rho$ abbiamo che $x|_{[t_0 - \epsilon, t_0 + \epsilon]}$ è immerso nel campo di integrali $f|_\Gamma$. Dimostriamo che esso è un campo di Mayer.

Poichè Γ è semplicemente connesso, è sufficiente dimostrare che φ verifica la (1.4.10) e la (1.4.11). La prima è sicuramente verificata per costruzione, in quanto φ risolve il sistema (1.4.14). Per quanto riguarda la (1.4.11), ponendo come in precedenza $e(t, c) := (t, \varphi(t, c), \dot{\varphi}(t, c))$ e usando la notazione L_x, L_v, L_{xx} , etc. per $L_x(e), L_v(e), L_{xx}(e)$, etc., abbiamo $L_x = \frac{\partial}{\partial t}L_v$ e dunque

$$\begin{aligned}
\frac{\partial}{\partial t}[c^i, c^j] &= \frac{\partial}{\partial c^i} \frac{\partial}{\partial t} L_v \cdot \frac{\partial \varphi}{\partial c^j} + \frac{\partial}{\partial c^i} L_v \cdot \frac{\partial \dot{\varphi}}{\partial c^j} - \frac{\partial}{\partial c^j} \frac{\partial}{\partial t} L_v \cdot \frac{\partial \varphi}{\partial c^i} - \frac{\partial}{\partial c^j} L_v \cdot \frac{\partial \dot{\varphi}}{\partial c^i} \\
&= \frac{\partial}{\partial c^i} L_x \cdot \frac{\partial \varphi}{\partial c^j} + \frac{\partial}{\partial c^i} L_v \cdot \frac{\partial \dot{\varphi}}{\partial c^j} - \frac{\partial}{\partial c^j} L_x \cdot \frac{\partial \varphi}{\partial c^i} - \frac{\partial}{\partial c^j} L_v \cdot \frac{\partial \dot{\varphi}}{\partial c^i} \\
&= L_{xx} \frac{\partial \varphi}{\partial c^i} \cdot \frac{\partial \varphi}{\partial c^j} + L_{xv} \frac{\partial \dot{\varphi}}{\partial c^i} \cdot \frac{\partial \varphi}{\partial c^j} + L_{vx} \frac{\partial \varphi}{\partial c^i} \cdot \frac{\partial \dot{\varphi}}{\partial c^j} + L_{vv} \frac{\partial \dot{\varphi}}{\partial c^i} \cdot \frac{\partial \dot{\varphi}}{\partial c^j} + \\
&\quad - L_{xx} \frac{\partial \varphi}{\partial c^j} \cdot \frac{\partial \varphi}{\partial c^i} - L_{xv} \frac{\partial \dot{\varphi}}{\partial c^j} \cdot \frac{\partial \varphi}{\partial c^i} - L_{vx} \frac{\partial \varphi}{\partial c^j} \cdot \frac{\partial \dot{\varphi}}{\partial c^i} - L_{vv} \frac{\partial \dot{\varphi}}{\partial c^j} \cdot \frac{\partial \dot{\varphi}}{\partial c^i} \\
&= 0.
\end{aligned}$$

Quindi le parentesi di Lagrange dipendono solo da c e non da t . Dalla (1.4.14) $\varphi(t^*, c) = z^*$, da cui $\frac{\partial \varphi}{\partial c^i}(t^*, c) = 0$ per ogni $i = 1, \dots, n$. Di conseguenza

$$[c^i, c^j](t, c) = [c^i, c^j](t^*, c) = 0.$$

In questo modo vale la (1.4.11) e dunque il teorema è dimostrato. \square

Ricordando il teorema 1.4 otteniamo il

Teorema 1.6. *Sia L una Lagrangiana di classe \mathcal{C}^3 e sia $x(t)$, $a \leq t \leq b$ un estremo di L che soddisfa la condizione di ellitticit *

$$L_{v^i v^j}(t, x(t), v) \xi^i \xi^j \geq \epsilon |\xi|^2 \text{ per ogni } \xi \in \mathbb{R}^n \quad (1.4.16)$$

per un certo $\epsilon > 0$ e per ogni $t \in [a, b]$, $v \in \mathbb{R}^n$ e $\xi \in \mathbb{R}^n$. Allora per ogni $t_0 \in [a, b]$ esistono un intervallo $I_0 = [a_0, b_0] \subset [a, b]$, $t_0 \in \overset{\circ}{I}_0$ (se $t_0 = a$ o b allora $t_0 = a_0$ o b_0 rispettivamente) e un $\epsilon > 0$ tale che

$$\int_{a_0}^{b_0} L(t, x, \dot{x}) dt < \int_{a_0}^{b_0} L(t, y, \dot{y}) dt$$

per ogni $y \in \mathcal{C}^1([a_0, b_0], \mathbb{R}^n)$ per cui $y(a_0) = x(a_0)$, $y(b_0) = x(b_0)$ e $\|x - y\|_{L^\infty(I_0)} < \epsilon$. In altre parole, ogni tratto di x sufficientemente corto   un minimo locale.

Capitolo 2

Un teorema di regolarità parziale di Tonelli

In questo capitolo presentiamo un risultato di Tonelli, che dà regolarità parziale per minimi di integrali variazionali sotto certe ipotesi sulla Lagrangiana. Il teorema è il seguente:

Teorema 2.1 (di regolarità parziale di Tonelli). *Sia $L(t, x, v)$ una Lagrangiana di classe C^∞ tale che $L_{vv}(t, x, v)$ sia definita positiva per ogni (t, x, v) , e supponiamo che $x \in AC(I, \mathbb{R}^n)$ sia un minimo locale stretto dell'integrale variazionale*

$$\mathcal{L}(x) = \int_I L(t, x(t), \dot{x}(t)) dt$$

nella classe $\mathcal{C} := \{y \in AC(I) : y(a) = x(a) \text{ e } y(b) = x(b)\}$. Allora esiste una funzione

$$[\dot{x}] : \bar{I} \longrightarrow (\bar{\mathbb{R}})^n = (\mathbb{R} \cup \{-\infty, +\infty\})^n$$

che è continua e rappresenta la derivata di x (nel senso di limite dei rapporti incrementali). Inoltre l'insieme $\Sigma := \{t \in \bar{I} : [\dot{x}(t)] = \pm\infty\}$ è un chiuso di misura nulla e $x \in C^\infty(\bar{I} \setminus \Sigma)$.

Nella prima sezione enunceremo ed dimostreremo alcuni lemmi preliminari; la dimostrazione, che occuperà tutta la seconda sezione, verrà fatta solo per il caso $n = 1$. Per il caso generale, che segue uno schema identico al nostro, si rimanda a [8] e a [9].

Osservazione 2.1. *L'insieme Σ può effettivamente essere non vuoto: questo è il caso, ad esempio, della Lagrangiana $L(t, x, v) := (t^2 - x^3)^2 v^{14} + \varepsilon v^2$ con integrale variazionale associato*

$$\mathcal{L}(x) = \int_0^1 [(t^2 - x^3)^2 \dot{x}^{14} + \varepsilon \dot{x}^2] dt.$$

Si può dimostrare che per $\varepsilon > 0$ sufficientemente piccolo esiste $k \in]0, 1]$ tale che \mathcal{L} ha un minimo assoluto x nella classe $\mathcal{C} = \{y \in AC(I) : y(0) = 0 \text{ e } y(1) = k\}$ per il quale vale $x(t) \sim t^{2/3}$ vicino a 0 e $x \in C^\infty(]0, 1])$: dunque in questo caso $\Sigma = \{0\}$. Per una dimostrazione di questo fatto si veda [1].

Osserviamo infine come il teorema di Tonelli 2.1 sia ottimale, nel senso che dato un qualsiasi chiuso Σ di misura nulla possiamo sempre costruire una lagrangiana di classe C^∞ il cui funzionale associato ha un minimo locale stretto che risulta C^∞ fuori da Σ , con derivata a valori $\pm\infty$ su Σ . Per la dimostrazione di questo fatto si rimanda a [1] e a [4].

2.1 Alcuni lemmi preliminari

Dimostriamo in questa sezione due proposizioni che verranno utilizzate nella dimostrazione del teorema di Tonelli 2.1. Come nelle ipotesi di tale teorema, nel seguito supporremo $L \in C^\infty(\bar{I} \times \mathbb{R} \times \mathbb{R})$ e $L_{vv} > 0$ su $\bar{I} \times \mathbb{R} \times \mathbb{R}$. La prima di tali proposizioni è la seguente

Proposizione 2.1. *Sia $x \in Lip(I, \mathbb{R})$ un estrema debole di \mathcal{L} . Allora $x \in C^\infty(\bar{I}, \mathbb{R})$.*

Dimostrazione. Consideriamo la funzione

$$\begin{aligned} \psi : \bar{I} \times \mathbb{R} \times \mathbb{R} &\longrightarrow \bar{I} \times \mathbb{R} \times \mathbb{R} \\ (t, x, v) &\longmapsto (t, x, L_v(t, x, v)) \end{aligned}$$

Il determinante Jacobiano di ψ è pari a L_{vv} che è positivo per ipotesi, dunque ψ è un diffeomorfismo tra $\bar{I} \times \mathbb{R} \times \mathbb{R}$ e la sua immagine $\mathcal{A} := \psi(\bar{I} \times \mathbb{R} \times \mathbb{R})$. Per la proposizione 1.2 del capitolo 1 esiste $c \in \mathbb{R}$ tale che

$$L_v(t, x(t), \dot{x}(t)) = c + \int_a^t L_x(s, x(s), \dot{x}(s)) ds \quad \text{q.o. } t \in I. \quad (2.1.1)$$

Sia $\gamma(t)$ la quantità al secondo membro della (2.1.1): abbiamo $\gamma \in AC(I, \mathbb{R})$ e

$$\dot{\gamma}(t) = L_x(t, x(t), \dot{x}(t)) \quad \text{q.o. } t \in I$$

e, dato che L_x è continua e \dot{x} è limitata, $\dot{\gamma} \in L^\infty(I, \mathbb{R})$. Poniamo

$$\sigma(t) := (t, x(t), \dot{x}(t))$$

$$\eta(t) := (t, x(t), \gamma(t)).$$

dove σ è definita quasi ovunque su I , mentre η lo è su tutto \bar{I} . La (2.1.1) implica che

$$\psi \circ \sigma = \eta \quad \text{quasi ovunque su } I \quad (2.1.2)$$

da cui $\eta(t) \in \mathcal{A}$ per quasi ogni $t \in I$. Supponiamo invece di aver dimostrato che $\eta(t) \in \mathcal{A}$ per ogni $t \in \bar{I}$. Sia $\alpha : \bar{I} \rightarrow \mathbb{R}$ definita da

$$(t, x(t), \alpha(t)) := \psi^{-1}(\eta(t)).$$

Dal fatto che ψ^{-1} è ben definita e continua su \mathcal{A} e che $\eta(t) \in \mathcal{A}$ per ogni $t \in \bar{I}$ segue che α è ben definita e continua $t \in \bar{I}$. D'altra parte la (2.1.2) dice che

$$\psi^{-1} \circ \eta = \sigma = (\cdot, x(\cdot), \dot{x}(\cdot)) \quad \text{q.o. su } I$$

da cui $\dot{x}(t) = \alpha(t)$ q.o. su I . Poichè x è lipschitziana, essa coincide con l'integrale della propria derivata e quindi

$$x(t) = x(a) + \int_a^t \alpha(s) ds. \quad (2.1.3)$$

Dalla (2.1.3), vista la continuità di α , otteniamo che $x \in \mathcal{C}^1(\bar{I}, \mathbb{R})$, e anche $\gamma \in \mathcal{C}^1(I, \mathbb{R})$. Definiamo

$$\begin{aligned} F : \bar{I} \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (t, v) &\longmapsto L_v(t, x, v) - \gamma(t). \end{aligned}$$

F è di classe \mathcal{C}^1 e $F_v(t, \dot{x}(t)) = L_{vv}(t, \dot{x}(t)) > 0$. Poichè $(t, \dot{x}(t))$ è soluzione di $F(t, v) = 0$, il teorema delle funzioni implicite ci dice che \dot{x} è di classe \mathcal{C}^1 , ma allora γ e quindi F sono \mathcal{C}^2 , da cui anche \dot{x} lo è. Iterando questo argomento otteniamo $\dot{x} \in \mathcal{C}^\infty$: dunque la tesi è dimostrata modulo provare che $\eta(t) \in \mathcal{A}$ per ogni $t \in \bar{I}$.

A tale scopo, sia k la costante di lipschitzianità di x e sia $\Delta = \psi(\bar{I} \times \mathbb{R} \times \{|v| \leq k\})$: esso è un compatto di \mathcal{A} e $\eta(t) \in \Delta$ per q.o. t (dato che $\|\dot{x}\|_\infty \leq k$). Ma η è continua su \bar{I} , dunque $\eta(t) \in \Delta \subset \mathcal{A}$ per ogni $t \in \bar{I}$.

La proposizione è dunque dimostrata nella sua interezza. \square

Lemma 2.1. *Siano $\Omega \subset \mathbb{R}^2$ un aperto limitato e M, δ due costanti positive. Allora esiste $\epsilon > 0$ tale che per ogni $(t_0, x_0) \in \Omega$ e $\alpha, \beta \in [-M, M]$, si ha che per $|t - t_0| < \epsilon$ esiste ed è unica la soluzione $x^{(\alpha, \beta)}(t)$ del problema di Cauchy*

$$\begin{cases} \frac{d}{dt} L_v(t, x(t), \dot{x}(t)) - L_x(t, x(t), \dot{x}(t)) = 0 \\ x(t_0) = x_0 + \alpha \\ \dot{x}(t_0) = \beta. \end{cases} \quad (2.1.4)$$

Inoltre possiamo scegliere ϵ in modo che:

(i) x e \dot{x} sono di classe \mathcal{C}^1 su $S := \{(t, \alpha, \beta) : |t - t_0| < \epsilon, \alpha, \beta \in [-M, M]\}$

(ii) su S abbiamo

$$\begin{aligned} |\dot{x}^{(\alpha, \beta)}(t) - \beta| &< \delta \\ \frac{\partial x^{(\alpha, \beta)}}{\partial \alpha}(t) &> 0 \\ \operatorname{sgn} \frac{\partial x^{(\alpha, \beta)}}{\partial \beta}(t) &= \operatorname{sgn}(t - t_0) \end{aligned} \quad (2.1.5)$$

Dimostrazione. Dal momento che $L_{vv} > 0$ la risoluzione del sistema (2.1.4) è equivalente a quella di:

$$\ddot{x} = \frac{L_{tv} + \dot{x}L_{xv} - L_x}{L_{vv}}$$

con i medesimi dati al bordo. Il teorema segue da risultati standard su ODE e, per quanto riguarda il punto (ii), dalle seguenti osservazioni:

$$\begin{aligned} \frac{\partial x^{(\alpha, \beta)}}{\partial \alpha}(t_0) &= 1 \\ \frac{\partial x^{(\alpha, \beta)}}{\partial \beta}(t_0) &= 0, \quad \frac{\partial \dot{x}^{(\alpha, \beta)}}{\partial \beta}(t_0) = 1. \end{aligned}$$

□

Dimostriamo ora la seconda delle proposizioni che verranno utilizzate nella dimostrazione del teorema di Tonelli: in essa faremo uso della proposizione 2.1 e del lemma 2.1.

Proposizione 2.2. *Siano m, ϱ, N tre costanti positive. Allora esiste $\epsilon > 0$ tale che per ogni $(t_0, t_1) \subset [a, b]$ e ogni x_0, x_1 tali che $0 < t_1 - t_0 < \epsilon$ e $|x_0| \leq m$ e $|\frac{x_1 - x_0}{t_1 - t_0}| < N$, si ha che esiste ed è unica la soluzione $\tilde{x} \in \mathcal{C}^2([t_0, t_1])$ di*

$$\begin{cases} \frac{d}{dt} L_v(t, x(t), \dot{x}(t)) - L_x(t, x(t), \dot{x}(t)) = 0 \\ x(t_0) = x_0 \\ x(t_1) = x_1 \end{cases} \quad (2.1.6)$$

nella classe delle funzioni x tali che $\|x - x_0\|_{L^\infty([t_0, t_1])} < \varrho$.

Inoltre \tilde{x} risulta essere l'unico minimo del funzionale

$$\mathcal{L}^{(t_0, t_1)}(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt$$

nella classe

$$\mathcal{C} := \{x \in AC(t_0, t_1) : x(t_0) = x_0, x(t_1) = x_1, \|x - x_0\|_{L^\infty([t_0, t_1])} < \varrho\}.$$

Dimostrazione. Poniamo $\Omega := (a, b) \times (-m - \varrho, m + \varrho)$ e siano M, δ tali che $M > N \wedge 2\varrho$ e $0 < \delta < M - N$. Scegliamo allora ϵ come nel lemma 2.1; a meno di sceglierlo ancora più piccolo possiamo supporre che $3M\epsilon < \varrho$. Integrando la (2.1.5) del punto (ii) del lemma precedente otteniamo

$$|x^{(\alpha, \beta)}(t) - x_0 - \alpha - \beta(t - t_0)| \leq \delta(t - t_0) \quad (2.1.7)$$

che, assieme all'ipotesi $x_0 - N(t - t_0) \leq x_1 \leq x_0 + N(t - t_0)$, ci dà

$$\begin{aligned} x^{(0, -M)}(t_1) &\leq x_0 - N(t_1 - t_0) - (M - N - \delta)(t_1 - t_0) \\ &\leq x_1 - (M - N - \delta)(t_1 - t_0) \\ &< x_1 \end{aligned} \quad (2.1.8)$$

$$\begin{aligned} x^{(0, M)}(t_1) &\geq x_0 + N(t_1 - t_0) + (M - N - \delta)(t_1 - t_0) \\ &\geq x_1 + (M - N - \delta)(t_1 - t_0) \\ &> x_1 \end{aligned} \quad (2.1.9)$$

Poichè $\operatorname{sgn} \frac{\partial x^{(\alpha, \beta)}}{\partial \beta}(t) = \operatorname{sgn}(t - t_0)$ avremo che

$$\frac{\partial x^{(0, \beta)}}{\partial \beta}(t_1) > 0$$

per ogni $\beta \in [-M, M]$. Ricordando la (2.1.8) e la (2.1.9) deduciamo che esiste un unico β_0 tale che $x^{(0, \beta_0)}(t_1) = x_1$, e poniamo $\tilde{x}(t) := x^{(0, \beta_0)}(t)$. Il lemma 2.1 ci dice inoltre che \tilde{x} è di classe \mathcal{C}^2 .

Sostituendo $t = t_1$, $\alpha = 0$ e $\beta = \beta_0$ nella (2.1.7) otteniamo

$$|\tilde{x}(t_1) - x_0 - \beta_0(t_1 - t_0)| \leq \delta(t_1 - t_0)$$

$$|\beta_0(t_1 - t_0)| \leq \delta(t_1 - t_0) + |x_1 - x_0|$$

ed infine $|\beta_0| \leq \delta + N$. A questo punto riscriviamo la (2.1.7) per $\alpha = 0$ e $\beta = \beta_0$ ottenendo

$$|\tilde{x}(t) - x_0| \leq (\delta + |\beta_0|)(t - t_0) \leq (2\delta + N)\epsilon \leq (2(M - N) + N)\epsilon < \varrho.$$

Supponiamo ora che y sia un'altra soluzione in $\mathcal{C}^2([t_0, t_1])$ dell'equazione di Eulero con dati al bordo $y(t_0) = x_0$ e $y(t_1) = x_1$ e tale che $\|\tilde{y} - x_0\|_{L^\infty([t_0, t_1])} < \varrho$. Allora per il teorema di Lagrange esiste $\bar{t} \in (t_0, t_1)$ tale che $\dot{y}(\bar{t}) = (x_1 - x_0)/(t_1 - t_0) < M$; inoltre $|y(\bar{t})| \leq |x_0| + \varrho < m + \varrho$ e quindi $(\bar{t}, y(\bar{t})) \in \Omega$. Possiamo dunque applicare il lemma 2.1 con $(\bar{t}, y(\bar{t}))$ al posto di (t_0, x_0) e $\dot{y}(\bar{t})$ al posto di β , deducendo dalla parte (ii) che per ogni t

$$\left| \dot{y}(t) - \frac{x_1 - x_0}{t_1 - t_0} \right| \leq \delta,$$

la quale per $t = t_0$ dà $|\dot{y}(t_0)| \leq N + \delta < M$. Ma β_0 è l'unico numero in $[-M, M]$ tale che la soluzione di (2.1.4) per $\alpha = 0$ e $\beta = \beta_0$ vale x_1 al tempo t_1 : poichè anche $\beta := \dot{y}(t_0)$ soddisfa tali condizioni, abbiamo $\beta = \beta_0$ e quindi $y = \tilde{x}$. La prima parte della proposizione è dunque dimostrata.

Per la seconda parte, consideriamo la famiglia a un parametro di soluzioni $\{x^{(\alpha, \beta_0)}\}_{|\alpha| \leq M}$. Ricordando la (2.1.7) e il fatto che $|\beta_0| \leq \delta + N$, abbiamo

$$\begin{aligned} x^{(-M, \beta_0)}(t) - x_0 &\leq -M + (\beta_0 + \delta)(t - t_0) \\ &\leq -M + (2\delta + N)\epsilon \\ &< -\varrho \\ x^{(M, \beta_0)}(t) - x_0 &\geq M + (\beta_0 - \delta)(t - t_0) \\ &\geq M - (2\delta + N)\epsilon \\ &> \varrho. \end{aligned}$$

Poichè $\frac{\partial x^{(\alpha, \beta_0)}}{\partial \alpha}(t) > 0$ segue che \tilde{x} è immerso in un campo di estremali che copre la regione $[t_0, t_1] \times [x_0 - \varrho, x_0 + \varrho]$. Poichè $L_{vv} > 0$, dal teorema 1.6 segue che \tilde{x} è l'unico minimo di $\mathcal{L}^{(t_0, t_1)}$ in \mathcal{C} . \square

2.2 Il teorema di Tonelli

Procediamo ora alla dimostrazione del teorema di regolarità parziale di Tonelli 2.1.

Dimostrazione. Come in precedenza, sia

$$\mathcal{C}(a, b) := \{y \in AC(a, b) : y(a) = x(a) = x_a \text{ e } y(b) = x(b) = x_b\}$$

Per ipotesi esiste $\lambda > 0$ tale che $\mathcal{L}(x) \leq \mathcal{L}(y)$ per ogni $y \in \mathcal{C}(a, b)$ con $\|y - x\|_\infty < \lambda$. Sia \bar{t} un punto in cui i rapporti incrementali di x sono limitati, ovvero

$$N(\bar{t}) := \liminf_{t \rightarrow \bar{t}} \left| \frac{x(t) - x(\bar{t})}{t - \bar{t}} \right| < +\infty.$$

Poichè x è assolutamente continua, essa è derivabile quasi ovunque; in particolare i punti di $[a, b]$ per cui $N < \infty$ sono quasi tutti.

Supponiamo $\bar{t} \neq b$ e sia $\bar{t}_1 > \bar{t}$ tale che $|x(t) - x(\bar{t})| < \lambda/2$ per ogni $t \in [\bar{t}, \bar{t}_1]$, e sia $N > N(\bar{t})$. Possiamo allora applicare la proposizione 2.2 con $t_0 = \bar{t}$, $x_0 = x(\bar{t})$, $\varrho = \lambda/2$ e $x_1 = x(t_1)$ per un certo $t_1 \in (\bar{t}, \bar{t}_1)$ tale che $t_1 - \bar{t} < \epsilon$ (dove ϵ è preso come nella proposizione) e

$$\left| \frac{x(t_1) - x(\bar{t})}{t_1 - \bar{t}} \right| < N.$$

Sia allora \tilde{x} , come nella proposizione, la soluzione dell'equazione di Eulero su $[\bar{t}, t_1]$ con dati al bordo $x_0 = x(\bar{t})$ e $x_1 = x(t_1)$, e sia $\hat{x} \in \mathcal{C}(a, b)$ definita da

$$\hat{x}(t) := \begin{cases} \tilde{x}(t) & \text{se } t \in [\bar{t}, t_1] \\ x(t) & \text{altrimenti.} \end{cases}$$

Allora $\max_{[a, b]} |\hat{x}(t) - x(t)| \leq \lambda$ (infatti \hat{x} e x coincidono fuori da $[\bar{t}, t_1]$, mentre al suo interno $|\hat{x}(t) - x(t)| = |\tilde{x}(t) - x(t)| \leq |\tilde{x}(t) - x(\bar{t})| + |x(\bar{t}) - x(t)| \leq \lambda/2 + \lambda/2$) e quindi, sempre per 2.2

$$\mathcal{L}(\hat{x}) - \mathcal{L}(x) = \mathcal{L}^{(\bar{t}_1, t_1)}(\hat{x}) - \mathcal{L}^{(\bar{t}_1, t_1)}(x) \geq 0$$

ma poichè \tilde{x} è l'unico minimo di $\mathcal{L}^{(\bar{t}_1, t_1)}$ con tali dati al bordo e $\max_{[\bar{t}, t_1]} |\tilde{x}(t) - \tilde{x}(\bar{t})| < \lambda/2$, ne segue che $x = \tilde{x}$ su $[\bar{t}, t_1]$ e quindi $x \in \mathcal{C}^2([\bar{t}, t_1])$.

Un discorso analogo ci dice che x è di classe \mathcal{C}^2 in un intorno sinistro chiuso di \bar{t} : ne segue che x è Lipschitziana in un intorno di \bar{t} per ogni \bar{t} tale che $N(\bar{t}) < +\infty$. La proposizione 2.1 ci dice allora che x è \mathcal{C}^∞ in un intorno di ogni tale \bar{t} . In particolare x è \mathcal{C}^∞ su un aperto di misura piena.

Dobbiamo ancora dimostrare che la derivata $[\dot{x}]$ (intesa come limite a valori in $\overline{\mathbb{R}}$ dei rapporti incrementali) esiste anche nei punti dove N non è finita, ed è una funzione continua (sempre a valori in $\overline{\mathbb{R}}$). Sia dunque \bar{t} tale che $N(\bar{t}) = +\infty$, e supponiamo inizialmente che $\bar{t} \in (a, b)$. Allora esiste una successione $\{t_j\}_{j \in \mathbb{N}}$, $t_j \rightarrow \bar{t}$ tale che si verifica uno dei seguenti quattro casi:

1. $t_j < \bar{t}$ e $\lim_{j \rightarrow +\infty} \frac{x(t_j) - x(\bar{t})}{t_j - \bar{t}} = +\infty$
2. $t_j < \bar{t}$ e $\lim_{j \rightarrow +\infty} \frac{x(t_j) - x(\bar{t})}{t_j - \bar{t}} = -\infty$
3. $t_j > \bar{t}$ e $\lim_{j \rightarrow +\infty} \frac{x(t_j) - x(\bar{t})}{t_j - \bar{t}} = +\infty$
4. $t_j > \bar{t}$ e $\lim_{j \rightarrow +\infty} \frac{x(t_j) - x(\bar{t})}{t_j - \bar{t}} = -\infty$.

Supponiamo ad esempio di essere nel caso 1: gli altri si trattano in maniera del tutto analoga. Per i nostri scopi è sufficiente provare che, date due successioni arbitrarie $\{r_j\}$ e $\{s_j\}$ tali che $r_j \rightarrow \bar{t}$, $s_j \rightarrow \bar{t}$ e $r_j \geq s_j$, allora

$$\lim_{j \rightarrow +\infty} \frac{x(r_j) - x(s_j)}{r_j - s_j} = +\infty.$$

Da questo infatti seguono l'esistenza di $[\dot{x}(\bar{t})] = +\infty$ e la continuità di $[\dot{x}]$ in \bar{t} . Siano M e δ costanti arbitrarie e applichiamo il lemma 2.1 con $x_0 = x(\bar{t})$. La famiglia di soluzioni $\{x^{(\alpha, M)}\}_{|\alpha| \leq M}$ forma un campo di estremali che ricoprono un qualche intorno in \mathbb{R}^2 di $(\bar{t}, x(\bar{t}))$, dunque se t è sufficientemente vicino a \bar{t} esiste un unico $\alpha(t) \in [-M, M]$ tale che $x(t) = x^{(\alpha(t), M)}(t)$. Per il teorema delle funzioni implicite (che possiamo applicare in quanto il lemma ci dice che $\frac{\partial x^{(\alpha, M)}}{\partial \alpha}(t) > 0$), α è una funzione continua di t con $\alpha(\bar{t}) = 0$. Dimostriamo che α è non decrescente in un intorno di \bar{t} . Supponiamo per assurdo che ciò sia falso: esistono allora tre successioni $\{a_j\}$, $\{b_j\}$ e $\{c_j\}$ tutte tendenti a \bar{t} e tali che $a_j < b_j < c_j$ e $\alpha(a_j) = \alpha(c_j) \neq \alpha(b_j)$ (abbiamo usato anche la continuità di α). Per j sufficientemente grande la soluzione dell'equazione di Eulero $y_j := x^{(\alpha(a_j), M)} = x^{(\alpha(c_j), M)}$ soddisfa

$$y_j(a_j) = x(a_j), \quad y_j(b_j) \neq x(b_j), \quad y_j(c_j) = x(c_j)$$

$$\max_{[a_j, c_j]} |x(t) - y_j(t)| \leq \lambda.$$

Poichè y_j è immerso in un campo di integrali il teorema 1.6 ci dice che

$$\int_{a_j}^{c_j} L(t, y_j, \dot{y}_j) dt < \int_{a_j}^{c_j} L(t, x, \dot{x}) dt$$

che contraddice l'ipotesi di minimalità di x su $\{y \in \mathcal{C}(a, b) : \|y - x\|_\infty < \lambda\}$. Dunque α è monotona in un intorno di \bar{t} : dimostriamo che non può essere non crescente. Ricordando la (2.1.7), otteniamo

$$|x^{(\alpha(t_j), M)}(t_j) - \bar{t} - \alpha(t_j) - M(t_j - \bar{t})| \leq \delta(t_j - \bar{t})$$

da cui (tenendo conto che $x^{(\alpha(t_j), M)}(t_j) = x(t_j)$)

$$\frac{\alpha(t_j)}{\bar{t} - t_j} \leq \delta + M - \frac{x(\bar{t}) - x(t_j)}{\bar{t} - t_j} \longrightarrow -\infty$$

ovvero, avendo $t_j \leq \bar{t}$ e $\alpha(\bar{t}) = 0$, α è non decrescente (in un intorno di \bar{t}).

Questo fatto ci consente di concludere: infatti se r_j e s_j sono come sopra, ricordando che $\alpha(r_j) \geq \alpha(s_j)$ e $\frac{\partial x^{(\alpha, M)}}{\partial \alpha}(t) > 0$, otteniamo $x^{(\alpha(r_j), M)} \geq x^{(\alpha(s_j), M)}$ e quindi

$$\begin{aligned} \frac{x(r_j) - x(s_j)}{r_j - s_j} &= \frac{x^{(\alpha(r_j), M)}(r_j) - x^{(\alpha(s_j), M)}(s_j)}{r_j - s_j} \\ &\geq \frac{x^{(\alpha(s_j), M)}(r_j) - x^{(\alpha(s_j), M)}(s_j)}{r_j - s_j} \\ &= \dot{x}^{(\alpha(s_j), M)}(p_j) \quad (*) \\ &\geq M - \delta \end{aligned}$$

per un certo $p_j \in [s_j, r_j]$. Nel passaggio (*) abbiamo usato la parte (ii) del lemma 2.1, e precisamente la (2.1.5). Per l'arbitrarietà di M e δ concludiamo che

$$\lim_{j \rightarrow +\infty} \frac{x(r_j) - x(s_j)}{r_j - s_j} = +\infty$$

come volevamo.

Se $\bar{t} = a$ o $\bar{t} = b$ si procede in maniera del tutto analoga (si possono avere solo due dei quattro casi). Il teorema è dunque dimostrato nella sua interezza. \square

Capitolo 3

Esistenza dei minimi: un teorema di semicontinuità

In questo capitolo dimostreremo un teorema che garantisce, sotto certe condizioni, l'esistenza dei minimi nel Calcolo delle Variazioni. Esso ci servirà nel seguito, quando tratteremo il problema dell'esistenza delle geodetiche nelle varietà sub-riemanniane. Per una trattazione più generale di questi argomenti si veda [3].

Definizione 3.1. Dato un insieme misurabile $\Omega \subset \mathbb{R}^a$ diciamo che una funzione misurabile $f : \Omega \times \mathbb{R}^b \times \mathbb{R}^c \rightarrow]-\infty, +\infty]$ è un **integrando normale-convesso** se esiste un insieme \mathcal{H}^a -trascurabile A tale che

- (i) $f(t, \cdot, \cdot)$ è semicontinua inferiormente per ogni $t \in \Omega \setminus A$
- (ii) $f(t, x, \cdot)$ è una funzione convessa su \mathbb{R}^c per ogni $t \in \Omega \setminus A$ e ogni $x \in \mathbb{R}^b$.

Il risultato che vogliamo presentare nel capitolo è il seguente

Teorema 3.1. Sia $L :]a, b[\times \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, +\infty]$ un integrando normale-convesso. Allora il funzionale

$$\mathcal{L}(x) = \int_a^b L(t, x(t), \dot{x}(t)) dt$$

è sequenzialmente semicontinuo inferiormente rispetto alla topologia debole nello spazio $AC([a, b[, \mathbb{R}^n)$.

Nella prima sezione esporremo alcuni lemmi che sfrutteremo nella seconda, dove daremo la dimostrazione del teorema 3.1.

3.1 Alcuni lemmi preliminari

Per dimostrare il teorema 3.1 abbiamo bisogno di alcuni strumenti. Nel seguito, se non specificato diversamente, $(\Omega, \mathcal{A}, \mu)$ indicherà uno spazio di misura dove \mathcal{A} è μ -completa e μ è priva di atomi.

Iniziamo con i seguenti due lemmi:

Lemma 3.1 (di selezione misurabile). *Sia X uno spazio metrico completo e separabile, e sia $\Gamma : \Omega \rightarrow \mathcal{P}(X)$ una multifunzione. Supponiamo che:*

- (i) $\Gamma(t)$ è chiuso e non vuoto per ogni $t \in \Omega$;
- (ii) il grafico di Γ appartiene a $\mathcal{A} \otimes \mathcal{B}_X$ (\mathcal{B}_X indica la σ -algebra di Borel su X).

Allora esiste una successione di funzione misurabili $z_n : (\Omega, \mathcal{A}) \rightarrow (X, \mathcal{B}_X)$ tale che per ogni $t \in \Omega$ l'insieme $\cup_n \{z_n(t)\}$ è denso in $\Gamma(t)$.

Lemma 3.2. *Sia $f : \mathbb{R}^b \times \mathbb{R}^c \rightarrow]-\infty, +\infty]$ una funzione semicontinua inferiormente con $f(x, \cdot)$ convessa per ogni $x \in \mathbb{R}^b$. Supponiamo che valga almeno una delle seguenti condizioni:*

- (i) esiste una funzione continua $\tilde{v} : \mathbb{R}^b \rightarrow \mathbb{R}^c$ tale che la funzione (definita su \mathbb{R}^b) $x \mapsto f(x, \tilde{v}(x))$ è continua e finita;
- (ii) esiste una funzione $\theta : \mathbb{R}^+ \rightarrow \mathbb{R}$ tale che

$$\lim_{t \rightarrow +\infty} \frac{\theta(t)}{t} \quad e \quad f(x, v) \geq \theta(|v|) \quad \text{per ogni } x, v.$$

Allora esistono due successioni di funzioni continue $a_n : \mathbb{R}^b \rightarrow \mathbb{R}$ e $b_n : \mathbb{R}^b \rightarrow \mathbb{R}^c$ tali che

$$f(x, v) = \sup_n \{a_n(x) + \langle b_n(x), v \rangle\} \quad \text{per ogni } x, v. \quad (3.1.1)$$

Per la dimostrazione del lemma 3.1 si consulti [6]. Per quanto riguarda il lemma 3.2, si veda il lemma 2.2.3 in [3]. Useremo anche il seguente teorema, per cui si rimanda a [6], teorema III.23:

Teorema 3.2 (di proiezione misurabile). *Siano (T, \mathcal{T}) uno spazio misurabile e Σ uno spazio di Suslin (cioè uno spazio di Hausdorff tale che esiste una mappa continua e suriettiva da uno spazio polacco, ovvero metrizzabile, completo e separabile, in Σ). Se $U \subset T \times \Sigma$ appartiene a $\mathcal{T} \otimes \mathcal{B}_\Sigma$ (dove \mathcal{B}_Σ indica la σ -algebra di Borel di Σ), allora la sua proiezione su T appartiene a \mathcal{T} .*

Ad esempio sono di Suslin gli spazi \mathbb{R}^n .

Definizione 3.2. Una funzione $\mathcal{A} \otimes \mathcal{B}_b$ -misurabile (\mathcal{B}_b indica la σ -algebra dei Boreliani su \mathbb{R}^b)

$$f : \Omega \times \mathbb{R}^b \rightarrow]-\infty, +\infty]$$

è un **integrando di Carathéodory** se $f(t, \cdot)$ è una funzione finita e continua su \mathbb{R}^b per μ -quasi ogni $t \in \Omega$.

Proposizione 3.1. Sia $f : \Omega \times \mathbb{R}^b \times \mathbb{R}^c \rightarrow]-\infty, +\infty]$ una funzione misurabile (rispetto alla σ -algebra prodotto $\mathcal{A} \otimes \mathcal{B}_b \otimes \mathcal{B}_c$) che soddisfa le seguenti condizioni:

1. $f(t, \cdot, \cdot)$ è semicontinua inferiormente per ogni $x \in \Omega$;
2. $f(t, x, \cdot)$ è convessa per ogni $(t, x) \in \Omega \times \mathbb{R}^b$

e supponiamo inoltre che per ogni $t \in \Omega$ la funzione $f(t, \cdot, \cdot)$ soddisfi una delle condizioni (i), (ii) del lemma 3.2. Allora esistono due successioni di integrandi di Carathéodory

$$\begin{aligned} a_n & : \Omega \times \mathbb{R}^b \rightarrow \mathbb{R} \\ b_n & : \Omega \times \mathbb{R}^b \rightarrow \mathbb{R}^c \end{aligned}$$

tali che per ogni $(t, x, v) \in \Omega \times \mathbb{R}^b \times \mathbb{R}^c$ si abbia

$$f(t, x, v) = \sup_n \{a_n(t, x) + \langle b_n(t, x), v \rangle\}.$$

Dimostrazione. Sia $\mathcal{C} := \mathcal{C}^0(\mathbb{R}^b, \mathbb{R} \times \mathbb{R}^c)$ munito della usuale topologia della convergenza uniforme sui compatti: esso è completo e separabile.

Data una funzione $u \in \mathcal{C}$ denotiamo con u^1 e u^2 le sue componenti rispettivamente sul primo fattore \mathbb{R} e sul secondo \mathbb{R}^c , e poniamo

$$\begin{aligned} w_u & : \mathbb{R}^b \times \mathbb{R}^c \rightarrow \mathbb{R} \\ (x, v) & \mapsto u^1(x) + \langle u^2(x), v \rangle. \end{aligned}$$

Dato un aperto $A \subset \mathbb{R}^b \times \mathbb{R}^c$, poniamo

$$\begin{aligned} w_A(u) & := \inf_{(x,v) \in A} w_u(x, v) \\ f_A(t) & := \inf_{(x,v) \in A} f(t, x, v) \end{aligned}$$

Dal momento che si tratta di un inf di funzioni continue, w_A risulta semicontinua superiormente su \mathcal{C} . D'altra parte, f_A è misurabile su Ω : infatti, posto $A_r := \{f_A < r\}$ per ogni $r \in \mathbb{R}$, abbiamo

$$\begin{aligned} A_r & = \{t \in \Omega : f(t, x, v) < r \text{ per almeno un } (x, v) \in A\} \\ & = \pi_\Omega(\{(t, x, v) \in \Omega \times A : f(t, x, v) < r\}) \end{aligned}$$

dove $\pi_\Omega : \Omega \times \mathbb{R}^b \times \mathbb{R}^c$ denota la proiezione sul primo fattore Ω . Per il teorema di proiezione misurabile 3.2 abbiamo $A_r \in \mathcal{A}$ per ogni r .

A questo punto definiamo la multifunzione

$$\begin{aligned} \Gamma &: \Omega \rightarrow \mathcal{P}(\mathcal{C}) \\ t &\longmapsto \{u \in \mathcal{C} : f(t, x, v) \geq w_u(x, v) \text{ per ogni } (x, v) \in \mathbb{R}^b \times \mathbb{R}^c\}. \end{aligned}$$

Vogliamo dimostrare che Γ è $\mathcal{A} \otimes \mathcal{B}_{\mathcal{C}}$ -misurabile ($\mathcal{B}_{\mathcal{C}}$ indica la σ -algebra dei boreliani su \mathcal{C}). Data infatti una base numerabile \mathcal{U} degli aperti di $\mathbb{R}^b \times \mathbb{R}^c$, abbiamo

$$\{(t, u) \in \Omega \times \mathcal{C} : u \in \Gamma(t)\} = \bigcap_{A \in \mathcal{U}} \{(t, u) \in \Omega \times \mathcal{C} : f_A(t) \geq w_A(u)\}$$

e quindi il grafico di Γ appartiene a $\mathcal{A} \otimes \mathcal{B}_{\mathcal{C}}$, dato che il secondo membro vi appartiene in quanto intersezione numerabile di misurabili (qui usiamo le proprietà di misurabilità di w_A e di f_A). Per il lemma di selezione misurabile 3.1 esiste una successione di funzioni misurabili $z_n : \Omega \rightarrow \mathcal{C}$ tali che

$$\Gamma(t) = \overline{\{z_n(t) : n \in \mathbb{N}\}} \quad \text{per ogni } t \in \Omega. \quad (3.1.2)$$

Infine definiamo per ogni $n \in \mathbb{N}$, $t \in \Omega$ e $x \in \mathbb{R}^b$ le funzioni

$$\begin{aligned} a_n(t, x) &:= [z_n(t)]^1(x) \\ b_n(t, x) &:= [z_n(t)]^2(x). \end{aligned}$$

Le a_n e b_n sono di Carathéodory; inoltre, in virtù della (3.1.2) e del lemma 3.2, otteniamo, come si voleva,

$$f(t, x, v) = \sup_n \{a_n(t, x) + \langle b_n(t, x), v \rangle\}$$

per ogni $(t, x, v) \in \Omega \times \mathbb{R}^b \times \mathbb{R}^c$. □

Osservazione 3.1. Siano f , a_n e b_n come nel teorema 3.1, con l'ipotesi aggiuntiva $f \geq 0$. Definiamo per $n, m \in \mathbb{N}$, $t \in \Omega$ e $x \in \mathbb{R}^b$ le funzioni

$$\begin{aligned} d_{n,m}(t, x) &:= \frac{m}{m \vee (|a_n(t, x)| + |b_n(t, x)|)} \\ a_{n,m}(t, x) &:= d_{n,m}(t, x)a_n(t, x) \\ b_{n,m}(t, x) &:= d_{n,m}(t, x)b_n(t, x) \end{aligned}$$

É facile verificare la seguente formula di approssimazione per f

$$f(t, x, v) = \sup_{n,m} [a_{n,m}(t, x) + \langle b_{n,m}(t, x), v \rangle]^+$$

che ci consente di approssimare f tramite gli integrandi di Carathéodory limitati $a_{n,m}$ e $b_{n,m}$.

3.2 La dimostrazione del teorema di semicontinuità

Ora abbiamo tutti gli strumenti per dimostrare il teorema 3.1. Ricordando che $W^{1,1}(a,b)$ si immerge con compattezza in $L^p(a,b)$ per ogni $p \in [1, +\infty[$ (teorema di Rellich-Kondrachov: vedi [2]) e che la condizione di compattezza richiesta è automaticamente verificata nel caso di integrando positivo, otteniamo il teorema come corollario del seguente risultato provato da De Giorgi nel caso di integrandi di Carathéodory positivi (vedi [10]) e poi da Ioffe nel caso generale (cfr. [14]).

Teorema 3.3. *Siano $p, q \in [1, +\infty]$ e Ω un sottinsieme misurabile di \mathbb{R}^a di misura finita e sia $f : \Omega \times \mathbb{R}^b \times \mathbb{R}^c \rightarrow]-\infty, +\infty]$ una funzione misurabile che soddisfi le seguenti condizioni:*

- (i) $f(t, \cdot, \cdot)$ è semicontinua inferiormente su $\mathbb{R}^b \times \mathbb{R}^c$ per ogni $t \in \Omega$
- (ii) $f(t, x, \cdot)$ è una funzione convessa su \mathbb{R}^c per ogni $t \in \Omega$ e ogni $x \in \mathbb{R}^b$.

Supponiamo inoltre che valga la seguente condizione di compattezza: per ogni successione $\{x_n\}$ convergente in $L^p(\Omega, \mathbb{R}^b)$ e per ogni successione $\{v_n\}$ debolmente convergente in $L^q(\Omega, \mathbb{R}^c)$ (debolmente-* se $q = +\infty$) tale che

$$\int_{\Omega} f^+(t, x_n(t), v_n(t)) d\mathcal{H}^a(t) \leq c + \int_{\Omega} f^-(t, x_n(t), v_n(t)) d\mathcal{H}^a(t) \text{ per ogni } n$$

per una certa costante $c > 0$, allora la successione $f^-(t, x_n(t), v_n(t))$ è debolmente compatta in $L^1(\Omega)$.

Allora il funzionale

$$F : L^p(\Omega, \mathbb{R}^b) \times L^q(\Omega, \mathbb{R}^c) \rightarrow \mathbb{R}$$

$$(x, v) \mapsto \int_{\Omega} f(t, x(t), v(t)) d\mathcal{H}^a(t)$$

è ben definito a valori in $]-\infty, +\infty]$ ed è sequenzialmente semicontinuo inferiormente rispetto alla topologia prodotto fra quella forte su $L^p(\Omega, \mathbb{R}^b)$ e quella debole su $L^q(\Omega, \mathbb{R}^c)$ (debole-* se $q = +\infty$).

Dimostrazione. Il fatto che F sia ben definito a valori in $]-\infty, +\infty]$ segue dalla condizione di compattezza. Date infatti $(x, v) \in L^p \times L^q$, nel caso esista $c > 0$ tale che

$$\int_{\Omega} f^+(t, x(t), v(t)) d\mathcal{H}^a \leq c + \int_{\Omega} f^-(t, x(t), v(t)) d\mathcal{H}^a \quad (3.2.1)$$

allora $f^-(t, x(t), v(t)) \in L^1(\Omega)$ e quindi $f(t, x(t), v(t)) \in L^1(\Omega)$ per la (3.2.1). Nel caso contrario in cui non esista un tale c devono valere entrambe le seguenti condizioni

$$\int_{\Omega} f^-(t, x(t), v(t)) d\mathcal{H}^a < +\infty$$

$$\int_{\Omega} f^+(t, x(t), v(t)) d\mathcal{H}^a = +\infty$$

da cui

$$\int_{\Omega} f(t, x(t), v(t)) d\mathcal{H}^a = +\infty.$$

Dimostriamo ora il teorema sotto l'ipotesi aggiuntiva che esista una funzione $\theta : \mathbb{R} \rightarrow \mathbb{R}$ tale che

$$\lim_{t \rightarrow +\infty} \frac{\theta(t)}{t} = +\infty$$

$$f(t, x, v) \geq \theta(|v|).$$

La proposizione 3.1 e l'osservazione 3.1 ci dicono allora che esistono due opportune successioni di integrandi di Carathéodory limitati $\{a_n\}$ e $\{b_n\}$ tali che

$$f(t, x, v) = \sup \{ [a_n(t, x) + \langle b_n(t, x), v \rangle]^+ : n \in \mathbb{N} \}.$$

Supponiamo che

$$\int_{\Omega} f(t, x, v) d\mathcal{H}^a = \sup \left\{ \sum_{i \in I} \int_{A_i} [a_i(t, x) + \langle b_i(t, x), v \rangle]^+ d\mathcal{H}^a \right\} \quad (3.2.2)$$

dove il sup è preso fra tutte le partizioni finite di Ω in insiemi boreliani A_i a due a due disgiunti. In tal caso è sufficiente dimostrare che per ogni n naturale e ogni A boreliano il funzionale

$$\int_A a_n(t, x) + \langle b_n(t, x), v \rangle d\mathcal{H}^a$$

è semicontinuo inferiormente rispetto alla topologia forte di $L^p(\Omega, \mathbb{R}^b)$ e debole di $L^q(\Omega, \mathbb{R}^c)$ (debole-* se $q = +\infty$), e questo è banale dal momento che a_n e b_n sono integrali di Carathéodory limitati.

Per dimostrare la (3.2.2), dimostriamo più in generale che, date delle funzioni misurabili g, g_n a valori in $] -\infty, +\infty]$ e tali che $g = \sup g_n$ e $g_n \geq h$ per una certa $h \in L^1(\Omega)$, allora

$$\int_{\Omega} g d\mathcal{H}^a = \sup \left\{ \sum_{i \in I} \int_{A_i} g_i d\mathcal{H}^a \right\} \quad (3.2.3)$$

dove il sup è preso fra tutte le partizioni finite di Ω in insiemi boreliani A_i a due a due disgiunti. In particolare, questo implicherà che per ogni $g \in L^1(\Omega)$

$$\int_{\Omega} g^+ d\mathcal{H}^a = \sup \left\{ \int_A g d\mathcal{H}^a : A \text{ boreliano} \right\}.$$

La disuguaglianza \geq nella (3.2.3) è immediata. Per quanto riguarda quella opposta, poniamo $h_n := \sup_{i=1, \dots, n} g_i$; abbiamo allora $h_n \uparrow g$ e quindi, per il teorema di convergenza monotona, abbiamo

$$\int_{\Omega} g d\mathcal{H}^a = \sup_n \int_{\Omega} h_n d\mathcal{H}^a \quad (3.2.4)$$

Fissato n , definiamo gli insiemi misurabili a due a due disgiunti A_1, \dots, A_n in questo modo:

$$\begin{aligned} A_1 &:= \{h_n = g_1\} \\ A_i &:= \{h_n = g_i\} \setminus \bigcup_{j=1}^{i-1} A_j \end{aligned}$$

In questo modo

$$\int_{\Omega} h_n d\mathcal{H}^a = \sum_{i=1}^n \int_{A_i} g_i d\mathcal{H}^a$$

e quindi, ricordando la (3.2.4), la nostra affermazione è dimostrata.

Tornando al nostro teorema, ci accingiamo a dimostrarlo nell'ipotesi ulteriore che f sia limitata dal basso: per semplicità supporremo f positiva. Siano $x_n \xrightarrow{L^p} x$ e $v_n \xrightarrow{L^q} v$ ($v_n \xrightarrow{*} v$ se $q = +\infty$); per il teorema di compattezza di Dunford-Pettis, che richiamiamo qui sotto (per la sua dimostrazione si veda [11]), esiste una $\theta : \mathbb{R} \rightarrow \mathbb{R}$ tale che

$$\lim_t \frac{\theta(t)}{t} = +\infty \quad e$$

$$\int_{\Omega} \theta(|v_n|) d\mathcal{H}^a \leq 1 \quad \text{per ogni } n.$$

Fissato $\epsilon > 0$ poniamo $f_{\epsilon}(t, x, v) = f(t, x, v) + \epsilon\theta(|v|)$. Abbiamo allora

$$\begin{aligned} F(x, v) = \int_{\Omega} f(t, x, v) d\mathcal{H}^a &\leq \int_{\Omega} f_{\epsilon}(t, x, v) d\mathcal{H}^a \quad (*) \\ &\leq \liminf_{n \rightarrow +\infty} \int_{\Omega} f_{\epsilon}(t, x_n, v_n) d\mathcal{H}^a \\ &\leq \epsilon + \liminf_{n \rightarrow +\infty} \int_{\Omega} f(t, x_n, v_n) d\mathcal{H}^a \\ &= \epsilon + \liminf_{n \rightarrow +\infty} F(x_n, v_n) \end{aligned}$$

dove nel passaggio contrassegnato da (*) abbiamo usato quanto dimostrato in precedenza in quanto, per definizione, si ha $f_\epsilon(t, x, v) \geq \theta(|v|)$.

A questo punto possiamo dimostrare il risultato nel caso generale. Siano ancora $x_n \xrightarrow{L^p} x$ e $v_n \xrightarrow{L^q} v$ ($v_n \xrightarrow{*} v$ se $q = +\infty$); per ogni $h, k \in \mathbb{N}$ poniamo

$$\begin{aligned} f_m &:= \sup \{f, -m\} \\ F_m(u, v) &:= \int_{\Omega} f_m(t, x, v) d\mathcal{H}^a \\ g_n(t) &:= f^-(t, x_n(t), v_n(t)) \\ A_{n,m} &:= \{g_n > m\} \end{aligned}$$

Possiamo supporre che $F(x_n, v_n)$ abbia limite finito per $n \rightarrow +\infty$; in tal caso, per la proprietà di compattezza, la successione $\{g_n\}$ risulta debolmente compatta in $L^1(\Omega)$. Per ogni m abbiamo

$$\begin{aligned} F(x, v) &\leq F_m(x, v) \quad (*) \\ &\leq \liminf_n F_m(x_n, v_n) \\ &= \liminf_n \left[F(x_n, v_n) + \int_{A_{n,m}} (g_n - m) d\mathcal{H}^a \right] \\ &\leq \liminf_n F(x_n, v_n) + \limsup_n \int_{A_{n,m}} g_n d\mathcal{H}^a \end{aligned}$$

dove nel passaggio (*) abbiamo sfruttato la versione del teorema che abbiamo dimostrato, quella con f limitata dal basso. In virtù del teorema di Dunford-Pettis che enunciamo qui sotto, le g_n sono uniformemente integrabili, e il teorema segue prendendo $m \rightarrow +\infty$. \square

Teorema 3.4 (Dunford-Pettis). *Sia $(\Omega, \mathcal{A}, \mu)$ uno spazio di misura con μ positiva e finita, e sia \mathcal{C} un sottinsieme di $L^1(\Omega, \mathcal{A}, \mu)$. Allora le seguenti condizioni sono equivalenti:*

- (i) \mathcal{C} è relativamente compatto nella topologia debole di $L^1(\Omega, \mathcal{A}, \mu)$
- (ii) \mathcal{C} è uniformemente integrabile, ovvero \mathcal{C} è limitato in $L^1(\Omega, \mathcal{A}, \mu)$ e per ogni $\epsilon > 0$ esiste $\delta > 0$ tale che

$$\mu(A) < \delta \Rightarrow \int_A |u| d\mu < \epsilon \text{ per ogni } u \in \mathcal{C}$$

(iii) esiste una funzione $\theta : [0, +\infty[\rightarrow [0, +\infty[$, che può essere scelta convessa e crescente, tale che

$$\lim_{t \rightarrow +\infty} \frac{\theta(t)}{t} = +\infty \quad e \quad \sup_{u \in C} \int_{\Omega} \theta(|u|) d\mu < +\infty.$$

Per la sua dimostrazione si veda [11], pagine 24-27.

Capitolo 4

Spazi di Carnot-Carathéodory

In questo capitolo cominciamo lo studio degli spazi di Carnot-Carathéodory: inizieremo fissando le direzioni lungo cui potremo muoverci, che chiameremo “orizzontali”. In questo modo le curve ammissibili saranno quelle la cui velocità è quasi ovunque orizzontale: definiremo diverse nozioni di lunghezza di una curva, essenzialmente nello stesso modo in cui definiamo le diverse norme L^p di una funzione. La distanza di Carnot-Carathéodory fra due punti risulterà definita come l’inf delle lunghezze delle curve che li connettono: vedremo che tale quantità risulterà indipendente dalla particolare nozione di lunghezza presa in considerazione. Infine, nella seconda sezione, dimostreremo il teorema di Chow, il quale garantisce che esistono sempre curve orizzontali fra due punti quando i campi che generano le direzioni orizzontali sono di classe C^∞ e tali che, in ogni punto, l’algebra di Lie da essi generata ricopre l’intero spazio tangente.

4.1 La distanza di Carnot-Carathéodory

Consideriamo un aperto connesso Ω di \mathbb{R}^n ed una famiglia $\{X_1, \dots, X_m\}$ di campi di vettori localmente Lipschitziani; questo significa che, indicando con ∂_{x_i} , $i = 1, \dots, n$, i vettori della base canonica di \mathbb{R}^n , e se le funzioni $a_{ij} : \Omega \rightarrow \mathbb{R}$ sono tali che

$$X_j(x) = \begin{pmatrix} a_{1j}(x) \\ \vdots \\ a_{nj}(x) \end{pmatrix} = \sum_{i=1}^n a_{ij}(x) \partial_{x_i}$$

allora $a_{ij} \in Lip_{loc}(\Omega)$ per ogni $i \in \{1, \dots, n\}$ e ogni $j \in \{1, \dots, m\}$. Il sottospazio vettoriale generato dai vettori $X_1(x), \dots, X_m(x)$ (che naturalmente può essere visto come un sottospazio dello spazio tangente ad Ω nel punto x) si chiama **sottospazio**

orizzontale nel punto x , e verrà indicato con $H_x\Omega$; con l'espressione **direzione orizzontale** indicheremo invece un qualsiasi vettore di tale sottospazio.

Sia $\mathcal{M}_{n,m}$ lo spazio vettoriale delle matrici a n righe e m colonne a valori reali; definiamo la funzione a valori matrici

$$\mathcal{A} : \Omega \rightarrow \mathcal{M}_{n,m}$$

$$x \mapsto \text{col} [X_1(x), \dots, X_m(x)] = \begin{pmatrix} a_{11}(x) & \cdots & a_{1m}(x) \\ \vdots & \ddots & \vdots \\ a_{n1}(x) & \cdots & a_{nm}(x) \end{pmatrix}$$

Definizione 4.1. Una curva $\gamma : [0, T] \rightarrow \Omega$ si dice **ammissibile** (o **orizzontale**) se è assolutamente continua ed esiste una funzione misurabile $h : [0, T] \rightarrow \mathbb{R}^m$ tale che

$$\dot{\gamma}(t) = \mathcal{A}(\gamma(t))h(t) = \sum_{j=1}^m h_j(t) X_j(\gamma(t)) \text{ per q.o. } t \in [0, T]$$

Diremo poi che γ è **subunitaria** se è ammissibile e $\|h\|_{L^\infty} \leq 1$.

Definizione 4.2. Definiamo la **distanza di Carnot-Carathéodory** d_{CC} su Ω come

$$d_{CC}(x, y) := \inf \{ T : \text{esiste } \gamma : [0, T] \rightarrow \Omega \text{ subunitaria tale che} \\ \gamma(0) = x \text{ e } \gamma(T) = y \}$$

con la convenzione che $\inf \emptyset = +\infty$ (ovvero diciamo che la distanza fra due punti che non sono connessi con cammini orizzontali è infinita).

Lo spazio (Ω, d_{CC}) si chiama **spazio di Carnot-Carathéodory**.

Dimostreremo fra breve che si tratta di uno spazio metrico, ovvero che d_{CC} è una distanza. Per distinguerla da quella Euclidea, indicheremo nel seguito quest'ultima con la notazione d_E ; coerentemente useremo il pedice E quando parleremo di oggetti in riferimento alla metrica Euclidea, quali il diametro $diam_E$, le palle $B_E(x, r)$, etc.

Osservazione 4.1. Più in generale si possono dare le stesse definizioni a partire da una varietà M : in questo caso si partirà dal dato, per ogni $p \in M$, di una distribuzione $H_p M$ di sottospazi di $T_p M$ che chiameremo **distribuzione orizzontale** e che sarà generata localmente da una famiglia di campi di vettori Lipschitziani. Le curve orizzontali saranno quindi quelle tali che $\dot{\gamma}(t) \in T_{\gamma(t)} M$ per quasi ogni t .

La scelta di partire da un aperto di \mathbb{R}^n è dettata da ragioni di maggior semplicità e dalla considerazione che, in ogni caso, non si perde in generalità, essendo i nostri discorsi a livello locale.

Proposizione 4.1. Se $d_{CC}(x, y) < +\infty$ per ogni $x, y \in \Omega$, allora (Ω, d) è uno spazio metrico.

Dimostrazione. Dobbiamo dimostrare che d_{CC} è una distanza; infatti:

- $d_{CC}(x, x) = 0$
- la proprietà di simmetria $d_{CC}(x, y) = d_{CC}(y, x)$ è ovvia
- la disuguaglianza triangolare $d_{CC}(x, y) \leq d_{CC}(x, z) + d_{CC}(z, y)$ è anch'essa una proprietà ovvia.

Resta quindi da dimostrare che se $x \neq y$ allora $d_{CC}(x, y) > 0$: questo segue immediatamente se dimostriamo, ad esempio, che se $K \Subset \Omega$ allora esiste $C = C(K) > 0$ tale che $d_{CC}(x, y) > C|x - y|$ per ogni $x, y \in K$.

Siano dunque fissati K e $x, y \in K$, e sia $\epsilon > 0$ sufficientemente piccolo in modo che si abbia

$$K_\epsilon := \{z \in \Omega : d_E(z, K) \leq \epsilon\} \Subset \Omega.$$

Poniamo $M := \sup_{K_\epsilon} \|\mathcal{A}\|$ e sia $r := \min\{\epsilon, |x - y|\}$. Consideriamo una qualsiasi curva subunitaria $\gamma : [0, T] \rightarrow \Omega$ tale che $\gamma(0) = x$ e $\gamma(T) = y$; dimostriamo che $MT \geq r$. Poniamo

$$\bar{t} := \inf \{t \in [0, T] : d_E(\gamma(t), x) \geq r\}.$$

Osserviamo che, essendo $d_E(x, y) \geq r$, si ha $\bar{t} \leq T$; inoltre $d_E(\gamma(\bar{t}), x) = r$ per continuità. Abbiamo allora

$$\begin{aligned} r &= |\gamma(\bar{t}) - \gamma(0)| = \left| \int_0^{\bar{t}} \dot{\gamma}(s) ds \right| \\ &= \left| \int_0^{\bar{t}} \mathcal{A}(\gamma(s))h(s) ds \right| \\ &\leq \int_0^{\bar{t}} \|\mathcal{A}(\gamma(s))\| |h(s)| ds \\ &\leq \bar{t}M \leq TM. \end{aligned}$$

Dunque $T \geq r/M$ e quindi

$$\begin{aligned} r = \epsilon &\quad \Rightarrow \quad T \geq \frac{\epsilon}{M} \geq \frac{\epsilon}{M \operatorname{diam}_E(K)} |x - y| \\ r = |x - y| &\quad \Rightarrow \quad T \geq \frac{1}{M} |x - y| \end{aligned}$$

e quindi è sufficiente prendere $C(K) = \min \left\{ \frac{\epsilon}{M \operatorname{diam}_E(K)}, \frac{1}{M} \right\}$. □

Nel dimostrare la proposizione 4.1 abbiamo implicitamente dimostrato che la metrica Euclidea è continua rispetto a quella di Carnot-Carathéodory d_{CC} . Non è vero in generale il viceversa, come mostra il seguente esempio.

Esempio 4.1. Consideriamo i seguenti campi di vettori in \mathbb{R}^2 :

$$X_1 = \partial_x$$

$$X_2 = \max(0, x) \cdot \partial_y.$$

È chiaro che $d_{CC}(A, B) < +\infty$ per ogni $A, B \in \mathbb{R}^2$: infatti basta considerare due curve che uniscono A e B al semipiano destro (queste due curve esistono sempre, basta considerare quelle parallele all'asse delle ascisse, che sono sempre ammissibili) e unire queste due curve rimanendo all'interno del semipiano destro. Inoltre è facile vedere che, se $x < 0$, allora la distanza di Carnot-Carathéodory fra il punto (x, y) e il semipiano destro è maggiore o uguale a $|x|$; poichè per connettere due punti del semipiano sinistro con diversa la seconda coordinata dobbiamo necessariamente passare per il semipiano destro, ne segue che

$$\lim_{\substack{y \rightarrow 0 \\ y \neq 0}} d_{CC}((x, y), (x, 0)) \geq 2|x| > 0$$

e quindi, in questo caso, la distanza di Carnot-Carathéodory non è continua rispetto a quella Euclidea.

4.1.1 Una definizione equivalente della distanza d_{CC}

Sia $x \in \Omega$ e $v \in \mathbb{R}^n$; definiamo la “norma orizzontale” di v come

$$|v|_H := \min \{ |h| : h \in \mathbb{R}^m, \mathcal{A}(x)h = v \}$$

con la convenzione $\min \emptyset = +\infty$: in questo modo un vettore non orizzontale risulta avere norma orizzontale pari a $+\infty$. Osserviamo inoltre che si tratta di un vero minimo, in quanto l'insieme $\mathcal{A}(x)^{-1}(v)$ è un sottospazio affine (eventualmente vuoto) di \mathbb{R}^m : se è non vuoto questo implica che il minimo esiste ed è unico. Essendo tale sottospazio affine parallelo al sottospazio (vettoriale) $\ker \mathcal{A}(x)$ deduciamo che il minimo è raggiunto nel punto $h - P_{\ker \mathcal{A}(x)}(h)$ per ogni $h \in \mathcal{A}(x)^{-1}(v)$, dove per ogni sottospazio vettoriale V indichiamo con P_V la proiezione su V .

La norma orizzontale ha la proprietà di essere convessa: siano infatti dati $v_1, v_2 \in H_x \Omega$ e $\lambda \in [0, 1]$, e siano h_1, h_2 gli unici vettori di \mathbb{R}^m tali che

$$\mathcal{A}(x)h_j = v_j \quad \text{e}$$

$$|v_j|_H = |h_j|$$

per $j = 1, 2$. Ne segue che $\mathcal{A}(x)(\lambda h_1 + (1 - \lambda)h_2) = \lambda v_1 + (1 - \lambda)v_2$ da cui

$$\begin{aligned} |\lambda v_1 + (1 - \lambda)v_2|_H &\leq |\lambda h_1 + (1 - \lambda)h_2| \\ &\leq \lambda|h_1| + (1 - \lambda)|h_2| \\ &= \lambda|v_1|_H + (1 - \lambda)|v_2|_H \end{aligned} \quad (4.1.1)$$

e quindi la norma orizzontale è convessa se v_1 e v_2 sono orizzontali; d'altra parte se uno dei due non lo è avrà norma orizzontale $+\infty$ e quindi la (4.1.1) risulta banalmente verificata.

Definizione 4.3. *Data una curva $\gamma : [0, T] \rightarrow \Omega$ definiamo la sua **lunghezza orizzontale** come la quantità*

$$l_H(\gamma) := \int_0^T |\dot{\gamma}(t)|_H dt. \quad (4.1.2)$$

Ovviamente una curva con lunghezza orizzontale finita risulta ammissibile.

Notiamo poi che la definizione è ben posta, in quanto l'integrando che compare nella (4.1.2) è una funzione positiva e misurabile. La misurabilità si può ottenere in questo modo: sappiamo che esiste una funzione misurabile $h : [0, T] \rightarrow \mathbb{R}^m$ tale che $\dot{\gamma}(t) = \mathcal{A}(\gamma(t))h(t)$; allora

$$|\dot{\gamma}(t)|_H = |h(t) - P_{\ker \mathcal{A}(\gamma(t))}h(t)|$$

che risulta misurabile grazie al seguente

Lemma 4.1. *Siano $\mathcal{A} : [0, 1] \rightarrow L(\mathbb{R}^m, \mathbb{R}^n) \simeq \mathcal{M}_{n,m}$ misurabile e $h : [0, 1] \rightarrow \mathbb{R}^m$ misurabile. Allora la funzione*

$$\mathbb{R} \ni t \longmapsto h(t) - P_{\ker \mathcal{A}(t)}h(t) \in \mathbb{R}^m \quad (4.1.3)$$

risulta misurabile.

Dimostrazione. Innanzitutto osserviamo che il secondo membro della (4.1.3) è uguale a $P_{\ker \mathcal{A}(t)^\perp}h(t)$; inoltre si ha che $\ker \mathcal{A}(t)^\perp = \text{Im } \mathcal{A}(t)^T$, e quindi tale sottospazio è generato dai vettori che si ottengono considerando le righe della matrice \mathcal{A} , ovvero

$$V_j(t) = \begin{pmatrix} \mathcal{A}(t)_{j1} \\ \vdots \\ \mathcal{A}(t)_{jm} \end{pmatrix} \quad j = 1, \dots, n$$

che sono evidentemente funzioni misurabili. Definita la funzione $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^m$ data da

$$\varphi(v) = \begin{cases} \frac{v}{|v|} & \text{se } v \neq 0 \\ 0 & \text{se } v = 0 \end{cases}$$

utilizziamo il procedimento di Gram-Schmidt per ottenere un sistema di generatori $W_1(t), \dots, W_n(t)$ di $\ker \mathcal{A}(t)^\perp$ tale che i suoi vettori non nulli siano ortonormali:

$$\begin{aligned} W_1(t) &:= \varphi(V_1(t)) \\ W_j(t) &:= \varphi \left(V_j(t) - \sum_{i=1}^{j-1} \langle V_j(t), W_i(t) \rangle W_i(t) \right). \end{aligned}$$

Dimostriamo che le funzioni $W_j : [0, 1] \rightarrow \mathbb{R}^m$ sono misurabili. É sufficiente provare che, per ogni funzione ψ misurabile e a valori in \mathbb{R}^m , la funzione $\varphi \circ \psi$ è a sua volta misurabile: a quel punto W_1 sarà misurabile e concluderemo per induzione.

Sia A un aperto di \mathbb{R}^m : dobbiamo provare che $(\varphi \circ \psi)^{-1}(A)$ è un insieme misurabile. Abbiamo però

$$(\varphi \circ \psi)^{-1}(A) = \psi^{-1}(A \cap \{0\}) \cup (\varphi \circ \psi)^{-1}(A \setminus \{0\})$$

che è misurabile perchè unione di due insiemi misurabili: il primo lo è in maniera ovvia, il secondo perchè $\varphi \in \mathcal{C}^\infty(\mathbb{R}^m \setminus \{0\}, \mathbb{R}^m \setminus \{0\})$, e sappiamo che la composizione di una funzione continua con una misurabile (in quest'ordine!) è misurabile. A questo punto abbiamo, grazie all'ortogonalità di quei W_j che non sono nulli,

$$\begin{aligned} P_{\ker \mathcal{A}(t)^\perp} h(t) &= P_{\text{span}\{W_1(t), \dots, W_n(t)\}} h(t) \\ &= P_{\text{span}\{W_1(t)\}} h(t) + \dots + P_{\text{span}\{W_n(t)\}} h(t) \\ &= \langle W_1(t), h(t) \rangle W_1(t) + \dots + \langle W_n(t), h(t) \rangle W_n(t) \end{aligned}$$

che è misurabile. Il lemma risulta dunque dimostrato. \square

Osservazione 4.2. *É facile verificare che la lunghezza orizzontale è invariante per riparametrizzazione della curva γ .*

Siamo dunque dati $x, y \in \Omega$; poniamo

$$d_H(x, y) := \inf \{ l_H(\gamma) : \gamma \in AC([0, 1], \Omega), \gamma(0) = x \text{ e } \gamma(1) = y \}. \quad (4.1.4)$$

Osserviamo che una definizione equivalente di d_H si ottiene considerando, nella (4.1.4), l'inf sulle sole funzioni Lipschitziane: infatti una curva assolutamente continua si può sempre riparametrizzare in modo che sia Lipschitziana, ad esempio prima parametrizzandola per lunghezza (euclidea!) d'arco e poi riportandola su $[0, 1]$. In questa osservazione è fondamentale l'invarianza di l_H per riparametrizzazioni. Il nostro obiettivo è la seguente

Proposizione 4.2. Per ogni $x, y \in \Omega$ si ha

$$d_{CC}(x, y) = d_H(x, y).$$

Dimostrazione. Innanzitutto osserviamo che

$$d_{CC}(x, y) = \inf \{ \|h\|_{L^\infty} \quad : \quad \text{esiste } \gamma \in \text{Lip}([0, 1], \Omega) \text{ orizzontale tale che} \\ \gamma(0) = x \text{ e } \gamma(1) = y \}. \quad (4.1.5)$$

Data infatti una curva $\gamma : [0, T] \rightarrow \Omega$ subunitaria con estremi x e y definiamo $\tilde{\gamma}(t) := \gamma(Tt)$; se \tilde{h} è la funzione a valori in \mathbb{R}^m tale che $\dot{\tilde{\gamma}} = \mathcal{A}(\tilde{\gamma})\tilde{h}$ abbiamo $\|\tilde{h}\|_{L^\infty} = T\|h\|_{L^\infty} \leq T$, da cui si ricava che il secondo membro della (4.1.5) è minore o uguale a $d_{CC}(x, y)$.

La disuguaglianza opposta si ottiene in maniera analoga: data $\tilde{\gamma} \in \text{Lip}([0, 1], \Omega)$ con estremi x e y definiamo

$$\gamma(t) := \tilde{\gamma} \left(\frac{t}{\|\tilde{h}\|_{L^\infty}} \right).$$

Abbiamo che γ va da x a y , è subunitaria ed è definita su $[0, \|\tilde{h}\|_{L^\infty}]$; ne segue che $d_{CC}(x, y)$ è minore o uguale al secondo membro della (4.1.5).

A questo punto è facile dedurre la disuguaglianza $d_H(x, y) \leq d_{CC}(x, y)$: data infatti $\gamma \in \text{Lip}([0, 1], \Omega)$ orizzontale abbiamo

$$\begin{aligned} l_H(\gamma) &= \int_0^T |\dot{\gamma}(t)|_H dt \\ &\leq \int_0^T |h| dt \\ &= \|h\|_{L^1} \leq \|h\|_{L^\infty} \end{aligned}$$

che ci dà proprio tale disuguaglianza.

Resta quindi da dimostrare la disuguaglianza $d_{CC}(x, y) \leq d_H(x, y)$; a tale scopo sia data una $\tilde{\gamma} : [0, 1] \rightarrow \Omega$ ammissibile che va da x a y , e sia \tilde{h} tale che $\dot{\tilde{\gamma}} = \mathcal{A}(\tilde{\gamma})\tilde{h}$ e $\|\tilde{h}\|_{L^1} < +\infty$. Inoltre possiamo supporre che $|\tilde{h}| = |\dot{\tilde{\gamma}}|_H$: al più basterà sostituire \tilde{h} con $\tilde{h} - P_{\ker \mathcal{A}(\tilde{\gamma})}\tilde{h}$. Per i nostri scopi è sufficiente trovare una $\gamma : [0, 1] \rightarrow \Omega$ ammissibile che va da x a y e tale che se $\dot{\gamma} = \mathcal{A}(\gamma)h$ allora $\|h\|_{L^\infty} \leq \|\tilde{h}\|_{L^1} = l_H(\gamma)$.

Consideriamo dunque la funzione assolutamente continua e non decrescente $\varphi : [0, 1] \rightarrow [0, 1]$ data da

$$\varphi(t) = \frac{1}{\|\tilde{h}\|_{L^1}} \int_0^t |\tilde{h}(s)| ds.$$

Sia poi $\psi : [0, 1] \rightarrow [0, 1]$ la pseudo-inversa di ϕ , ovvero la funzione data da $\psi(s) := \inf \{t : \phi(t) = s\}$. Posto $S := \{t : \phi \text{ non è differenziabile in } \psi(t)\}$ si ha $S \subset \phi(\{t : \phi \text{ non è differenziabile in } t\})$; essendo S contenuto nell'immagine secondo una funzione assolutamente continua di un insieme di misura nulla, avrà anch'esso misura nulla e quindi $\phi \circ \psi$ è differenziabile quasi ovunque. Tuttavia tale funzione coincide per costruzione con l'identità; ne deduciamo che

$$\dot{\psi}(s) = \frac{1}{\dot{\phi}(\psi(s))} = \frac{\|\tilde{h}\|_{L^1}}{|h(\psi(s))|} \quad \text{q.o. } s.$$

Definiamo ora $\gamma : [0, 1] \rightarrow \Omega$ tramite $\gamma(s) := \tilde{\gamma}(\psi(s))$: l'assoluta continuità di $\tilde{\gamma}$ ci dice che γ è differenziabile quasi ovunque e quindi

$$\begin{aligned} \dot{\gamma}(s) &= \dot{\tilde{\gamma}}(\psi(s))\dot{\psi}(s) \\ &= \mathcal{A}(\tilde{\gamma}(\psi(s)))\tilde{h}(\psi(s))\dot{\psi}(s) \\ &= \mathcal{A}(\gamma(s))\|\tilde{h}\|_{L^1}\frac{\tilde{h}(\psi(s))}{|\tilde{h}(\psi(s))|} \quad \text{q.o. } s. \end{aligned}$$

Infine per $j = 1, \dots, m$ definiamo

$$h_j(s) := \begin{cases} \|\tilde{h}\|_{L^1}\frac{\tilde{h}_j(\psi(s))}{|\tilde{h}(\psi(s))|} & \text{se } \tilde{h}(\psi(s)) \neq 0 \\ 0 & \text{se } \tilde{h}(\psi(s)) = 0. \end{cases}$$

Si ha evidentemente che γ va da x a y e $\dot{\gamma}(s) = \mathcal{A}(\gamma(s))h(s)$ quasi ovunque, ed è immediato verificare che vale la disuguaglianza che cercavamo, $\|h\|_{L^\infty} \leq \|\tilde{h}\|_{L^1}$. \square

Osservazione 4.3. Osserviamo che la dimostrazione della proposizione 4.2 fornisce anche un procedimento per parametrizzare una curva orizzontale in modo che la sua velocità orizzontale sia costante. Questo può tornare utile in molti casi.

Osservazione 4.4. Dato $p \in [1, \infty]$ definiamo per ogni curva $\gamma : [0, 1] \rightarrow \Omega$ la **lunghezza orizzontale p -esima** $l_p(\gamma)$ data da

$$l_p(\gamma) := \|\dot{\gamma}|_H\|_{L^p(0,1)}.$$

Definiamo infine la **distanza orizzontale p -esima** $d_p(x, y)$ fra due punti $x, y \in \Omega$ come

$$d_p(x, y) := \inf \{l_p(\gamma) : \gamma \in AC([0, 1]), \gamma(0) = x \text{ e } \gamma(1) = y\}.$$

Ripercorrendo la dimostrazione della proposizione 4.2 osserviamo che

$$d_1(x, y) = d_H(x, y) = d_{CC}(x, y) = d_\infty(x, y)$$

che, assieme alla disuguaglianza di Holder

$$l_p(\gamma) \leq l_q(\gamma) \quad \text{per ogni } \gamma \text{ e ogni } p < q,$$

ci dice che $d_p = d_{CC}$ per ogni $p \in [1, +\infty]$. Useremo d_2 nel capitolo successivo, quando parleremo di geodetiche normali e anormali.

4.2 Il teorema di Chow

Il problema che ci poniamo in questa sezione è quello di dare condizioni sufficienti affinché uno spazio di Carnot-Carathéodory risulti connesso, ovvero affinché due suoi punti qualsiasi siano connettabili tramite una curva orizzontale. Nel caso di campi vettoriali X_1, \dots, X_m di classe C^∞ si ha un noto risultato dimostrato da Chow [7] nel 1939; ci accingiamo dunque a dimostrarlo seguendo l'approccio di [15].

Sia Y un campo vettoriale C^∞ su Ω ; per $x \in \Omega$ consideriamo il seguente problema di Cauchy

$$\begin{cases} \dot{\gamma}_x(t) = Y(\gamma_x(t)) \\ \gamma_x(0) = x. \end{cases} \quad (4.2.1)$$

Per risultati standard sulle equazioni differenziali ordinarie, γ_x è definita per $t \in]-\delta, \delta[$ per un certo $\delta = \delta(x) > 0$; inoltre se K è un compatto contenuto in Ω possiamo scegliere δ in maniera indipendente da $x \in K$. Per $|t| < \delta$ e $x \in K$ è quindi ben definita la **mappa esponenziale**

$$\exp(tY)(x) = e^{tY}(x) := \gamma_x(t)$$

dove γ_x è la soluzione di 4.2.1. Indicheremo con $\exp(Y)(x)$ il punto $\gamma_x(1)$: ovviamente il campo Y dovrà essere sufficientemente piccolo in modo che $\gamma_x(1)$ risulti definito.

È ben noto che la funzione $t \mapsto e^{tY}(x)$ è di classe C^∞ , e lo sviluppo in serie di Taylor dà

$$e^{tY}(x) = x + tY(x) + O(t^2).$$

Ricordiamo che se $X = \sum_{i=1}^n a_i(x)\partial_{x_i}$ e $Y = \sum_{i=1}^n b_i(x)\partial_{x_i}$ sono due campi vettoriali di classe C^∞ è definito il loro **commutatore** (o **parentesi di Lie**) $[X, Y]$, che è il campo vettoriale C^∞ dato da

$$[X, Y] = XY - YX$$

ovvero, in coordinate,

$$[X, Y] = \sum_{i=1}^n \sum_{j=1}^n \left(a_j(x) \frac{\partial}{\partial x_j} b_i(x) - b_j(x) \frac{\partial}{\partial x_j} a_i(x) \right) \partial_{x_i}.$$

Tale prodotto è antisimmetrico ($[X, Y] = -[Y, X]$) e soddisfa la relazione di Jacobi

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Seguendo il formalismo delle algebre di Lie definiamo l'applicazione aggiunta ad X , quella cioè che ad un campo Y di classe C^∞ associa il campo $\text{ad}X(Y) := [X, Y]$.

Con l'applicazione iterata delle parentesi di Lie i campi X_1, \dots, X_m generano un'algebra di Lie (per la cui definizione si veda il capitolo 6), che indicheremo con $\mathfrak{L}(X_1, \dots, X_m)$. Per ogni $x \in \Omega$ quest'algebra di Lie ci dà un sottospazio vettoriale $\mathfrak{L}(X_1, \dots, X_m)(x)$ dello spazio tangente.

Definizione 4.4. Diciamo che i campi vettoriali X_1, \dots, X_m soddisfano la **condizione di Chow-Hörmander** se per ogni $x \in \Omega$ è massimo il rango del sottospazio $\mathfrak{L}(X_1, \dots, X_m)(x)$, ovvero

$$\text{rank } \mathfrak{L}(X_1, \dots, X_m)(x) = n \text{ per ogni } x \in \Omega.$$

Se i campi X_1, \dots, X_m soddisfano la condizione di Chow-Hörmander, allora (Ω, d_{CC}) si chiama **spazio sub-Riemanniano**; se lavoriamo su varietà invece che su aperti di \mathbb{R}^n parleremo di **varietà sub-Riemanniane**.

Il teorema di Chow dice che se X_1, \dots, X_m soddisfano la condizione di Chow-Hörmander, allora due punti qualsiasi di Ω sono connettabili tramite una curva orizzontale. Prima di dimostrarlo abbiamo però bisogno di alcuni altri strumenti, che andiamo ad esporre brevemente.

Se $\alpha = (\alpha_1, \dots, \alpha_k)$ è un multi-indice di interi non negativi definiamo

$$\begin{aligned} |\alpha| &:= \alpha_1 + \dots + \alpha_k \\ \alpha! &:= \alpha_1! \dots \alpha_k! \end{aligned}$$

e diremo che k è la lunghezza di α . Se $\beta = (\beta_1, \dots, \beta_k)$ è un altro multi-indice di lunghezza k tale che $\alpha_k + \beta_k \geq 1$ e X, Y sono due campi C^∞ definiamo

$$C_{\alpha\beta}(X, Y) := \begin{cases} (\text{ad}X)^{\alpha_1}(\text{ad}Y)^{\beta_1} \dots (\text{ad}X)^{\alpha_k}(\text{ad}Y)^{\alpha_k-1}Y & \text{se } \beta_k \neq 0 \\ (\text{ad}X)^{\alpha_1}(\text{ad}Y)^{\beta_1} \dots (\text{ad}X)^{\alpha_k-1}X & \text{se } \beta_k = 0. \end{cases}$$

Notiamo che se $k = 1$ e $\alpha = \beta = (1)$ allora $C_{\alpha\beta}(X, Y)$ coincide con il commutatore $[X, Y]$; qui come nel seguito intendiamo che $(\text{ad}X)^0$ è l'identità.

La composizione di due esponenziali si può ricondurre ad un opportuno esponenziale tramite il seguente

Teorema 4.1 (formula di Baker-Campbell-Hausdorff). *Dati due campi vettoriali X, Y sia $C(X, Y)$ il campo tale che $\exp(X)\exp(Y) = \exp(C(X, Y))$. Indicando con $\alpha + \beta$ il multi-indice $(\alpha_1 + \beta_1, \dots, \alpha_k + \beta_k)$, allora si ha che*

$$C(X, Y) := \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sum_{\substack{\alpha=(\alpha_1, \dots, \alpha_k) \\ \beta=(\beta_1, \dots, \beta_k) \\ \alpha_j + \beta_j \geq 1 \forall j}} \frac{1}{\alpha! \beta! |\alpha + \beta|} C_{\alpha\beta}(X, Y) \quad (4.2.2)$$

quando la serie al secondo membro ha senso.

Osservazione 4.5. *Nel contesto dei gruppi di Lie la sommatoria del teorema precedente converge uniformemente se i campi X, Y in questione sono invarianti a sinistra e sufficientemente piccoli: si veda a tal proposito [25].*

Osservazione 4.6. *Sviluppando la formula di Baker-Campbell-Hausdorff si ottiene*

$$C(X, Y) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] + \frac{1}{12}[Y, [Y, X]] + R_4(X, Y)$$

dove $R_4(X, Y)$ indica una serie formale di commutatori ciascuno di lunghezza 4 o più. Indicando invece con $R_3(X, Y)$ serie di commutatori di lunghezza almeno 3 otteniamo

$$\begin{aligned} & \exp(-Y) \exp(-X) \exp(Y) \exp(X) \\ = & \exp\left(-X - Y + \frac{1}{2}[X, Y] + R_3(X, Y)\right) \circ \\ & \circ \exp\left(X + Y + \frac{1}{2}[X, Y] + R_3(X, Y)\right) \\ = & \exp\left(-X - Y + \frac{1}{2}[X, Y] + X + Y + \frac{1}{2}[X, Y] + \right. \\ & \left. - \frac{1}{2}[Y, X] - \frac{1}{2}[X, Y] + R_3(X, Y)\right) \\ = & \exp([X, Y] + R_3(X, Y)). \end{aligned} \tag{4.2.3}$$

Intuitivamente, se seguiamo prima il flusso del campo X , quindi quello di Y , poi di $-X$ ed infine quello di $-Y$, sarà approssimativamente come l'aver seguito il campo $[X, Y]$. Usando più volte questo procedimento riusciremo ad approssimare il flusso lungo un commutatore di lunghezza h dei X_1, \dots, X_m utilizzando al più 4^{h-1} linee di flusso di tali campi: questo fatto ci consentirà di muoverci lungo le direzioni proibite seguendo solo flussi di campi orizzontali, e questa è l'idea che sta alla base della dimostrazione del teorema di Chow.

Teorema 4.2 (di Chow). *Se i campi X_1, \dots, X_m soddisfano la condizione di Chow-Hörmander allora $d_{CC}(x, y) < +\infty$ per ogni $x, y \in \Omega$: in particolare esistono cammini orizzontali che uniscono x e y .*

Dimostrazione. Per ogni $x \in \Omega$ esiste $k = k(x)$ tale che nel punto x l'intero spazio tangente è generato da commutatori di lunghezza minore o uguale a k dei campi X_1, \dots, X_m ; d'altra parte per la continuità di X_1, \dots, X_m (e dei loro commutatori) questo k sarà tale che in tutto un intorno (euclideo) U_x di x i commutatori di lunghezza minore o uguale a k di tali campi hanno rango massimo.

Sia dunque K un compatto (rispetto alla topologia euclidea) contenuto in Ω : ricoprendolo con una famiglia finita $\{U_{x_1}, \dots, U_{x_q}\}$ di tali intorno deduciamo che la condizione di Chow-Hörmander è soddisfatta su tutto K da commutatori di lunghezza minore o uguale a $\sup\{k(x_1), \dots, k(x_q)\}$. Considerando allora una successione di compatti che invadono Ω , e tenendo conto del fatto che Ω è connesso, il teorema risulterà dimostrato grazie alla seguente proposizione. \square

Proposizione 4.3. *Sia K un insieme compatto rispetto alla topologia euclidea e contenuto in Ω e supponiamo che per ogni $x \in K$ la condizione di Chow-Hörmander sia soddisfatta da commutatori di lunghezza minore o uguale a k . Allora esiste $C = C(K)$ tale che per ogni $x, y \in K$*

$$d_{CC}(x, y) \leq C |x - y|^{\frac{1}{k}}.$$

Dimostrazione. Sia $I = (I_1, \dots, I_r)$ una r -upla (ordinata) di campi I_j scelti fra i $2m$ campi $\pm X_1, \dots, \pm X_m$; se t è sufficientemente piccolo siamo in grado di definire

$$E(I, t)(x) := e^{tI_r} \circ \dots \circ e^{tI_1}(x).$$

È immediato verificare che

$$d_{CC}(x, E(I, t)(x)) \leq |t|r. \quad (4.2.4)$$

Dato un multi-indice $\alpha = (\alpha_1, \dots, \alpha_h)$ di lunghezza h e tale che $1 \leq \alpha_j \leq m$, definiamo il campo X_α dato da

$$X_\alpha = [X_{\alpha_1}, [X_{\alpha_2}, \dots [X_{\alpha_{h-1}}, X_{\alpha_h}] \dots]]$$

e anche di X_α diremo che ha lunghezza h .

La formula di Baker-Campbell-Hausdorff implica che esiste una r -upla $I = (I_1, \dots, I_r)$ tale che $r \leq 4^{h-1}$ e

$$\exp(t^h X_\alpha)(x) = E(I, t)(x) + t^{h+1} O(1) \quad (4.2.5)$$

dove $O(1)$ indica una funzione limitata: per una dimostrazione rigorosa di questo fatto si veda [20]. Ne segue che

$$\exp(\tau X_\alpha) = \begin{cases} E(I, \tau^{1/h})(x) + \tau^{\frac{h+1}{h}} O(1) & \text{se } \tau \geq 0 \\ E(I_-, (-\tau)^{1/h})(x) + (-\tau)^{\frac{h+1}{h}} O(1) & \text{se } \tau < 0 \end{cases}$$

dove I_- indica la sequenza di campi che corrispondono al commutatore $-X_\alpha$ in modo che valga la (4.2.5) con I_- e $-X_\alpha$ al posto di I e X_α rispettivamente.

Definiamo ora

$$E_\alpha(\tau) = \begin{cases} E(I, \tau^{1/h}) & \text{se } \tau \geq 0 \\ E(I_-, (-\tau)^{1/h})(x) & \text{se } \tau < 0. \end{cases}$$

Dimostriamo che $(\tau, x) \mapsto E_\alpha(\tau)(x)$ è di classe \mathcal{C}^1 . È sufficiente mostrare la continuità di $\frac{\partial}{\partial \tau} E_\alpha(\tau)(x)$ per $\tau = 0$; fissato allora $\tau > 0$ e ponendo $\tau = t_0^h$ e $\tau + \Delta\tau = t^h$ abbiamo

$$\begin{aligned} \frac{\partial E_\alpha(\tau)(x)}{\partial \tau} &= \lim_{\Delta\tau \rightarrow 0} \frac{E_\alpha(\tau + \Delta\tau) - E_\alpha(\tau)}{\Delta\tau} \\ &= \lim_{\Delta\tau \rightarrow 0} \frac{E(I, t)(x) - E(I, t_0)(x)}{t^h - t_0^h} \\ &= \frac{1}{ht_0^{h-1}} \frac{\partial E(I, t_0)(x)}{\partial t} \quad (*) \\ &= \frac{1}{ht_0^{h-1}} \frac{\partial \exp(t_0^h X_\alpha)(x)}{\partial t} + \tau^{1/h} O(1) \\ &= X_\alpha(e^{\tau X_\alpha}(x)) + \tau^{1/h} O(1) \end{aligned} \quad (4.2.6)$$

dove nel passaggio contrassegnato da (*) abbiamo usato la (4.2.5). Analogamente per $\tau < 0$ otteniamo

$$\frac{\partial E_\alpha(\tau)(x)}{\partial \tau} = X_\alpha(e^{\tau X_\alpha}(x)) + \tau^{1/h} O(1)$$

che, assieme alla 4.2.6, ci dà la continuità di $\frac{\partial}{\partial \tau} E_\alpha(\tau)(x)$, da cui deduciamo che $(\tau, x) \mapsto E_\alpha(\tau)(x)$ è di classe \mathcal{C}^1 .

Fissiamo ora un punto $x_0 \in K$: poichè è soddisfatta la condizione di Chow-Hörmander possiamo scegliere n campi $X_{\alpha_1}, \dots, X_{\alpha_n}$, fra i commutatori dei campi X_1, \dots, X_m di lunghezza minore o uguale a k , che nel punto x_0 siano linearmente indipendenti. In un intorno dell'origine in \mathbb{R}^n siamo in grado di definire la seguente funzione

$$F(t) = F(t_1, \dots, t_n) := E_{\alpha_n}(t_n) \circ \dots \circ E_{\alpha_1}(t_1)(x_0).$$

Ovviamente F è di classe \mathcal{C}^1 , e la (4.2.6) implica che la matrice

$$\frac{\partial F}{\partial t}(0) = \text{col}[X_{\alpha_1}(x_0), \dots, X_{\alpha_n}(x_0)]$$

è invertibile per come abbiamo scelto gli X_{α_j} . Per il teorema delle funzioni implicite la funzione F è localmente invertibile: esistono cioè $\rho, \sigma, M > 0$ tali che

$$\{|x - x_0| < \sigma\} \subset F(\{|t| < \rho\}) \quad \text{e}$$

$$|F(t) - F(0)| \geq M|t| \text{ per ogni } |t| < \rho.$$

Osserviamo che questo è sufficiente per la dimostrazione del teorema di Chow: abbiamo infatti dimostrato che l'insieme dei punti raggiungibili da x_0 con curve orizzontali è un aperto (della topologia euclidea), dunque deve coincidere con l'intero Ω dal momento che questo è connesso.

Sia ora $x \in \{|x - x_0| < \sigma\}$ e sia $t \in \mathbb{R}^n$, $|t| < \rho$ tale che $F(t) = x$. Definendo per $j = 1, \dots, n$ i punti $x_j := E_{\alpha_j}(t_j)(x_{j-1})$ abbiamo $x = x_n$ e

$$\begin{aligned} d_{CC}(x_0, x) &\leq \sum_{j=1}^n d_{CC}(x_{j-1}, x_j) \\ &= \sum_{j=1}^n d_{CC}(x_{j-1}, E_{\alpha_j}(t_j)(x_{j-1})) \\ &= \sum_{j=1}^n d_{CC}(x_{j-1}, E(I_j, |t_j|^{1/k_j})) \quad (*) \\ &= C_1 \sum_{j=1}^n |t_j|^{1/k_j} \leq C_2 |t|^{1/k} \\ &\leq \frac{C_2}{M^{1/k}} |F(t) - F(t_0)|^{1/k} = C |x - x_0|^{1/k}. \end{aligned}$$

dove nel passaggio contrassegnato da (*) abbiamo sfruttato la (4.2.4). Dal momento che le costanti ρ, σ, M possono essere scelte in modo che dipendano solo dal compatto K , la tesi risulta provata. \square

Osservazione 4.7. *Modificando opportunamente la dimostrazione della proposizione 4.3 si riesce a provare che la distanza di Carnot-Carathéodory si comporta, lungo una direzione che si ottiene come combinazione vettoriale di commutatori di lunghezza h dei campi X_1, \dots, X_m (ma senza che si possa ottenere solo a partire da commutatori di lunghezza inferiore), come la distanza Euclidea elevata alla potenza $1/h$.*

La condizione di Chow-Hörmander è dunque sufficiente a garantire la connessione orizzontale dello spazio di Carnot; tuttavia essa non è in generale necessaria, come dimostra l'esempio 4.1. È stato tuttavia dimostrato che se i campi X_1, \dots, X_m sono analitici essa è anche necessaria: si veda ad esempio [23].

Capitolo 5

Geodetiche negli spazi di Carnot-Carathéodory

Il capitolo è dedicato allo studio delle geodetiche, ovvero delle curve di minima lunghezza, negli spazi di Carnot-Carathéodory. Inizieremo col dimostrarne l'esistenza sotto condizioni piuttosto generali: essenzialmente esiste sempre una curva di lunghezza minima fra due punti se i campi orizzontali X_1, \dots, X_m (quelli che ci servono a definire la metrica di Carnot-Carathéodory) non “scoppiano” all'infinito: tale teorema di esistenza è valido, naturalmente, anche nell'ambito della Geometria Riemanniana. Nella seconda sezione introduciamo il problema del controllo ottimale e il Principio di Pontryagin: saranno questi strumenti a permetterci di dividere la geodetiche in normali e singolari, e di ricavare la regolarità delle prime. Infine, l'ultima sezione è dedicata ad una prima presentazione del gruppo di Heisenberg ed allo studio delle sue geodetiche.

5.1 Esistenza

In questa sezione ci occupiamo di capire se è possibile congiungere due punti assegnati con una curva orizzontale la cui lunghezza è pari esattamente alla distanza di Carnot-Carathéodory fra i due punti. Per semplicità considereremo solo il caso $\Omega = \mathbb{R}^n$.

Definizione 5.1. Una *geodetica* fra due punti $x, y \in \mathbb{R}^n$ è una curva ammissibile $\gamma : [0, T] \rightarrow \mathbb{R}^n$ tale che $\gamma(0) = x$, $\gamma(T) = y$ e $l_H(\gamma) = d_{CC}(x, y)$.

È immediato verificare che una geodetica lo è rispetto a due suoi punti qualsiasi.

Il seguente teorema dà l'esistenza locale della geodetiche negli spazi di Carnot-Carathéodory.

Teorema 5.1. *Sia $U \subset \Omega$ un aperto limitato e sia $M := \sup_{x \in U} \|\mathcal{A}(x)\|$. Siano $x, y \in U$ tali che $Md_{CC}(x, y) < d_E(x, \partial U)$. Allora esiste una geodetica da x a y .*

Dimostrazione. Sia $\{\gamma_j\}_{j \in \mathbb{N}}$ una successione di curve ammissibili da $[0, 1]$ in Ω con $\gamma_j(0) = x$ e $\gamma_j(1) = y$, e tali che $l_H(\gamma_j) \rightarrow d_{CC}(x, y)$. Se $h_j : [0, 1] \rightarrow \mathbb{R}^m$ sono funzioni misurabili tali che $\dot{\gamma}_j = \mathcal{A}(\gamma_j)h_j$ e $|\dot{\gamma}_j|_H = |h_j|$ q.o., allora

$$\|h_j\|_{L^1} = l_H(\gamma_j) \longrightarrow d_{CC}(x, y). \quad (5.1.1)$$

Vogliamo dimostrare che per j sufficientemente grande γ_j è interamente contenuta in U . Per la (5.1.1) e usando l'ipotesi $Md_{CC}(x, y) < d_E(x, \partial U)$ per j sufficientemente grande si ha $M\|h_j\|_{L^1} < d_E(x, \partial U)$. Ragioniamo per assurdo: non è restrittivo supporre che per j sufficientemente grande l'insieme $F_j := \{t : \gamma_j(t) \notin U\}$ sia non vuoto, e quindi possiamo definire $\bar{t}_j := \inf F_j$. Abbiamo allora $\bar{t}_j \in \partial U$ e $\gamma_j(t) \in U$ per ogni $t < \bar{t}_j$ e quindi

$$\begin{aligned} d_E(x, \gamma_j(\bar{t}_j)) &\leq \int_0^{\bar{t}_j} |\dot{\gamma}_j| dt \\ &\leq \int_0^{\bar{t}_j} \|\mathcal{A}(\gamma_j(t))\| |h_j(t)| dt \\ &\leq M\|h_j\|_{L^1} \\ &< d_E(x, \partial U) \end{aligned}$$

che ci dà un assurdo; dunque $F_j = \emptyset$ per ogni j .

A meno di sottosuccessioni le γ_j sono quindi contenute in U , dunque sono equilimitate in L^∞ e quindi in L^1 . Tuttavia anche le $\dot{\gamma}_j$ sono limitate in L^1 , in quanto

$$\|\dot{\gamma}_j\|_{L^1} \leq M\|h_j\|_{L^1} \rightarrow Md_{CC}(x, y) < +\infty.$$

Riparametrizzando se necessario le γ_j in modo che abbiano velocità (orizzontale) costante otteniamo che anche le $\dot{\gamma}_j$ sono equilimitate in L^∞ . In definitiva sia le γ_j che le $\dot{\gamma}_j$ sono equilimitate e in L^1 e in L^∞ ; sfruttando il teorema di Dunford-Pettis 3.4 possiamo quindi estrarre una sottosuccessione, che chiameremo ancora γ_j , tale che $\gamma_j \rightarrow \gamma$ in $AC([0, 1], \Omega)$. Ovviamente $\gamma(0) = x$ e $\gamma(1) = y$: basta quindi dimostrare che $l_H(\gamma) \leq d_{CC}(x, y)$, cosa che sarà vera se dimostreremo che il funzionale l_H è sequenzialmente semicontinuo inferiormente per la convergenza debole in $AC([0, 1], \Omega)$. A tale scopo definiamo l'integrando variazionale

$$\begin{aligned} L &: [0, 1] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, +\infty] \\ &(t, x, v) \longmapsto |v|_H. \end{aligned}$$

Notiamo che $l_H(\gamma) = \int_0^1 L(t, \gamma(t), \dot{\gamma}(t)) dt$; ricordando il teorema 3.1 saremo in grado di concludere se mostreremo che

1. per ogni t la funzione $L(t, \cdot, \cdot)$ è semicontinua inferiormente
2. $L(t, x, \cdot)$ è convessa per ogni t e ogni x .

Il punto 2. segue dalla convessità della norma orizzontale che abbiamo già mostrato. Dato che L non dipende da t , il punto 1. sarà vero se mostreremo che la funzione $(x, v) \mapsto |v|_H$ è semicontinua inferiormente: si noti che si ha dipendenza effettiva dal punto x ($|\cdot|_H$ dipende dal punto in cui la si calcola).

Sia dunque data una successione (x_j, v_j) di punti di $\mathbb{R}^n \times \mathbb{R}^n$ convergenti a (x, v) ; vogliamo dimostrare che $|v|_H \leq \liminf |v_j|_H$. Possiamo naturalmente supporre che il liminf sia invece un lim, e possiamo supporre che sia finito (altrimenti non abbiamo nulla da dimostrare): a meno di escluderne un numero finito, i vettori v_j sono dunque orizzontali (ovviamente visti come elementi del tangente in x_j). Siano quindi $h_j \in \mathbb{R}^m$ tali che $v_j = \mathcal{A}(x_j)h_j$ e $|v_j|_H = |h_j|$: in questo modo gli h_j sono limitati in \mathbb{R}^m , e dunque (a meno di sottosuccessioni) $h_j \rightarrow h \in \mathbb{R}^m$. Per la continuità di \mathcal{A} abbiamo $v = \mathcal{A}(x)h$ e dunque, per definizione di norma orizzontale,

$$|v|_H \leq |h| = \lim |v_j|_H.$$

La dimostrazione è dunque completa. □

5.2 Regolarità

Nello studio della regolarità delle geodetiche degli spazi di Carnot-Carathéodory uno degli strumenti più potenti è il cosiddetto Principio di Pontryagin, che è per certi versi una generalizzazione dell'equazione di Eulero al cosiddetto problema del controllo ottimale. Nella seguente sottosezione introduciamo brevemente tali questioni seguendo la linea di [12], al quale rimandiamo per eventuali approfondimenti.

5.2.1 Il problema del controllo ottimo e il Principio di Pontryagin

Siano dati:

- un sottoinsieme chiuso $V \subset \mathbb{R}^m$;
- una funzione $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ continua e tale che f_x sia continua (dove t, x, v indicano le variabili di \mathbb{R}, \mathbb{R}^n e \mathbb{R}^m rispettivamente);
- una funzione $\phi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^k$ (dove t_0, t_1, x_0, x_1 indicano le variabili di $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$)

Indichiamo inoltre con \mathcal{V} il cosiddetto insieme dei **controlli**, ovvero l'insieme delle funzioni $v : [t_0, t_1] \rightarrow V$ continue a tratti definite sull'intervallo chiuso $[t_0, t_1]$ (che può variare da controllo a controllo).

Per $x_0 \in \mathbb{R}^n$ e $v \in \mathcal{V}$ definita su $[t_0, t_1]$ consideriamo la soluzione $x(t)$ del problema di Cauchy

$$\begin{cases} \dot{x}(t) = f(t, x(t), v(t)) \\ x(t_0) = x_0. \end{cases} \quad (5.2.1)$$

Tale soluzione $x(t)$ si dirà **traiettoria** corrispondente al controllo $v(t)$ ed alla condizione iniziale x_0 , mentre il valore $x(t)$ si dice **stato** del sistema al tempo t . In questo senso si capisce il perchè del termine controllo: data infatti la condizione iniziale, la conoscenza di v ci permette di “controllare” tutta l'evoluzione successiva della traiettoria.

Definizione 5.2. Una coppia $(x_0, v) \in \mathbb{R}^n \times \mathcal{V}$ si dice **ammissibile** se esiste una soluzione dell'equazione del moto (5.2.1) corrispondente al controllo $v(t)$ ed alla condizione iniziale x_0 e tale che

$$\phi_j(t_0, t_1, x(t_0), x(t_1)) = 0 \quad \text{per ogni } j = 2, \dots, k \quad (5.2.2)$$

In pratica vogliamo imporre delle condizioni (per esempio delle condizioni agli estremi) che devono essere soddisfatte dalle coppie (x_0, v) che ammettiamo a competere, e che vengono tradotte nell'annullarsi delle componenti di ϕ esclusa la prima; il **problema del controllo ottimo** consiste invece nel minimizzare questa prima componente fra le coppie ammissibili. Introduciamo perciò il funzionale

$$J(x_0, v) := \phi_1(t_0, t_1, x(t_0), x(t_1)) \quad (5.2.3)$$

dove $[t_0, t_1]$ è l'intervallo di definizione di v e $x(t)$ è la soluzione dell'equazione del moto (5.2.1) relativa a x_0 e v . Per brevità indicheremo con e l'argomento della funzione a secondo membro in (5.2.3), ovvero la quadrupla $(t_0, t_1, x(t_0), x(t_1))$.

Come l'equazione di Eulero, il Principio di Pontryagin dà delle condizioni necessarie affinché una coppia ammissibile (x_0, v) sia **ottimale**, ovvero di minimo per J :

Teorema 5.2 (Principio di Pontryagin). Se la coppia (x_0, v) è ottimale e $x(t)$ è la traiettoria ad essa associata, allora esistono

- un vettore $\lambda \in \mathbb{R}^k$, $\lambda \neq 0$ e $\lambda_1 \leq 0$
- una funzione misurabile $\xi : [t_0, t_1] \rightarrow \mathbb{R}^n$

tali che, indicando con A^T la trasposta di una matrice A , si ha

1. $\dot{\xi}(t)^T = -\xi(t)f_x(t, x(t), v(t))$ per ogni $t \in [t_0, t_1]$;
2. $\xi(t) [f(t, x(t), w) - f(t, x(t), u(t))] \leq 0$ per ogni $w \in V$ e ogni $t \in]t_0, t_1[$;
3. $\xi(t_1)^T = \lambda^T \phi_{x_1}(e)$;
4. $\xi(t_0)^T = -\lambda^T \phi_{x_0}(e)$;
5. $\xi(t_1)^T f(t_1, x(t_1), v(t_1)) = -\lambda^T \phi_{t_1}(e)$;
6. $\xi(t_0)^T f(t_0, x(t_0), v(t_0)) = \lambda^T \phi_{t_0}(e)$;
7. se $f_t(t, x, u)$ è continua, allora per ogni $t \in [t_0, t_1]$ si ha

$$\xi(t)^T f(t, x(t), v(t)) = \lambda^T \phi_{t_0}(e) + \int_{t_0}^t \xi(s)^T f_t(s, x(s), u(s)) ds.$$

La 1. ci dà n condizioni che vengono chiamate **condizioni aggiunte**, mentre le 3.-4.-5.-6. sono dette **condizioni di trasversalità**.

Definiamo l'**Hamiltoniana** $H(t, x, v) := \xi(t)^T f(t, x, v)$; possiamo allora riscrivere la 2. del teorema 5.2 in questo modo:

$$\max_{w \in V} \{H(t, x(t), w)\} = H(t, x(t), v(t))$$

che è quello che si chiama comunemente **Principio del Massimo di Pontryagin**.

Una coppia (x_0, v) si dice **estremale** se verifica le condizioni 1., . . . , 7 del teorema 5.2; un estremale per cui si abbia $\lambda_1 = 0$ si dice **anormale** o **singolare**. Osserviamo che, se $\lambda_1 \neq 0$, possiamo supporre, a meno di rinormalizzare λ e ξ , che si abbia $\lambda_1 = -1$.

5.2.2 Geodetiche normali e singolari

Le geodetiche degli spazi di Carnot-Carathéodory sono soluzioni del seguente problema di controllo nel quale:

- i controlli sono le funzioni misurabili $h : [0, 1] \rightarrow \mathbb{R}^m$;
- l'equazione del moto è data da $\dot{x}(t) = \mathcal{A}(x(t))h(t)$;
- le condizioni agli estremi sono date da $x(0) = x_0$ e $x(1) = x_1$, dove x_0 e x_1 sono due punti assegnati che vogliamo congiungere con una curva orizzontale che realizza la distanza di Carnot-Carathéodory;

- il funzionale da minimizzare è la lunghezza orizzontale della curva x . Ricordando l'osservazione 4.4, sarà equivalente minimizzare la lunghezza orizzontale 2-esima della curva, ovvero il seguente funzionale:

$$J(h) = \frac{1}{2} l_2(x)^2 = \frac{1}{2} \int_0^1 |\dot{\gamma}(t)|_H^2 dt. \quad (5.2.4)$$

Sappiamo che il membro di destra della (5.2.4) è minore o uguale di $\frac{1}{2} \|h\|_{L^2}^2$, ma anche che, fissata una curva x , esiste una h per cui vale il segno di uguale. Possiamo quindi ridurci a studiare il funzionale

$$J(h) = \frac{1}{2} \|h\|_{L^2}^2 = \frac{1}{2} \int_0^1 |h(t)|^2 dt. \quad (5.2.5)$$

Il Principio di Pontryagin applicato a questo problema ci consente di enunciare il seguente

Teorema 5.3. *Se il controllo h è ottimale e x è la soluzione dell'equazione del moto relativa ad h , allora esistono $\lambda \in \{0, 1\}$ e $\xi \in Lip([0, 1], \mathbb{R}^n)$ tali che:*

1. $|\xi(t)| + \lambda \neq 0$ per ogni $t \in [0, 1]$;
2. $\dot{\xi} = -\frac{\partial}{\partial x} \langle \mathcal{A}(x)h, \xi \rangle$ quasi ovunque in $[0, 1]$;
3. $\langle \mathcal{A}(x)h, \xi \rangle - \frac{1}{2} \lambda |h|^2 = \max_{w \in \mathbb{R}^m} \left\{ \langle \mathcal{A}(x)w, \xi \rangle - \frac{1}{2} \lambda |w|^2 \right\}$ quasi ovunque in $[0, 1]$

Se $\lambda = 1$ la geodetica si dice **normale**, se $\lambda = 0$ **singolare** o anche **anormale**.

Le geodetiche sub-riemanniane normali sono localmente minimizzanti: si veda a tal proposito [18].

Osservazione 5.1. *Nel caso Riemanniano, in cui il rango di $\mathcal{A}(x)$ è massimo per ogni x (ovvero tutte le direzioni sono orizzontali), tutte le geodetiche sono normali. Se infatti fosse $\lambda = 0$ avremmo che la 3. del teorema 5.3 diventa*

$$\langle \mathcal{A}(x)h, \xi \rangle = \max_{w \in \mathbb{R}^m} \langle \mathcal{A}(x)w, \xi \rangle. \quad (5.2.6)$$

Ma per ogni x la matrice $\mathcal{A}(x)$, vista come applicazione lineare da \mathbb{R}^m in \mathbb{R}^n , è suriettiva: ne segue che il secondo membro della (5.2.6) non è finito, a meno che non si abbia $\xi \equiv 0$: anche questo caso tuttavia conduce ad un assurdo perchè contraddice la 1. dello stesso teorema 5.3.

Nel caso subriemanniano possono effettivamente verificarsi casi di geodetiche singolari: si vedano ad esempio [17] e [24].

L'Hamiltoniana per il nostro problema è

$$H(x, \xi) = \sum_{j=1}^m \langle X_j(x), \xi \rangle^2 = |\mathcal{A}(x)^T \xi|^2. \quad (5.2.7)$$

All'Hamiltoniana è collegato il seguente sistema di equazioni, che viene chiamato sistema di **equazioni di Hamilton**:

$$\begin{cases} \dot{x} = \frac{1}{2} \frac{\partial H(x, \xi)}{\partial \xi} \\ \dot{\xi} = -\frac{1}{2} \frac{\partial H(x, \xi)}{\partial x}. \end{cases} \quad (5.2.8)$$

Il seguente teorema risulta fondamentale nello studio della regolarità delle geodetiche degli spazi di Carnot-Carathéodory:

Teorema 5.4. *Una geodetica normale x e la sua variabile duale ξ risolvono le equazioni di Hamilton (5.2.8). Una geodetica singolare e la sua variabile duale ξ soddisfano l'equazione $H(x(t), \xi(t)) = 0$.*

Dimostrazione. Se $\lambda = 1$ dalla 3. del teorema 5.3 abbiamo

$$\begin{aligned} \langle \mathcal{A}(x)h, \xi \rangle - \frac{1}{2}|h|^2 &= \max_{w \in \mathbb{R}^m} \left\{ \langle \mathcal{A}(x)w, \xi \rangle - \frac{1}{2}|w|^2 \right\} \\ &= \max_{w \in \mathbb{R}^m} \left\{ \langle w, \mathcal{A}(x)^T \xi \rangle - \frac{1}{2}|w|^2 \right\} \end{aligned}$$

e il massimo a secondo membro viene raggiunto in $w = \mathcal{A}(x)^T \xi$. Dunque $h(t) = \mathcal{A}(x(t))^T \xi(t)$: sostituendo questa nella 2. e nell'equazione di stato otteniamo le equazioni di Hamilton (5.2.8).

Se invece $\lambda = 0$ la 3. diventa

$$\langle h, \mathcal{A}(x)^T \xi \rangle = \max_{w \in \mathbb{R}^m} \langle w, \mathcal{A}(x)^T \xi \rangle$$

che implica $\mathcal{A}(x)^T \xi = 0$. Ricordando la definizione di Hamiltoniana (5.2.7) otteniamo $H(x(t), \xi(t)) = 0$. \square

Il seguente corollario si ottiene immediatamente dal teorema 5.4:

Corollario 5.1. *Le geodetiche degli spazi di Carnot-Carathéodory normali hanno una regolarità pari almeno a quella dell'Hamiltoniana H cioè, in ultima analisi, a quella dei campi X_j .*

Il problema della regolarità delle geodetiche degli spazi di Carnot-Carathéodory risulta dunque notevolmente semplificato: infatti le uniche che potrebbero creare dei problemi sono, per il corollario 5.1, solo quelle singolari.

Introduciamo ora una condizione, più forte di quella di Chow-Hörmander, che appare più volte nella letteratura relativa agli spazi sub-riemanniani:

Definizione 5.3. *Indicando con \mathbb{H} il sottofibrato del tangente a \mathbb{R}^n generato dai campi X_1, \dots, X_m e con $\mathbb{H}(x)$ il sottospazio vettoriale di $T_x\mathbb{R}^n$ generato dai campi di \mathbb{H} valutati in x , diciamo che vale la **strong bracket generating condition** se*

$$\mathbb{H}(x) + [V, \mathbb{H}](x) = \mathbb{R}^n \quad \text{per ogni } x \text{ e ogni } V \in \mathbb{H} \setminus \{0\} \quad (5.2.9)$$

dove $[V, \mathbb{H}]$ indica il sottofibrato del tangente generato dai commutatori fra il campo V e i campi di \mathbb{H} , mentre $[V, \mathbb{H}](x)$ indica il sottospazio di $T_x\mathbb{R}^n$ generato dalla valutazione in x dei campi di $[V, \mathbb{H}]$.

Il seguente risultato dovuto a Strichartz ([21] e [22]) dà una condizione affinché tutte le geodetiche siano normali e quindi, in definitiva, regolari:

Teorema 5.5. *Se vale la strong bracket generating condition (5.2.9) ogni geodetica sub-riemanniana è normale.*

5.3 Un esempio concreto: il gruppo di Heisenberg

In questa sezione introdurremo quella che probabilmente è la varietà sub-riemanniana più famosa e studiata: il gruppo di Heisenberg. Il nostro scopo sarà quello di calcolare esplicitamente le geodetiche di tale struttura.

L' n -esimo gruppo di Heisenberg \mathbb{H}^n è lo spazio di Carnot-Carathéodory che si ottiene considerando i $2n$ campi orizzontali su $\mathbb{R}^{2n+1} = \mathbb{R}_x^n \times \mathbb{R}_y^n \times \mathbb{R}_t$

$$\begin{aligned} X_j(x, y, t) &= \partial_{x_j} + 2y_j \partial_t & j &= 1, \dots, n \\ Y_j(x, y, t) &= \partial_{y_j} - 2x_j \partial_t & j &= 1, \dots, n \end{aligned}$$

Per ogni $j = 1, \dots, n$ si ha che $[X_j, Y_j] = -4\partial_t$, mentre tutti gli altri commutatori sono nulli.

Il sostantivo “gruppo” deriva dal fatto che, come vedremo più avanti, \mathbb{H}^n può essere dotato della struttura di gruppo di Carnot. L’unica proprietà che ci servirà in questo contesto è l’invarianza della metrica rispetto alla moltiplicazione a sinistra per un elemento qualsiasi del gruppo: questo ci consentirà di ricondurci al caso in cui il punto di partenza è l’origine, mentre quello di arrivo resta arbitrario.

Verifichiamo che tutte le geodetiche del gruppo di Heisenberg sono normali. Come al solito sia $\mathcal{A}(x) = \text{col}[X_1(x), \dots, X_n(x), Y_1(x), \dots, Y_n(x)]$, e sia $\gamma : [0, 1] \rightarrow \mathbb{H}^n$ una geodetica tale che $\dot{\gamma} = \mathcal{A}(\gamma)h$. Scriveremo $\gamma(s) = (x(s), y(s), t(s)) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ e $h(s) = (h_1(s), h_2(s)) \in \mathbb{R}^n \times \mathbb{R}^n$: in particolare $\dot{x} = h_1$ e $\dot{y} = h_2$. Supponiamo per assurdo che γ sia singolare, ovvero che si abbia $\lambda = 0$ nel teorema 5.3; lo stesso teorema ci dice che, posto $\xi = (\xi, \eta, \tau) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$, si ha

$$\begin{cases} \dot{\xi} = 2\tau h_2 \\ \dot{\eta} = -2\tau h_1 \\ \dot{\tau} = 0, \end{cases} \quad (5.3.1)$$

e quindi $\tau(s) \equiv \beta \in \mathbb{R}$. Inoltre il teorema 5.4 ci dice che, essendo la geodetica singolare, si deve avere

$$\begin{cases} \xi + 2\beta y = 0 \\ \eta - 2\beta x = 0. \end{cases} \quad (5.3.2)$$

Se $\beta = 0$ allora $\xi \equiv \eta \equiv 0$, ma questo non è possibile per la 1. del teorema 5.3. Dunque $\beta \neq 0$. Differenziando allora le equazioni (5.3.2), andando a sostituire il risultato nelle (5.3.1) e semplificando $2\beta \neq 0$ otteniamo

$$\begin{cases} \dot{x} = -h_1 \\ \dot{y} = -h_2. \end{cases}$$

Tuttavia sappiamo che $\dot{x} = h_1$ e $\dot{y} = h_2$ e quindi $h = 0$ quasi ovunque; questo non è possibile a meno che γ non sia una curva costante.

Abbiamo dunque dimostrato che tutte le geodetiche del gruppo di Heisenberg sono normali; il teorema 5.4 ci dice allora che esse soddisfano le equazioni di Hamilton (5.2.8), e questo ci consente di calcolarle esplicitamente.

Posto

$$\begin{aligned} x &= (x, y, t) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \\ \xi &= (\xi, \eta, \tau) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \end{aligned}$$

l'Hamiltoniana diventa

$$H((x, y, t), (\xi, \eta, \tau)) = \sum_{j=1}^n (\xi_j + 2y_j\tau)^2 + (\eta_j - 2x_j\tau)^2$$

Le equazioni di Hamilton si scrivono allora come

$$\begin{cases} \dot{x} = \xi + 2\tau y & x(0) = 0 \\ \dot{y} = \eta - 2\tau x & y(0) = 0 \\ \dot{t} = 4\tau(|x|^2 + |y|^2) + 2\langle x, \eta \rangle - 2\langle y, \xi \rangle & t(0) = 0 \\ \dot{\xi} = 2\tau\eta - 4\tau^2 x & \xi(0) = A \\ \dot{\eta} = -2\tau\xi - 4\tau^2 y & \eta(0) = B \\ \dot{\tau} = 0 & \tau(0) = \varphi/4. \end{cases}$$

La risoluzione di tale sistema ci dà

$$\begin{cases} x(s) = \frac{A(1 - \cos \varphi s) + B \sin \varphi s}{\varphi} \\ y(s) = \frac{-B(1 - \cos \varphi s) + A \sin \varphi s}{\varphi} \\ t(s) = 2 \frac{\varphi s - \sin \varphi s}{\varphi^2} (|A|^2 + |B|^2) \end{cases} \quad (5.3.3)$$

nel caso $\varphi \neq 0$; se invece $\varphi = 0$ si ha

$$\begin{cases} x(s) = Bs \\ y(s) = As \\ t(s) = 0. \end{cases} \quad (5.3.4)$$

Se $\gamma(s)$ è la curva di componenti $(x(s), y(s), t(s))$ come in (5.3.3) o in (5.3.4), si ha

$$\dot{\gamma}(s) = \sum_{j=1}^n (A_j \sin \varphi s + B_j \cos \varphi s) X_j(\gamma(s)) + (A_j \cos \varphi s - B_j \sin \varphi s) Y_j(\gamma(s)).$$

Dunque la parametrizzazione per lunghezza d'arco si ha per $|A|^2 + |B|^2 = 1$.

Nel resto della sezione assumeremo fissata la parametrizzazione per lunghezza d'arco $|A|^2 + |B|^2 = 1$; vogliamo dimostrare che ogni curva γ del tipo (5.3.3) è minimizzante fino al tempo $s = 2\pi/\varphi$, cioè fino a quando non ripassa per l'asse t nel punto di coordinate $(0, 0, t)$, dove $t = 4\pi/\varphi^2$. Per comodità supponiamo $t > 0$. Essendo γ subunitaria abbiamo $d_{CC}(0, (0, 0, t)) \leq 2\pi/\varphi$. Sia $\gamma_1 : [0, T] \rightarrow \mathbb{H}^n$ una geodetica minimale subunitaria che congiunge 0 e $(0, 0, t)$ e supponiamo che essa sia relativa ai parametri A_1, B_1 e φ_1 , con $|A_1|^2 + |B_1|^2 = 1$: è facile verificare che γ non può essere del tipo (5.3.4) e dunque $\varphi_1 \neq 0$. Dimostriamo che $T = 2\pi/\varphi_1$. Dato che $t > 0$ e $\alpha \geq \sin \alpha$ se e solo se $\alpha \geq 0$ otteniamo $\varphi_1 > 0$; inoltre

$$\begin{cases} A_{1j}(1 - \cos \varphi_1 T) + B_{1j} \sin \varphi_1 T = 0 \\ A_{1j} \sin \varphi_1 T + B_{1j}(1 - \cos \varphi_1 T) = 0 \end{cases}$$

per ogni $j = 1, \dots, n$. Dato che $|A_1|^2 + |B_1|^2 = 1$, il determinante di questo sistema deve annullarsi; svolgendo i calcoli si ottiene $1 - \cos \varphi_1 T = 0$ e quindi $\varphi_1 T = 2k\pi$ per un certo $k \in \mathbb{N}^*$. Dato che $\gamma_1(T) = (0, 0, t)$ dobbiamo poi avere $T = \sqrt{k\pi t}$, e per la minimalità di γ si ha $\sqrt{k\pi t} \leq \sqrt{\pi t}$. In definitiva abbiamo $k = 1$, da cui $T = 2\pi/\varphi_1$. Concludiamo notando che $4\pi/\varphi^2 = t = 4\pi/\varphi_1^2$, da cui $\varphi_1 = \varphi$ e quindi $T = 2\pi/\varphi$, che implica che γ è minimizzante.

Le curve del tipo (5.3.4) sono invece sempre minimizzanti, in quanto le geodetiche del primo tipo incontrano l'iperpiano $\{t = 0\}$ solo nell'origine.

Tuttavia l'osservazione più interessante è forse la seguente: a φ fissato, tutte le geodetiche parametrizzate per lunghezza d'arco passano, al tempo $s = \frac{2\pi}{\varphi}$, per il punto $(0, 0, \frac{4\pi}{\varphi^2})$. Questo significa che esiste una famiglia continua di curve (tutte minimizzanti) che congiungono l'origine col punto $(0, 0, \frac{4\pi}{\varphi^2})$. Poichè al variare di φ otteniamo tutto l'asse $(0, 0, t)$, questo significa che troviamo punti arbitrariamente vicini all'origine che sono congiunti ad essa da più curve minimizzanti: contrariamente a quanto succede nella geometria riemanniana, le geodetiche degli spazi di Carnot-Carathéodory non sono uniche neppure localmente.

Osservazione 5.2. *Dal momento che le curve orizzontali in Heisenberg sono solo quelle tali che $\dot{t} = 2\langle y, \dot{x} \rangle - 2\langle x, \dot{y} \rangle$, è chiaro che siamo in grado di ricostruire una curva orizzontale a partire dalla sua proiezione su $\mathbb{R}_x^n \times \mathbb{R}_y^n$ e da un suo punto fissato (in \mathbb{R}^{2n+1}). Inoltre è facile verificare che la lunghezza (di Carnot-Carathéodory) della curva coincide con quella Euclidea della sua proiezione.*

In particolare per $n = 1$ sia $\gamma(s) = (x(s), y(s), t(s))$ una curva orizzontale tale che $\gamma(0) = 0$, e consideriamo l'area $A(s)$ spazzata (fino al tempo s) dal raggio vettore (in \mathbb{R}^2) che congiunge 0 col punto $(x(\cdot), y(\cdot))$. Inoltre $\alpha := dt = 2x dy - 2y dx$ è una uno-forma tale che $d\alpha = 4 dx \wedge dy$ è pari a 4 volte la forma d'area, e quindi per il teorema di Stokes

$$A(s) = \int_{(x(\cdot), y(\cdot))} \alpha = t(s) - t(0).$$

Inoltre le geodetiche, proiettate, appaiono o come rette (quelle del tipo (5.3.4)) o come circonferenze (quelle del tipo (5.3.3)). Questo ci dice che sull'asse t la distanza dall'origine è, a meno di costanti, pari a $t^{1/2}$; infatti la proiezione su \mathbb{R}^2 di una geodetica minimizzante che congiunge 0 e $(0, 0, t)$ è una circonferenza che racchiude un'area pari a t , mentre la sua lunghezza è la radice di tale area (il tutto, ovviamente, a meno di costanti moltiplicative).

Capitolo 6

Gruppi di Carnot

In questo capitolo presentiamo alcune particolari varietà sub-riemanniane: i gruppi di Carnot. Essi sono la classe di gruppi di Lie (la cui teoria viene accennata nella prima sezione) di dimensione finita, connessi, semplicemente connessi e la cui algebra ammette una stratificazione, ovvero è della forma $\mathfrak{g} = V_1 \oplus \cdots \oplus V_k$ con $V_j = [V_1, V_{j-1}]$ per $j = 2, \dots, k$ e $[V_1, V_k] = \{0\}$: questo porta alla definizione di un particolare gruppo di automorfismi a un parametro, le dilatazioni δ_λ . La presenza di una stratificazione ci permette anche di definire, nella sezione 3, le cosiddette coordinate graduate, ovvero un isomorfismo con un \mathbb{R}^n (dotato di una certa struttura di gruppo) che risulta anche un diffeomorfismo. Come già accennato nel precedente capitolo, fra i gruppi di Carnot vi è l'esempio più famoso e studiato di varietà sub-riemanniana: il gruppo di Heisenberg. Dopo aver visto quale forma assumono, tramite le coordinate graduate, i campi invarianti a sinistra, andremo a studiare, nella quarta sezione, la struttura di Carnot-Carathéodory che risulta dal considerare orizzontali i campi del primo strato V_1 , e presteremo particolare attenzione al legame fra la distanza e le dilatazioni. Infine, nella quinta e ultima sezione, enunceremo il teorema di Pansu, che costituisce la generalizzazione naturale ai gruppi di Carnot del classico risultato di Rademacher sulla derivabilità delle funzioni Lipschitziane; di esso daremo la dimostrazione solo nel caso di mappe definite su \mathbb{R} .

6.1 Gruppi di Lie e algebre relative

In questa sezione esponiamo brevemente la teoria dei gruppi e delle algebre di Lie: per approfondimenti si veda [25].

Definizione 6.1. *Un **gruppo di Lie** \mathbb{G} è una varietà differenziale dotata della struttura di gruppo differenziabile, nella quale cioè le mappe*

$$\begin{aligned}\mathbb{G} \times \mathbb{G} \ni (x, y) &\longmapsto xy \in \mathbb{G} \\ \mathbb{G} \ni x &\longmapsto x^{-1} \in \mathbb{G}\end{aligned}$$

sono di classe \mathcal{C}^∞ .

Indicheremo con e l'identità del gruppo, mentre per ogni $g \in \mathbb{G}$ denoteremo con l_g la traslazione a sinistra, ovvero la mappa (che risulta di classe \mathcal{C}^∞)

$$\begin{aligned}l_g &: \mathbb{G} \longrightarrow \mathbb{G} \\ x &\longmapsto g \cdot x.\end{aligned}$$

Definizione 6.2. *Uno spazio vettoriale \mathfrak{g} è un'algebra di Lie se esiste una mappa $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ bilineare, antisimmetrica e soddisfacente l'identità di Jacobi $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ per ogni $X, Y, Z \in \mathfrak{g}$.*

Ad esempio lo spazio $\Gamma(TM)$ dei campi vettoriali su una varietà differenziale M è un'algebra di Lie se munita del prodotto $[X, Y] = XY - YX$ dato dal commutatore dei campi.

Definizione 6.3. *Un campo $X \in \Gamma(T\mathbb{G})$ si dice **invariante a sinistra** se per ogni $g \in \mathbb{G}$ si ha*

$$X(g) = dl_g(X(e)).$$

Indichiamo con \mathcal{G} il sottinsieme di $\Gamma(T\mathbb{G})$ formato dai campi invarianti a sinistra: dato che i commutatori di tali campi sono ancora invarianti a sinistra, \mathcal{G} risulta un'algebra di Lie. Tale algebra è canonicamente isomorfa a $T_e\mathbb{G}$, dove l'isomorfismo è dato da

$$T_e\mathbb{G} \ni v \longleftrightarrow X \in \mathcal{G} \text{ tale che } X(g) = dl_g(v).$$

Dato $g \in \mathbb{G}$ e $X \in \mathcal{G}$ consideriamo la curva γ_g^X soluzione della seguente equazione differenziale:

$$\begin{cases} \dot{\gamma}_g^X(t) = X(\gamma_g^X(t)) \\ \gamma_g^X(0) = g \end{cases} \quad (6.1.1)$$

La curva γ_g^X è definita per ogni $t \in \mathbb{R}$, dato che i campi invarianti a sinistra sono completi: in effetti si ha $\gamma_g^X(t+s) = \gamma_g^X(s) \cdot \gamma_g^X(t)$, e questo è sufficiente per ottenere la completezza.

Definizione 6.4. *Dati $g \in \mathbb{G}$ e $X \in \mathcal{G}$ poniamo $\exp(X)(g) := \gamma_g^X(1)$, dove γ_g^X è la soluzione dell'equazione differenziale (6.1.1). La **mappa esponenziale** $\exp : \mathcal{G} \rightarrow \mathbb{G}$ è definita tramite $\exp(X) := \exp(X)(e)$.*

In questo modo si ha $\exp(X)(g) = g \cdot \exp(X)$ e quindi per ogni $X, Y \in \mathcal{G}$ vale

$$\exp(X)(\exp(Y)) = \exp(Y) \cdot \exp(X).$$

Ricordiamo inoltre che $\exp(X) \cdot \exp(Y) = \exp(C(X, Y))$, dove $C(X, Y)$ è dato dalla formula di Baker-Campbell-Hausdorff (4.2.2).

Date due sottoalgebre $\mathfrak{a}, \mathfrak{b}$ di un'algebra di Lie \mathfrak{g} indichiamo con $[\mathfrak{a}, \mathfrak{b}]$ lo spazio vettoriale generato dagli elementi di $\{[X, Y] : X \in \mathfrak{a}, Y \in \mathfrak{b}\}$. Infine poniamo $\mathfrak{g}_1 := \mathfrak{g}$ e per induzione $\mathfrak{g}_{j+1} := [\mathfrak{g}, \mathfrak{g}_j]$.

Definizione 6.5. Diciamo che \mathfrak{g} è **nilpotente** di passo k se $\mathfrak{g}_k \neq \{0\}$ e $\mathfrak{g}_{k+1} = \{0\}$.

Analogamente diremo che un gruppo di Lie \mathbb{G} è nilpotente di passo k se tale è l'algebra \mathcal{G} ad esso associata.

Infine riportiamo il seguente, fondamentale risultato:

Teorema 6.1. Sia \mathbb{G} un gruppo di Lie nilpotente, connesso e semplicemente connesso; allora $\exp : \mathcal{G} \rightarrow \mathbb{G}$ è un diffeomorfismo.

Nel seguito supporremo sempre che \mathbb{G} sia connesso e semplicemente connesso.

6.2 Gruppi di Carnot

Definizione 6.6. Un'algebra di Lie \mathfrak{g} si dice **stratificata** se ammette una stratificazione, cioè se esistono dei sottospazi lineari V_1, \dots, V_k di \mathfrak{g} tali che

$$\begin{aligned} \mathfrak{g} &= V_1 \oplus \dots \oplus V_k \\ V_j &= [V_1, V_{j-1}] \quad \text{per } j = 2, \dots, k \\ [V_1, V_k] &= \{0\}. \end{aligned}$$

Un gruppo \mathbb{G} si dice stratificato se la sua algebra di Lie \mathcal{G} ammette una stratificazione. Se \mathbb{G} ha dimensione finita ed è stratificato, allora è anche nilpotente di passo pari al numero di sottospazi che compaiono nella stratificazione.

La presenza di una stratificazione permette di definire un importante gruppo a un parametro $\{\delta_\lambda\}$ di dilatazioni dell'algebra. Fissato $\lambda \geq 0$ definiamo infatti $\delta_\lambda X = \lambda^j X$ se $X \in V_j$, mappa che estendiamo poi per linearità a tutta l'algebra. Infine definiamo δ_λ anche per $\lambda < 0$ tramite $\delta_\lambda X = -\delta_{|\lambda|} X$.

È immediato verificare le seguenti proprietà delle dilatazioni:

1. $\delta_{\lambda\mu} = \delta_\lambda \circ \delta_\mu$;
2. $\delta_\lambda([X, Y]) = [\delta_\lambda X, \delta_\lambda Y]$;

$$3. \delta_\lambda(C(X, Y)) = C(\delta_\lambda X, \delta_\lambda Y).$$

Definizione 6.7. Un **gruppo di Carnot** \mathbb{G} è un gruppo di Lie di dimensione finita, connesso, semplicemente connesso e stratificato (e dunque nilpotente). Se k è come nella definizione 6.6 diremo che \mathbb{G} è di passo k .

La proprietà fondamentale dei gruppi di Carnot è quella che la mappa $\exp : \mathcal{G} \rightarrow \mathbb{G}$ risulta un diffeomorfismo grazie al teorema 6.1: la sua inversa \exp^{-1} ci consente allora di definire un gruppo a un parametro di automorfismi di \mathbb{G} , che per comodità indicheremo ancora con $\{\delta_\lambda\}$, tramite la $\delta_\lambda := \exp_* \delta_\lambda$, ovvero

$$\delta_\lambda(x) = \exp(\delta_\lambda(\exp^{-1}(x))).$$

Dalle proprietà delle dilatazioni delle algebre di Lie discendono direttamente le seguenti proprietà delle dilatazioni dei gruppi di Carnot:

$$1. \delta_{\lambda\mu} = \delta_\mu \circ \delta_\lambda, \text{ infatti}$$

$$\begin{aligned} \delta_{\lambda\mu}(x) &= \exp(\delta_{\lambda\mu} \exp^{-1}(x)) \\ &= \exp(\delta_\lambda \delta_\mu \exp^{-1}(x)) \\ &= \exp(\delta_\lambda \exp^{-1}(\exp \delta_\mu \exp^{-1}(x))) \\ &= \exp(\delta_\lambda \exp^{-1}(\delta_\mu(x))) = \delta_\lambda \delta_\mu(x) \end{aligned}$$

$$2. \delta_\lambda(x \cdot y) = \delta_\lambda(x) \cdot \delta_\lambda(y), \text{ infatti}$$

$$\begin{aligned} \delta_\lambda(x \cdot y) &= \exp \delta_\lambda \exp^{-1}(x \cdot y) \\ &= \exp \delta_\lambda (C(\exp^{-1} x, \exp^{-1} y)) \\ &= \exp (C(\delta_\lambda \exp^{-1} x, \delta_\lambda \exp^{-1} y)) \\ &= \exp(\delta_\lambda \exp^{-1}(x)) \cdot \exp(\delta_\lambda \exp^{-1}(y)) = \delta_\lambda(x) \cdot \delta_\lambda(y). \end{aligned}$$

6.3 Coordinate graduate

Ci proponiamo ora di dimostrare che un gruppo di Carnot \mathbb{G} è sempre isomorfo (come gruppo) e diffeomorfo (come varietà) a \mathbb{R}^n munito di un certo prodotto di gruppo, dove n è la dimensione topologica di \mathbb{G} .

Sia infatti X_1, \dots, X_n una base vettoriale dell'algebra di Lie \mathcal{G} dei campi vettoriali invarianti a sinistra; dati $X, Y \in \mathcal{G}$ allora $X = \sum_{j=1}^n x_j X_j$ e $Y = \sum_{j=1}^n y_j X_j$ per certi $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ in \mathbb{R}^n .

Definizione 6.8. *Un sistema di coordinate esponenziali relativo alla base X_1, \dots, X_n è la mappa*

$$F : \mathbb{R}^n \longrightarrow \mathbb{G}$$

$$x \longmapsto \exp \left(\sum_{j=1}^n x_j X_j \right). \quad (6.3.1)$$

Introduciamo su \mathbb{R}^n la legge di gruppo che si ricava dall'imporre la condizione

$$x \cdot y = z \iff \sum_{j=1}^n z_j X_j = C \left(\sum_{j=1}^n x_j X_j, \sum_{j=1}^n y_j X_j \right).$$

È semplice verificare che l'identità del gruppo è l'origine e che $x^{-1} = -x$.

Così facendo \mathbb{R}^n , munito di tale prodotto, risulta un gruppo di Lie con algebra di Lie isomorfa a \mathcal{G} ; inoltre la legge di gruppo è stata posta ad hoc in modo che la mappa F in (6.3.1) risulti un isomorfismo di gruppi. Ma poichè sia \mathbb{G} che \mathbb{R}^n sono nilpotenti e semplicemente connessi, il teorema 6.1 ci dice che sono entrambi diffeomorfi a \mathcal{G} , e in definitiva diffeomorfi essi stessi.

In quanto detto fino ad ora abbiamo sfruttato solamente la nilpotenza e la semplice connessione di \mathbb{G} : in effetti per un gruppo di Carnot esiste una scelta privilegiata della base X_1, \dots, X_n di \mathcal{G} che discende dalla presenza di una stratificazione e che consente di leggere in maniera molto semplice le dilatazioni di \mathbb{G} .

Sia dunque $\mathcal{G} = V_1 \oplus \dots \oplus V_k$ una stratificazione e poniamo, per $j = 1, \dots, k$, $n_j := \dim V_j$ e $m_j := n_1 + \dots + n_j$. Inoltre sia $m_0 := 0$. Diremo che una base X_1, \dots, X_n di \mathcal{G} è **adattata** alla stratificazione V_1, \dots, V_k se per ogni $j = 1, \dots, k$ gli n_j vettori $X_{m_{j-1}+1}, \dots, X_{m_j}$ formano una base di V_j .

Definizione 6.9. *Un sistema di coordinate esponenziali si dirà sistema di **coordinate graduate** se è relativo ad una base adattata ad una stratificazione di \mathcal{G} .*

Chiameremo **grado** della coordinata y_j l'unico intero positivo d_j tale che $m_{d_j-1} < j \leq m_{d_j}$.

Sia dunque $F : \mathbb{R}^n \rightarrow \mathbb{G}$ un sistema di coordinate graduate: per comodità indicheremo ancora con δ_λ le mappe da \mathbb{R}^n in sè indotte dalle dilatazioni δ_λ di \mathbb{G} , in modo da poter scrivere $\delta_\lambda \circ F = F \circ \delta_\lambda$. A questo punto le dilatazioni diventano semplicemente

$$\delta_\lambda : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$x \longmapsto \sum_{j=1}^n \lambda^{d_j} x_j e_j =$$

$$= (\lambda x_1, \dots, \lambda x_{m_1}, \lambda^2 x_{m_1+1}, \dots, \lambda^2 x_{m_2}, \dots, \lambda^k x_{m_{k-1}+1}, \dots, \lambda^k x_n)$$

per $\lambda \geq 0$, mentre come al solito $\delta_\lambda(x) = -\delta_{|\lambda|}(x)$ per $\lambda < 0$.

6.3.1 Il gruppo di Heisenberg

A titolo di esempio calcoliamo come appare il gruppo di Heisenberg letto in coordinate graduate, a partire dalla sua algebra di Lie.

Definizione 6.10. *Un'algebra di Lie con base $(X_1, \dots, X_n, Y_1, \dots, Y_n, T)$ che soddisfa le relazioni*

$$\begin{aligned} [X_j, X_k] &= [Y_j, Y_k] = 0 \\ [X_j, Y_k] &= -4\delta_{jk}T \end{aligned} \quad (6.3.2)$$

per ogni $j = 1, \dots, n$ si dice **algebra di Heisenberg** e si indica con \mathfrak{h}_n .

L'algebra di Heisenberg è nilpotente di passo 2 ed ammette la stratificazione $\mathfrak{h}_n = V_1 \oplus V_2$, dove $V_1 = \text{span}\{X_1, \dots, X_n, Y_1, \dots, Y_n\}$ e $V_2 = [V_1, V_1] = \text{span}\{T\}$.

Definizione 6.11. *Il gruppo di Heisenberg \mathbb{H}^n è il gruppo di Lie nilpotente e semplicemente connesso associato ad \mathfrak{h}_n .*

Essendo \mathfrak{h}_n nilpotente di passo 2, la formula di Baker-Campbell-Hausdorff (4.2.2) si riduce a

$$C(X, Y) = X + Y + \frac{1}{2}[X, Y]$$

In questo modo se

$$\begin{aligned} X &= \sum_{j=1}^n x_j X_j + \sum_{j=1}^n y_j Y_j + tT \\ Y &= \sum_{j=1}^n \xi_j X_j + \sum_{j=1}^n \eta_j Y_j + \tau T \end{aligned}$$

per certi $(x, y, t), (\xi, \eta, \tau) \in \mathbb{R}^{2n+1}$, allora

$$[X, Y] = 4(\langle \xi, y \rangle - \langle x, \eta \rangle) T$$

e quindi

$$C(X, Y) = \sum_{j=1}^n (x_j + \xi_j) X_j + \sum_{j=1}^n (y_j + \eta_j) Y_j + (t + \tau + 2\langle \xi, y \rangle - 2\langle x, \eta \rangle) T \quad (6.3.3)$$

Il gruppo di Heisenberg \mathbb{H}^n è dunque isomorfo a $\mathbb{R}^{2n+1} = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ con la legge di gruppo

$$\begin{pmatrix} x \\ y \\ t \end{pmatrix} \cdot \begin{pmatrix} \xi \\ \eta \\ \tau \end{pmatrix} = \begin{pmatrix} x + \xi \\ y + \eta \\ t + \tau + 2\langle \xi, y \rangle - 2\langle x, \eta \rangle \end{pmatrix}$$

In tale struttura di gruppo l'identità risulta essere l'origine di \mathbb{R}^{2n+1} e l'inverso di un elemento x è l'elemento $-x$.

Per $\lambda > 0$ le dilatazioni sono date da $\delta_\lambda(x, y, t) = (\lambda x, \lambda y, \lambda^2 t)$: si noti che sono effettivamente omomorfismi di gruppi.

Cerchiamo una base dei campi vettoriali invarianti a sinistra: sappiamo che un campo X invariante a sinistra deve soddisfare $X(g) = dl_g(X(e))$ per ogni elemento g del gruppo. In \mathbb{H}^n il differenziale della moltiplicazione a sinistra per un elemento (x, y, t) è la matrice $(2n+1) \times (2n+1)$ data da

$$dl_{(x,y,t)}(0) = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 2y^T & -2x^T & 1 \end{pmatrix}$$

dove I indica la matrice identità $n \times n$, 0 la matrice ($n \times n$ o $n \times 1$ a seconda dei casi) formata da soli zeri, mentre con l'apice T indichiamo al solito la trasposta di una matrice o di un vettore. A questo punto, indicando con $\partial_{x_1}, \dots, \partial_{x_n}, \partial_{y_1}, \dots, \partial_{y_n}, \partial_t$ la base canonica di \mathbb{R}^{2n+1} , una base dei campi invarianti a sinistra è data da

$$\begin{aligned} X_j(x, y, t) &= dl_{(x,y,t)}(\partial_{x_j}) = \partial_{x_j} + 2y_j \partial_t \\ Y_j(x, y, t) &= dl_{(x,y,t)}(\partial_{y_j}) = \partial_{y_j} - 2x_j \partial_t \\ T(x, y, t) &= dl_{(x,y,t)}(\partial_t) = \partial_t \end{aligned}$$

che sono i campi introdotti quando nella sezione 5.3 abbiamo calcolato le geodetiche del gruppo di Heisenberg. Notiamo infine che tali campi soddisfano le condizioni (6.3.2).

6.3.2 Campi invarianti a sinistra

Siano \mathbb{G} un gruppo di Carnot e $F : \mathbb{R}^n \rightarrow \mathbb{G}$ un sistema di coordinate graduate associato alla stratificazione $\mathcal{G} = V_1 \oplus \dots \oplus V_k$.

Definizione 6.12. *Una funzione $P : \mathbb{G} \rightarrow \mathbb{R}$ si dice **polinomio** su \mathbb{G} se la mappa composta $P \circ F$ risulta una funzione polinomiale su \mathbb{R}^n .*

Osservazione 6.1. *Osserviamo che la definizione è ben posta: se infatti G è un altro sistema di coordinate graduate allora $F^{-1} \circ G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è lineare (si tratta essenzialmente di un cambiamento di base in \mathcal{G}), e quindi se $P \circ F$ è una funzione polinomiale anche $P \circ G = (P \circ F) \circ (F^{-1} \circ G)$ lo sarà.*

Sia $p^j : \mathbb{R}^n \rightarrow \mathbb{R}$ la proiezione canonica sulla j -esima coordinata, e sia $\pi^j := p^j \circ F^{-1} : \mathbb{G} \rightarrow \mathbb{R}$. Infine dato un n -multiindice di interi non negativi $\alpha = (\alpha_1, \dots, \alpha_n)$ definiamo

$$\begin{aligned}\mathbb{R}^n \ni x &\xrightarrow{p^\alpha} \prod_{j=1}^n (p^j(x))^{\alpha_j} \in \mathbb{R} \\ \mathbb{G} \ni g &\xrightarrow{\pi^\alpha} \prod_{j=1}^n (\pi^j(g))^{\alpha_j} \in \mathbb{R}.\end{aligned}$$

Naturalmente ogni π^α è un polinomio su \mathbb{G} , e anzi ogni polinomio su \mathbb{G} si può scrivere come combinazione lineare finita di opportuni π^α . Diremo che il grado omogeneo di π^α è l'intero $\sum_{j=1}^n d_j \alpha_j$.

Definizione 6.13. *Il **grado omogeneo** $\deg_H(P)$ di un polinomio P è il più grande dei gradi omogenei dei π^α che compaiono nell'espressione minimale di P come combinazione lineare dei π^α .*

Ad esempio il polinomio $xy^2 - t^2$ nel gruppo di Heisenberg \mathbb{H}^1 ha grado omogeneo 4, in quanto il grado della coordinata t è 2.

Proposizione 6.1. *Il grado omogeneo di un polinomio P non dipende dalla scelta delle coordinate graduate.*

Dimostrazione. Siano $F : \mathbb{R}_x^n \rightarrow \mathbb{G}$ e $G : \mathbb{R}_y^n \rightarrow \mathbb{G}$ due sistemi di coordinate graduate associate alle basi X_1, \dots, X_n e Y_1, \dots, Y_n rispettivamente, basi che sono adattate alla stratificazione $\mathcal{G} = V_1 \oplus \dots \oplus V_k$. Abbiamo allora, sottintendendo le somme sugli indici quando sono ripetuti una volta “alzati” e una volta “abbassati”,

$$Y_j = A_j^l X_l.$$

In questo modo

$$F^{-1} \circ G(y) = A_j^l y^j e_l$$

dove e_1, \dots, e_n è la base canonica di \mathbb{R}_x^n . Inoltre, dal momento che si parla di basi adattate a una stratificazione, $A_j^l \neq 0$ solo se $m_{d_j-1} < l \leq m_{d_j}$: in questo modo la matrice $n \times n$ $A = (A_j^k)$ risulta della seguente forma

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_k \end{pmatrix} \quad (6.3.4)$$

dove A_j indica una opportuna matrice $n_j \times n_j$, mentre 0 indicano matrici della grandezza opportuna e con tutti i coefficienti nulli.

Per i nostri scopi è sufficiente dimostrare che i polinomi della forma $\pi^\alpha = p^\alpha \circ F^{-1}$, letti nelle coordinate G , mantengono invariato il grado omogeneo. Abbiamo

$$\pi^\alpha \circ G = (\pi^\alpha \circ F) \circ (F^{-1} \circ G) = \prod_{l=1}^n (A_j^l y^j)^{\alpha_l}.$$

Dato che A è diagonale della forma (6.3.4), il suo determinante sarà pari al prodotto dei determinanti delle matrici A_1, \dots, A_k ; dal momento che A è invertibile tutti questi determinanti saranno nonnulli. In particolare nessuna colonna delle A_j sarà completamente nulla, quindi per ogni $l \in \{1, \dots, n\}$ esiste j_l tale che $A_{j_l}^l \neq 0$. Ricordando poi che $A_j^l = 0$ se $d_j \neq d_l$ avremo

$$\deg_H (A_j^l y^j)^{\alpha_l} = d_l \alpha_l$$

dove però NON si fa la somma sugli indici l , in quanto sono entrambi abbassati. Infine

$$\begin{aligned} \deg_H(\pi^\alpha \circ F) &= \sum_{l=1}^n d_l \alpha_l = \sum_{l=1}^n \deg_H (A_j^l y^j)^{\alpha_l} \\ &= \deg_H \left(\prod_{l=1}^n (A_j^l y^j)^{\alpha_l} \right) = \deg_H(\pi^\alpha \circ G). \end{aligned}$$

□

Definizione 6.14. Un polinomio $P : \mathbb{G} \rightarrow \mathbb{R}$ si dice **omogeneo** di grado $\alpha > 0$ se $P(\delta_\lambda g) = \lambda^\alpha P(g)$ per ogni $g \in \mathbb{G}$ e ogni $\lambda > 0$.

Ad esempio il polinomio $xy^3 + t^2$ del gruppo di Heisenberg \mathbb{H}^1 è omogeneo di grado 4.

Nelle coordinate graduate la traslazione a sinistra per un elemento x si scrive

$$l_x(y) = F^{-1}(F(x)F(y)) = \sum_{j=1}^n C_j(x, y)e_j \quad (6.3.5)$$

dove le funzioni $C_j(x, y)$ sono polinomi che si ricavano dalla formula di Baker-Campbell-Hausdorff. Mostriamo che essi sono polinomi omogenei di grado d_j :

$$\begin{aligned} \sum_{j=1}^n \lambda^{d_j} C_j(x, y)e_j &= \delta_\lambda \left(\sum_{j=1}^n C_j(x, y)e_j \right) \\ &= \delta_\lambda (F^{-1}(F(x) \cdot F(y))) \\ &= F^{-1}(\delta_\lambda(F(x)) \cdot \delta_\lambda(F(y))) \\ &= F^{-1}(F(\delta_\lambda x) \cdot F(\delta_\lambda y)) = \sum_{j=1}^n C_j(\delta_\lambda x, \delta_\lambda y)e_j. \end{aligned}$$

Il confronto fra il primo e l'ultimo termine della catena di uguaglianze ci dice che $C_j(\delta_\lambda x, \delta_\lambda y) = \lambda^{d_j} C_j(x, y)$, ovvero che $C_j(x, y)$ è omogeneo di grado d_j .

Il nostro obiettivo è ora quello di dare una scrittura esplicita dei campi invarianti a sinistra letti attraverso coordinate graduate, ovvero dei campi vettoriali $\tilde{X}_j := F^* X_j$ su \mathbb{R}^n .

Utilizzando l'usuale identificazione fra campi e derivazioni otteniamo

$$\tilde{X}_j \varphi(x) = \partial_{x_j} (\varphi \circ l_x)(0) = \partial_{x_j} l_x^l(0) \partial_{x_l} \varphi(x)$$

per ogni $j = 1, \dots, n$, dove l_x^l indica la l -esima componente di l_x . Ancora sottintendiamo la somma sugli indici ripetuti, quando sono uno alzato e uno abbassato. Confrontando con la (6.3.5) si deduce

$$\partial_{x_j} l_x^l(0) = \partial_{x_j} C_l(x, \cdot)(0) =: a_{jl}(x).$$

Per l'omogeneità dei C_l abbiamo

$$\begin{aligned} \lambda^{d_l} a_{jl}(x) &= \lambda^{d_l} \partial_{x_j} C_l(x, y) \Big|_{y=0} = \partial_{x_j} C_l(\delta_\lambda x, \delta_\lambda y) \Big|_{y=0} \\ &= \lambda^{d_j} \partial_{x_j} C_l(x, y) \Big|_{y=0} = \lambda^{d_j} a_{jl}(\delta_\lambda x) \end{aligned}$$

e quindi i polinomi a_{jl} sono omogenei di grado $d_l - d_j$.

Questo implica che se $d_j > d_l$ allora $a_{jl} \equiv 0$. Osservando poi che i polinomi 0-omogenei sono solo le costanti otteniamo

$$\tilde{X}_j(x) = \sum_{d_l=d_j} c_{jl} \partial_{x_l} + \sum_{d_l>d_j} a_{jl}(x) \partial_{x_l}$$

dove c_{jl} sono costanti che possono essere calcolate notando che

$$\tilde{X}_j \varphi(0) = \frac{d}{dt} \varphi \circ F^{-1}(\exp t X_j) \Big|_{t=0} = \frac{d}{dt} \varphi(te_j) \Big|_{t=0} = \partial_{x_j} \varphi(0).$$

Questo ci dice che $c_{jl} = \delta_{jl}$ e quindi

$$\tilde{X}_j(x) = \partial_{x_j} + \sum_{l=m_{d_j+1}}^n a_{jl}(x) \partial_{x_l}. \quad (6.3.6)$$

Tale espressione dei campi invarianti a sinistra ci dice, fra l'altro, che essi sono linearmente indipendenti e di classe $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^n)$.

Dal momento che gli a_{jl} sono omogenei di grado $d_l - d_j$, le coordinate x_r con $d_r > d_l - d_j$ non possono apparire nell'espressione polinomiale di a_{jl} . Quindi gli a_{jl} non dipendono dalle x_r con $d_r \geq d_l$, cioè

$$a_{jl}(x) = a_{jl}(x_1, \dots, x_{m_{d_l-1}}).$$

In particolare avremo $a_{jl}(x) = a_{jl}(x_1, \dots, x_{l-1})$.

Si noti infine come tutto ciò risulti verificato nel gruppo di Heisenberg.

6.4 I gruppi di Carnot come spazi di Carnot-Carathéodory

Sia (\mathbb{R}^n, \cdot) un gruppo di Carnot con relativa algebra dotata della stratificazione $\mathcal{G} = V_1 \oplus \cdots \oplus V_k$. Sia $m = m_1 = \dim V_1$ e sia X_1, \dots, X_m una base di V_1 : essa verifica la condizione di Chow-Hörmander e quindi induce una metrica di Carnot-Carathéodory su \mathbb{R}^n . Inoltre tale condizione è verificata anche considerando i soli commutatori di lunghezza minore o uguale a k : la proposizione 4.3 ci dice allora che per ogni compatto $K \subset \mathbb{R}^n$ esiste una costante $C = C(K)$ tale che $d_{CC}(x, y) \leq |x - y|^{1/k}$ per ogni $x, y \in K$.

Tale distanza si comporta molto bene rispetto alle traslazioni a sinistra nonché alle dilatazioni del gruppo, come mostriamo nella seguente

Proposizione 6.2. *Per ogni $x, y, z \in \mathbb{R}^n$ ed ogni $\lambda \in \mathbb{R}$ si ha*

1. $d_{CC}(z \cdot x, z \cdot y) = d_{CC}(x, y)$;
2. $d_{CC}(\delta_\lambda x, \delta_\lambda y) = |\lambda| d_{CC}(x, y)$.

Dimostrazione. La parte 1. segue dal fatto che $\gamma : [0, T] \rightarrow \mathbb{R}^n$ è una curva subunitaria che congiunge x e y se e solo se $\tilde{\gamma} := l_z \circ \gamma$ è una curva subunitaria definita su $[0, T]$ che congiunge zx e zy . Infatti se $\dot{\gamma} = \mathcal{A}(\gamma)h$ allora

$$\begin{aligned} \dot{\tilde{\gamma}}(t) &= dl_z(\gamma(t))[\dot{\gamma}(t)] \\ &= dl_z(\gamma(t)) \left[\sum_{j=1}^m h_j(t) X_j(\gamma(t)) \right] \\ &= \sum_{j=1}^m h_j(t) dl_z(\gamma(t)) X_j(\gamma(t)) = \sum_{j=1}^m h_j(t) X_j(z \cdot \gamma(t)) \\ &= \sum_{j=1}^m h_j(t) X_j(\tilde{\gamma}(t)) \end{aligned}$$

da cui la tesi.

Per quanto riguarda il punto 2., basterà dimostrare l'enunciato nel caso $\lambda > 0$. Nello specifico sarà sufficiente mostrare che una curva $\gamma : [0, T] \rightarrow \mathbb{R}^n$ da x a y è subunitaria se e solo se è subunitaria la curva $\gamma_\lambda : [0, \lambda T] \rightarrow \mathbb{R}^n$ che unisce $\delta_\lambda x$ e $\delta_\lambda y$ ed è definita da $\gamma_\lambda(t) := \delta_\lambda(\gamma(t/\lambda))$. Abbiamo

$$\dot{\gamma}(t) = \sum_{j=1}^m h_j(t) X_j(\gamma(t)) = \sum_{l=1}^n \left(\sum_{j=1}^m h_j(t) a_{jl}(\gamma(t)) \right) \partial_{x_l}.$$

Dato che per tutti i j che compaiono nella sommatoria si ha $d_j = 1$ avremo che tutti gli a_{jl} interessati sono $d_l - 1$ omogenei e quindi

$$\begin{aligned}
\dot{\gamma}_\lambda(t) &= \sum_{l=1}^n \lambda^{d_l-1} \left(\sum_{j=1}^n h_j(t/\lambda) a_{jl}(\gamma(t/\lambda)) \right) \partial_{x_l} \\
&= \sum_{l=1}^n \left(\sum_{j=1}^m h_j(t/\lambda) a_{jl}(\gamma_\lambda(t)) \right) \partial_{x_l} \\
&= \sum_{j=1}^m h_j(t/\lambda) X_j(\gamma_\lambda(t)). \tag{6.4.1}
\end{aligned}$$

La relazione (6.4.1) ci dice che γ è subunitaria se e solo se γ_λ lo è, e da questo segue la parte 2. della proposizione. \square

Corollario 6.1. *Se γ è una geodetica fra due punti x e y di un gruppo di Carnot \mathbb{G} , allora per ogni $\lambda \in \mathbb{R}$ e ogni $z \in \mathbb{G}$ la curva $z \cdot \delta_\lambda(\gamma)$ è una geodetica fra i punti $z \cdot \delta_\lambda x$ e $z \cdot \delta_\lambda y$.*

La proposizione precedente ci consente di studiare i problemi connessi alla distanza di Carnot-Carathéodory (ad esempio l'andamento asintotico della distanza o la regolarità delle geodetiche) a partire da un punto iniziale a nostra scelta, che tipicamente sarà l'origine. Ad esempio è immediato verificare che, indicando con e_j i vettori della base canonica di \mathbb{R}^n , si ha

$$d_{CC}(0, te_j) = Ct^{1/d_j}.$$

dove C è una costante pari a $d_{CC}(0, e_j)$. A questo punto per ogni $x \in \mathbb{R}^n$ si ha

$$d_{CC}(x, x \cdot te_j) = Ct^{1/d_j}.$$

Definita la **dimensione omogenea** di (\mathbb{R}^n, \cdot) come l'intero $Q := \sum_{j=1}^k j \dim V_j$ è semplice verificare che per ogni $x \in \mathbb{R}^n$ e ogni insieme $E \subset \mathbb{R}^n$ \mathcal{L}^n -misurabile si ha

- $\mathcal{L}^n(x \cdot E) = \mathcal{L}^n(E \cdot x) = \mathcal{L}^n(E)$;
- $\mathcal{L}^n(\delta_\lambda E) = \lambda^Q \mathcal{L}^n(E)$;
- (\mathbb{R}^n, \cdot) ha dimensione di Hausdorff Q .

6.5 Il teorema di Pansu

Siano (\mathbb{G}, \cdot) e $(\overline{\mathbb{G}}, \bar{\cdot})$ due gruppi di Carnot; indicheremo con d e con \bar{d} (rispettivamente) le distanze indotte dalla struttura di Carnot-Carathéodory canonica, come nella sezione 6.4, e con δ_λ e $\bar{\delta}_\lambda$ le dilatazioni di gruppo su ciascuno di essi.

Definizione 6.15. Una mappa $\varphi : \mathbb{G} \rightarrow \overline{\mathbb{G}}$ è un **omomorfismo omogeneo** se è un omomorfismo di gruppi tale che $\varphi \circ \delta_\lambda = \bar{\delta}_\lambda \circ \varphi$.

Definizione 6.16. Una mappa $f : \mathbb{G} \rightarrow \overline{\mathbb{G}}$ è **Pansu-differenziabile** in $x \in \mathbb{G}$ se esiste un omomorfismo omogeneo $L : \mathbb{G} \rightarrow \overline{\mathbb{G}}$ tale che

$$\lim_{t \rightarrow 0} \frac{\bar{d}(f(x)^{-1} \bar{\cdot} f(x \cdot t), L(t))}{d(0, t)} = 0.$$

Chiameremo tale L **differenziale** di f in x e lo indicheremo con $Df(x)$. Si ha ovviamente $Df(x)(\xi) = \lim_{h \downarrow 0} \bar{\delta}_{1/h}(f(x)^{-1} \bar{\cdot} f(x \cdot \delta_h \xi))$ per ogni $\xi \in \mathbb{G}$.

Il seguente teorema estende il classico risultato di Rademacher sulla differenziabilità delle funzioni Lipschitziane; per una sua dimostrazione si vedano [16] e [19]. Con \mathcal{L}^n indicheremo la misura di Lebesgue n -dimensionale derivante dall'identificazione di \mathbb{G} con \mathbb{R}^n tramite le coordinate graduate $F : \mathbb{R}^n \rightarrow \mathbb{G}$: per comodità eviteremo la scrittura $F_* \mathcal{L}^n$.

Teorema 6.2 (Pansu). Sia $f : \mathbb{G} \rightarrow \overline{\mathbb{G}}$ una mappa Lipschitziana. Allora esiste $Df(x)$ per \mathcal{L}^n -quasi ogni $x \in \mathbb{G}$.

Diamo la dimostrazione del teorema 6.2 solo nel caso unidimensionale, in cui \mathbb{G} è \mathbb{R} con la struttura di gruppo canonica e in cui identificheremo $\overline{\mathbb{G}}$ con un certo \mathbb{R}^n tramite coordinate esponenziali; attraverso queste coordinate i campi invarianti a sinistra del primo strato orizzontale (quelli di V_1 per intenderci) assumono la forma

$$X_j(x) = \partial_{x_j} + \sum_{l=m+1}^n a_{jl}(x) \partial_{x_l} \quad j = 1, \dots, m. \quad (6.5.1)$$

Per comodità di notazione eviteremo le soprilineature quando ci riferiremo ad oggetti relativi al gruppo di arrivo.

Dimostrazione (del caso unidimensionale del teorema di Pansu). Sia $f : \mathbb{R} \rightarrow \mathbb{R}^n$ la nostra funzione: essendo essa Lipschitziana rispetto alla metrica di Carnot-Carathéodory, sarà Lipschitziana anche in senso euclideo e dunque il teorema di Rademacher implica che per quasi ogni $t \in \mathbb{R}$ esiste

$$\dot{f}(t) = \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h} = \sum_{j=1}^m a_j(t) X_j(f(t)). \quad (6.5.2)$$

In realtà Rademacher ci dice solo che la derivata di f esiste quasi ovunque; tuttavia la Lipschitzianità di f implica che \dot{f} è quasi ovunque orizzontale, da cui la sua forma particolare (6.5.2).

Inoltre il successivo teorema di derivabilità metrica 6.3 ci dice che per quasi ogni t si ha

$$\lim_{h \rightarrow 0} \frac{d_{CC}(f(t+h), f(t))}{|h|} = \sqrt{\sum a_j(t)^2}. \quad (6.5.3)$$

Dimostriamo che per quei t per cui valgono la (6.5.2) e la (6.5.3) (e quindi per quasi ogni t) allora f è Pansu-differenziabile in t ; in particolare non sarà restrittivo supporre $t = 0$ e $f(t) = 0$. Quello che vogliamo allora dimostrare è che

$$\lim_{h \rightarrow 0} \frac{d_{CC}(f(h), L(h))}{|h|} = 0 \quad (6.5.4)$$

dove $L : \mathbb{R} \rightarrow \mathbb{R}^n$ è l'omomorfismo omogeneo $L(h) := (a_1(0)h, \dots, a_m(0)h, 0, \dots, 0)$. Il primo membro della (6.5.4) si può riscrivere come $d_{CC}(\delta_{1/h}f(h), (a_1(0), \dots, a_m(0), 0, \dots, 0))$ e quindi basta dimostrare che $\delta_{1/h}f(h) \rightarrow (a_1(0), \dots, a_m(0), 0, \dots, 0)$ per $h \rightarrow 0$. Dato che f è Lipschitziana $\delta_{1/h}f(h)$ deve essere limitata, e quindi possiamo supporre che una sottosuccessione $\delta_{1/h_i}f(h_i)$ converga per $i \rightarrow \infty$ ad un certo punto $x = (x_1, \dots, x_n)$. Questo vuol dire che $f_j(h_i)/h_i^{d_j} \xrightarrow{i \rightarrow \infty} x_j$ per ogni $j = 1, \dots, n$ (f_j indica la j -esima componente di f): per via di (6.5.2) dovremo dunque avere $x_j = a_j(0)$ per ogni $j = 1, \dots, m$, dal momento che $d_j = 1$ per tali j . Quello che vogliamo dimostrare è che il punto limite non dipende dalla particolare successione h_i .

Riscrivendo la (6.5.3) otteniamo allora

$$\begin{aligned} \sqrt{\sum a_j(0)^2} &= \lim_{h \rightarrow 0} \frac{d_{CC}(f(h), 0)}{|h|} = \lim_{h \rightarrow 0} d_{CC}(\delta_{1/h}f(h), 0) \\ &= d_{CC}\left((a_1(0), \dots, a_m(0), x_{m+1}, \dots, x_n), 0\right); \end{aligned}$$

sarà dunque sufficiente dimostrare che se $x = (a_1(0), \dots, a_m(0), x_{m+1}, \dots, x_n)$ ha distanza dall'origine pari a $\sqrt{\sum a_j(0)^2}$, allora $x_{m+1} = \dots = x_n = 0$. Sia dunque $\gamma : [0, T] \rightarrow \mathbb{R}^n$ una geodetica minimizzante parametrizzata per lunghezza d'arco e tale che $\gamma(0) = 0$ e $\gamma(T) = x$. Avremo che $\dot{\gamma}(s) = \sum_{j=1}^m h_j(s)X_j(\gamma(x))$ per certe h_j misurabili; dato che i campi X_1, \dots, X_m assumono la forma particolare (6.5.1) avremo $\int_0^T h_j = a_j(0)$ e quindi

$$\begin{aligned} \sum_{j=1}^m a_j(0)^2 &= \sum_{j=1}^m \left(\int_0^T h_j \right)^2 & (*) \\ &\leq \sum_{j=1}^m T \left(\int_0^T h_j^2 \right) = T \int_0^T \sum_{j=1}^m h_j^2 = T^2 \end{aligned}$$

dove nel passaggio contrassegnato da (*) abbiamo sfruttato la disuguaglianza di Holder. Nella disuguaglianza il segno di uguale può valere solo se ciascuna h_j è quasi ovunque costante, dunque solo se $h_j \equiv a_j(0)/T$: questo implica facilmente che $x = \exp(\sum_{j=1}^m a_j(0)X_j) = (a_1(0), \dots, a_m(0), 0, \dots, 0)$. D'altra parte che $(a_1(0), \dots, a_m(0), 0, \dots, 0)$ abbia distanza dall'origine pari a $\sqrt{\sum_{j=1}^m a_j(0)^2}$ si evince dal fatto che la curva $t \mapsto \exp(t \sum a_j(0)X_j)$ congiunge x e l'origine ed ha lunghezza orizzontale $\sqrt{\sum a_j(0)^2}$.

Dunque $\delta_{1/h}f(h) \rightarrow (a_1(0), \dots, a_m(0), 0, \dots, 0)$ e la dimostrazione è completa. \square

Teorema 6.3 (di derivabilità metrica). *Sia (S, d) uno spazio metrico e sia $f :]a, b[\rightarrow S$ una funzione per cui esiste $m \in L^1(a, b)$ tale che*

$$d(f(s), f(t)) \leq \int_s^t m(\tau) d\tau \quad \text{per ogni } a \leq s \leq t \leq b. \quad (6.5.5)$$

Allora per quasi ogni $t \in]a, b[$ esiste

$$|f'(t)| := \lim_{s \rightarrow t} \frac{d(f(s), f(t))}{|s - t|}.$$

Inoltre $|f'|$ appartiene a $L^1(a, b)$, soddisfa la (6.5.5) se sostituitavi al posto di m , ed è tale che, per quasi ogni t , $|f'(t)| \leq m(t)$ per ogni funzione $m \in L^1(a, b)$ che soddisfa la (6.5.5).

Dimostrazione. Sia $\{y_n\} \subset S$ denso in $f(]a, b[)$, e sia $\varphi_n(t) := d(y_n, f(t))$. Dato che tutte le φ_n sono assolutamente continue (infatti $|\varphi_n(t) - \varphi_n(s)| = |d(y_n, f(t)) - d(y_n, f(s))| \leq d(f(t), f(s)) \leq \int_s^t m(\tau) d\tau$ per una certa $m \in L^1$), allora per quasi ogni t è ben definita la funzione

$$\lambda(t) := \sup_n |\dot{\varphi}_n(t)|.$$

Sia dunque t un punto in cui tutte le φ_n sono differenziabili; notiamo che

$$\liminf_{s \rightarrow t} \frac{d(f(s), f(t))}{|s - t|} \geq \sup_n \liminf_{s \rightarrow t} \frac{|\varphi_n(s) - \varphi_n(t)|}{|s - t|} = \lambda(t).$$

Questo, assieme alla (6.5.5), ci consente di concludere che, per ogni m che soddisfa la (6.5.5), si ha $\lambda \leq m$ nei punti di Lebesgue di m e in particolare quasi ovunque; estendendo λ a tutto $]a, b[$ in modo che risulti misurabile, abbiamo $\lambda \in L^1(a, b)$. D'altra parte abbiamo

$$d(f(s), f(t)) = \sup_n |\varphi_n(s) - \varphi_n(t)| \leq \int_s^t \lambda(\tau) d\tau$$

per ogni $a \leq s \leq t \leq b$, e quindi

$$\limsup_{s \rightarrow t} \frac{d(f(s), f(t))}{|s - t|} \leq \lambda(t)$$

nei punti di Lebesgue di λ , e quindi quasi ovunque. Questo implica che $|\dot{f}| = \lambda$ quasi ovunque, e la dimostrazione è completa. \square

Capitolo 7

Regolarità delle geodetiche nei Gruppi di Carnot

In questo capitolo vogliamo dare condizioni sufficienti a garantire la regolarità delle geodetiche sub-riemanniane nei gruppi di Carnot. Nella prima sezione spiegheremo cosa intendiamo per regolarità di una curva a valori in un gruppo di Carnot, mentre nella seconda dimostreremo che, dato un campo X appartenente al primo strato V_1 , la curva $t \mapsto \exp(tX)$ è, a meno di riparametrizzazione, l'unica geodetica fra un suo punto qualsiasi e l'identità del gruppo: in particolare essa risulterà anche minimizzante. Nella sezione 3 enunciamo una proprietà, che chiamiamo $(P1)$, che garantisce la regolarità delle geodetiche all'infuori di un chiuso di misura nulla: è evidente l'analogia col teorema di Tonelli discusso nel capitolo 2. La $(P1)$ richiede che esista unica e regolare una geodetica fra l'identità e un intorno dei punti dello strato orizzontale a distanza unitaria da essa. Nella quarta e ultima sezione dimostriamo che è garantita la regolarità globale quando vale la proprietà che chiamiamo $(P2)$, ovvero che le uniche geodetiche globalmente minimizzate sono solo quelle date dall'esponenziale di campi del primo strato.

7.1 Regolarità per curve orizzontali

Identificando un gruppo di Carnot \mathbb{G} con un certo (\mathbb{R}^n, \cdot) , per regolarità di una funzione intendiamo la regolarità della stessa rispetto alla struttura euclidea di \mathbb{R}^n . In questo senso potrà tornare utile il seguente

Lemma 7.1. *Siano $X_1, \dots, X_m \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ tali che per ogni $x \in \mathbb{R}^n$ i vettori $X_1(x), \dots, X_m(x)$ siano linearmente indipendenti. Sia poi $\gamma : [a, b] \rightarrow \mathbb{R}^n$ una curva derivabile e tale che*

$$\dot{\gamma}(t) = \sum_{j=1}^m h_j(t) X_j(\gamma(t)).$$

Allora $\gamma \in \mathcal{C}^k([a, b], \mathbb{R}^n) \iff h_j \in \mathcal{C}^{k-1}([a, b], \mathbb{R})$ per ogni $j = 1, \dots, m$.

Dimostrazione. L'implicazione \Rightarrow è del tutto banale; per dimostrare l'altra procediamo per induzione. Per comodità di notazione scriveremo $X_j(t)$ in luogo di $X_j(\gamma(t))$.

Supponiamo dunque $\gamma \in \mathcal{C}^1$; per ottenere la continuità delle h_j è sufficiente dimostrare che per ogni $t \in [a, b]$ e per ogni successione $\{t_n\}_n$ tale che $t_n \rightarrow t$ e $h_j(t_n) \rightarrow a_j \in \mathbb{R} \cup \{\pm\infty\}$, allora $a_j = h_j(t)$. Innanzitutto possiamo escludere il caso $|a_j| = \infty$, in quanto $\dot{\gamma}$ è continua e dunque limitata. Indicando con $o(1)$ una quantità (reale o vettoriale a seconda dei casi) che tende a 0 per $n \rightarrow \infty$ abbiamo poi

$$\dot{\gamma}(t_n) = \sum_{j=1}^m h_j(t_n) X_j(t_n) = \sum_{j=1}^m (a_j + o(1))(X_j(t) + o(1))$$

Ma, dato che $\gamma \in \mathcal{C}^1$, il primo membro convergerà a $\dot{\gamma}(t) = \sum_{j=1}^m h_j(t) X_j(t)$; mandando al limite l'ultimo membro e uguagliando otteniamo

$$\sum_{j=1}^m (a_j - h_j(t)) X_j(t) = 0.$$

Vista la lineare indipendenza dei vettori $X_j(t)$ otteniamo $a_j = h_j(t)$ per ogni j : quindi il lemma è provato nel caso $k = 1$.

Per quanto riguarda il passo induttivo, sia $\gamma \in \mathcal{C}^{k+1}$: quel che vogliamo dimostrare è che $h_j \in \mathcal{C}^k$. L'ipotesi induttiva ci dice innanzitutto che $h_j \in \mathcal{C}^{k-1}$: iniziamo col dimostrare che $h_j^{(k-1)}$ è derivabile, con derivata a valori in $\mathbb{R} \cup \{\pm\infty\}$. A tale scopo sarà sufficiente dimostrare che date due qualsiasi successioni $\{t_n\}_n, \{s_n\}_n$ tali che

- $t_n \rightarrow t$ e $s_n \rightarrow t$ per $n \rightarrow \infty$;
- $\frac{h_j^{(k-1)}(t_n) - h_j^{(k-1)}(t)}{t_n - t} \rightarrow a_j \in \mathbb{R} \cup \{\pm\infty\}$;
- $\frac{h_j^{(k-1)}(s_n) - h_j^{(k-1)}(t)}{s_n - t} \rightarrow b_j \in \mathbb{R} \cup \{\pm\infty\}$

allora $a_j = b_j$ per ogni j . Per comodità supponiamo $t = 0$; usando più volte la formula di derivazione di un prodotto di Leibnitz e tenendo conto che $h_j \in \mathcal{C}^{k-1}$, otteniamo che per certe costanti $C_{l,k}$ si ha

$$\begin{aligned}
\frac{\gamma^{(k)}(t_n) - \gamma^{(k)}(0)}{t_n} &= \frac{1}{t_n} \sum_{j=1}^m \left[\sum_{l=0}^{k-2} \left(C_{l,k} h_j^{(l)}(t_n) X_j^{(k-1-l)}(t_n) \right) + h_j^{(k-1)}(t_n) X_j(t_n) + \right. \\
&\quad \left. - \sum_{l=0}^{k-2} \left(C_{l,k} h_j^{(l)}(0) X_j^{(k-1-l)}(0) \right) - h_j^{(k-1)}(0) X_j(0) \right] \\
&= \sum_{j=1}^m \left[\sum_{l=0}^{k-2} \left(C_{l,k} \frac{h_j^{(l)}(t_n) X_j^{(k-1-l)}(t_n) - h_j^{(l)}(0) X_j^{(k-1-l)}(0)}{t_n} \right) + \right. \\
&\quad \left. + \frac{h_j^{(k-1)}(t_n) [X_j(0) + t_n \dot{X}_j(0) + o(|t_n|)] - h_j^{(k-1)}(0) X_j(0)}{t_n} \right] = \\
&= \sum_{j=1}^m \sum_{l=0}^{k-2} C_{l,k} \left(\frac{d h_j^{(l)} X_j^{(k-1-l)}}{dt} (0) + o(1) \right) + \\
&\quad + \sum_{j=1}^m \left(h_j^{(k-1)}(0) \dot{X}_j(0) + a_j X_j(0) + o(1) \right). \tag{7.1.1}
\end{aligned}$$

Analogamente avremo

$$\begin{aligned}
\frac{\gamma^{(k)}(s_n) - \gamma^{(k)}(0)}{s_n} &= \sum_{j=1}^m \sum_{l=0}^{k-2} C_{l,k} \left(\frac{d h_j^{(l)} X_j^{(k-1-l)}}{dt} (0) + o(1) \right) + \\
&\quad + \sum_{j=1}^m \left(h_j^{(k-1)}(0) \dot{X}_j(0) + b_j X_j(0) + o(1) \right). \tag{7.1.2}
\end{aligned}$$

Ma i primi membri della (7.1.1) e della (7.1.2) convergono entrambi a $\gamma^{(k+1)}(0)$: questo ci dice che $|a_j| \neq \infty$ e $|b_j| \neq \infty$. Mandando poi al limite gli ultimi membri e uguagliando otteniamo $\sum_{j=1}^m (a_j - b_j) X_j(0) = 0$, e quindi $a_j = b_j$. A questo punto derivando k volte la $\dot{\gamma}(t) = \sum_{j=1}^m h_j(t) X_j(\gamma(t))$ si ottiene

$$\gamma^{(k+1)} - \sum_{j=1}^m \sum_{l=0}^{k-1} \left(C_{l,k} h_j^{(j)} X_j^{(k-l)} \right) = \sum_{j=1}^m h_j^{(k)} X_j$$

ma il primo membro è continuo, quindi lo sarà anche il secondo. Questo implica che le $h_j^{(k)}$ sono continue, e dunque il lemma è dimostrato. \square

Osservazione 7.1. Osserviamo che il lemma precedente resta valido se i campi X_1, \dots, X_m sono di classe C^{k-1} .

Essendo interessati alle curve orizzontali, i campi X_1, \dots, X_m saranno naturalmente quelli formati da una base del primo strato orizzontale della stratificazione di \mathcal{G} . Abbiamo già osservato nella sottosezione 6.3.2 come essi siano linearmente indipendenti e di classe $C^\infty(\mathbb{R}^n, \mathbb{R}^n)$. Ovviamente tale lemma può tornare utile in un contesto più generale, in cui i campi X_1, \dots, X_m formano una base dei campi orizzontali di una varietà sub-riemanniana generica.

7.2 Una proprietà dei gruppi di Carnot

Prima di affrontare il problema della regolarità delle geodetiche nei Gruppi di Carnot, cerchiamo di studiare alcune proprietà di alcune di esse, e precisamente quelle che sono date dal flusso dei campi invarianti a sinistra. Iniziamo con la seguente

Definizione 7.1. Una **calibrazione** θ su una varietà sub-riemanniana M è una 1-forma chiusa tale che $|\theta(v)| \leq |v|_H$ per quasi ogni $v \in TM$.

Diremo poi che una curva orizzontale γ è **calibrata** da θ se θ è una calibrazione tale che $\theta(\dot{\gamma}(t)) = |\dot{\gamma}(t)|_H$ per quasi ogni t nell'intervallo di definizione di γ . Per l'applicazione alla geometria sub-riemanniana di tali strumenti, che sono piuttosto classici e trovano utilizzo in varie discipline della Matematica, si veda il libro di Montgomery [18].

Proposizione 7.1. Sia γ una curva orizzontale in un gruppo di Carnot \mathbb{G} calibrata da θ ; allora γ è una geodetica minimizzante.

Dimostrazione. Sia γ' una qualsiasi altra curva orizzontale con gli stessi estremi di γ . Essendo \mathbb{G} connesso e semplicemente connesso avremo che γ e γ' sono omotope, ed essendo θ chiusa la formula di Stokes ci dice che $\int_{\gamma'} \theta = \int_{\gamma} \theta$, e quindi

$$l_H(\gamma') = \int |\dot{\gamma}'|_H \geq \int_{\gamma'} \theta = \int_{\gamma} \theta = \int |\dot{\gamma}|_H = l_H(\gamma).$$

□

In generale l'uso delle calibrazioni permette di ricavare proprietà delle geodetiche per tutte le varietà sub-riemanniane: ad esempio

Proposizione 7.2. Sia γ una curva orizzontale calibrata; allora ogni arco sufficientemente breve di γ è una geodetica minimizzante.

o anche

Proposizione 7.3. *Sia γ un arco sufficientemente breve di una geodetica normale sub-riemanniana; allora ogni curva minimizzante con gli stessi estremi di γ coincide con γ a meno di riparametrizzazioni.*

Per le dimostrazioni di tali proposizioni si veda [18]. Tuttavia il nostro ambiente di lavoro, i gruppi di Carnot, ci permette di fare a meno di tali proprietà, in quanto la loro struttura semplifica notevolmente le dimostrazioni che faremo.

Scopo di questa sezione è dimostrare la seguente proprietà dei gruppi di Carnot, dovuta a V. Magnani.

Proposizione 7.4. *Sia \mathbb{G} un gruppo di Carnot con relativa algebra stratificata $\mathcal{G} = V_1 \oplus \cdots \oplus V_k$, e sia $X \in V_1$. Allora la curva $t \mapsto \exp(tX)$ è, a meno di riparametrizzazioni, l'unica geodetica fra l'identità e il punto $\exp(X)$. In particolare, $d_{CC}(e, \exp(X)) = |X|_H$.*

Dimostrazione. Ricordando il corollario 6.1 e il fatto che $\exp(\lambda X) = \delta_\lambda(\exp(X))$, possiamo senza perdere in generalità supporre che $|X|_H = 1$.

Completiamo X ad una base $X_1 = X, X_2, \dots, X_n$ di \mathcal{G} adattata alla stratificazione, e identifichiamo \mathbb{G} con (\mathbb{R}^n, \cdot) tramite il sistema di coordinate esponenziali $F : \mathbb{R}^n \rightarrow \mathbb{G}$ relativo a tale base. Avevamo già trovato una base dei campi vettoriali invarianti a sinistra, ed essa era della forma (6.3.6):

$$X_j(x) = \partial_{x_j} + \sum_{l=m_{d_j}+1}^n a_{jl}(x) \partial_{x_l}.$$

In particolare avremo che la matrice $A(x) = \text{col}[X_1, \dots, X_n]$ è triangolare inferiore e con diagonale formata da soli 1; per come è stata costruita avremo che $X_j(x) = A(x)_j^l \partial_{x_l}$.

Per ogni $j = 1, \dots, n$ consideriamo le 1-forme X_j^* definite tramite

$$X_j^*(x)[X_l(x)] = \delta_{jl}.$$

Allo stesso modo, data la base $\partial_{x_1}, \dots, \partial_{x_n}$ dei vettori di \mathbb{R}^n possiamo definire le uno-forme $\partial_{x_j}^*$ tali che

$$\partial_{x_j}^*(x)[\partial_{x_l}(x)] = \delta_{jl}.$$

É altresì chiaro che $\partial_{x_j}^* = dx_j$. Il successivo lemma 7.2 ci dice che

$$X_j^*(x) = B(x)_j^l dx_l$$

dove $B(x) = (A(x)^T)^{-1}$. Ma la forma particolare di $A(x)$ ci dice che $B(x)$ sarà triangolare superiore con diagonale formata da soli 1; in particolare

$$X_1^*(x) = dx_1$$

è una uno-forma chiusa. Vogliamo dimostrare che X_1^* calibra la curva

$$\begin{aligned} \gamma &: [0, 1] \rightarrow \mathbb{R}^n \\ t &\longmapsto \exp(tX) = (t, 0, \dots, 0). \end{aligned}$$

La prima cosa da fare è dimostrare che X_1^* è una calibrazione. Sia dunque $v \in T_x \mathbb{R}^n$; se v non è orizzontale (in x) allora $X_1^*(x)[v] \leq |v|_H = +\infty$. Se invece v è orizzontale allora $v = \sum_1^m v_j X_j(x)$ e quindi $X_1^*(x)[v] = v_1 \leq \sum_1^m v_j^2 = |v|_H$; notiamo che vale il segno di uguale se e solo se v è un multiplo di $X_1(x)$. Resta da dimostrare che $X_1^*(x)[\dot{\gamma}(t)] = |\dot{\gamma}(t)|_H$, ma questo è vero perchè $\dot{\gamma}(t)$ è un multiplo di $X_1(\gamma(t))$. Dunque X_1^* calibra γ , e quindi γ è una geodetica minimizzante fra 0 e il punto $(1, 0, \dots, 0)$.

Rimane solo da dimostrare che ogni altra curva orizzontale $\gamma' : [0, 1] \rightarrow \mathbb{R}^n$ da 0 a $(1, 0, \dots, 0)$ di lunghezza minima è una riparametrizzazione di γ . Abbiamo

$$l_H(\gamma') = \int_{\gamma'} |\dot{\gamma}'|_H \geq \int_{\gamma'} \theta = \int_{\gamma} \theta = \int |\dot{\gamma}|_H = l_H(\gamma) = l_H(\gamma')$$

e quindi $|\dot{\gamma}'(t)|_H = X_1^*(\gamma'(t))[\dot{\gamma}'(t)]$ per quasi ogni t . Questo implica che $\dot{\gamma}'(t)$ è un multiplo di $X_1(\gamma'(t))$ per quasi ogni t , ma questo è possibile solo se γ' è una riparametrizzazione di γ . \square

Lemma 7.2. *Supponiamo di avere dei campi vettoriali $\{R_j\}_{j=1}^n$ e $\{S_j\}_{j=1}^n$ su una varietà n -dimensionale M tali che, per ogni $x \in M$, $R_1(x), \dots, R_n(x)$ siano una base di $T_x M$, e così pure per S_1, \dots, S_n . Siano R_j^* e S_j^* le uno-forme definite da*

$$R_j^*(x)[R_l(x)] = \delta_{jl}$$

$$S_j^*(x)[S_l(x)] = \delta_{jl}.$$

Se $S_j(x) = A(x)^l_j R_l(x)$, allora $S_j^* = B_j^l R_l^*$, dove $B(x) = (A(x)^T)^{-1}$.

Dimostrazione. Sicuramente esiste una funzione $B : M \rightarrow \mathcal{M}_{n,n}$ tale che $S_j^* = B_j^l R_l^*$; a questo punto avremo

$$\begin{aligned} \delta_{jl} &= S_j^*[S_l] = B_j^m R_m^*[A_l^n R_n] \\ &= \delta_{mn} B_j^m A_l^n = \sum_m B_j^m A_l^m \\ &= (B^T)_m^j A_l^m \end{aligned}$$

cioè $B^T = A^{-1}$, da cui la tesi. \square

Osservazione 7.2. Osserviamo che in realtà una dimostrazione più semplice della proposizione 7.4 era già stata data, implicitamente, nella dimostrazione del caso unidimensionale del teorema di Pansu 6.2. Ripercorriamola: sia $x = (x_1, \dots, x_m, 0, \dots, 0) = \exp(X)$, e sia $\gamma : [0, T] \rightarrow \mathbb{R}^n$ una geodetica minimale da 0 a x parametrizzata per lunghezza d'arco (ancora identifichiamo \mathbb{G} e \mathbb{R}^n tramite coordinate esponenziali). Avremo $\dot{\gamma}(t) = \sum_1^m h_j(t)X_j(t)$; sfruttando la forma particolare dei campi invarianti a sinistra X_1, \dots, X_m è facile dedurre che $\int_0^T h_j = x_j$ per ogni $j = 1$. Utilizzando la disuguaglianza di Holder otteniamo

$$\begin{aligned} l_H(\exp(tX)|_{[0,1]})^2 &= \sum_{j=1}^m x_j^2 = \sum_{j=1}^m \left(\int_0^T h_j \right)^2 \\ &\leq \sum_{j=1}^m T \left(\int_0^T h_j^2 \right) = T^2 \end{aligned}$$

Abbiamo dunque ottenuto che la curva $t \mapsto \exp(tX)$ è minimale in $[0, 1]$. Per dimostrare che è anche l'unica geodetica minimale notiamo che il segno di uguale, nella disuguaglianza di Holder, vale solo se h_j è quasi ovunque uguale alla costante $x_j / \sum_1^m x_j^2$: questo implica facilmente che γ è una riparametrizzazione della geodetica $t \mapsto \exp(tX)$.

7.3 Regolarità delle geodetiche: un'estensione del teorema di Tonelli ai Gruppi di Carnot

In questa sezione cerchiamo di estendere il teorema di regolarità parziale di Tonelli 2.1 al problema delle geodetiche nei Gruppi di Carnot; in particolare daremo una condizione che, se soddisfatta, garantisce la regolarità delle geodetiche al di fuori di un chiuso di misura nulla. Nel nostro caso, il ruolo dell'ipotesi $L_{vv} > 0$ viene ricoperto dalla stretta convessità della norma orizzontale.

Teorema 7.1. Sia \mathbb{G} un gruppo di Carnot con algebra stratificata $\mathcal{G} = V_1 \oplus \dots \oplus V_k$, e sia $H := \{x \in \mathbb{G} : x = \exp(v), v \in V_1 \text{ e } |v|_H = 1\}$. Supponiamo inoltre che valga la seguente proprietà

(P1) Esiste $r > 0$ tale che, per ogni $y \in \mathbb{G}$ con $d_{CC}(y, H) < r$, esiste un'unica geodetica di classe C^q tra l'identità e y .

Sia $\gamma : [0, T] \rightarrow \mathbb{G}$ una geodetica minimizzante parametrizzata per lunghezza d'arco; allora esiste un chiuso di misura nulla $\Sigma \subset [0, T]$ tale che $\gamma \in C^q([0, T] \setminus \Sigma, \mathbb{G})$.

Dimostrazione. Il teorema di Pansu 6.2 ci dice che

$$\lim_{h \rightarrow 0} \delta_{1/h}(\gamma^{-1}(t)\gamma(t+h)) = \exp(\dot{\gamma}(t)) \in H$$

per quasi ogni $t \in [0, T]$ e quindi, posto

$$\Sigma := \left\{ t : \liminf_{h \rightarrow 0} d_{CC}(\delta_{1/h}(\gamma^{-1}(t)\gamma(t+h)), H) \geq r \right\},$$

abbiamo che Σ ha misura nulla; qui r è quello della proprietà (P1). Vogliamo dimostrare che Σ è chiuso e che γ è di classe \mathcal{C}^q in $[0, T] \setminus \Sigma$.

Sia dunque $t \notin \Sigma$ e supponiamo che $t \neq 0, T$ (in caso contrario la dimostrazione segue comunque la stessa linea); abbiamo che $d_{CC}(\delta_{1/h}(\gamma(t)^{-1}\gamma(t+h)), H) < r$ per un certo $h > 0$ sufficientemente piccolo. Per continuità esiste $\varepsilon > 0$ tale che $d_{CC}(\delta_{1/h}(\gamma(t-\varepsilon)^{-1}\gamma(t+h)), H) < r$; consideriamo allora la curva

$$\begin{aligned} \gamma_h &: [-\varepsilon, h] \rightarrow \mathbb{G} \\ s &\longmapsto \delta_{1/h}(\gamma(t-\varepsilon)^{-1}\gamma(t+s)). \end{aligned}$$

Abbiamo che $\gamma_h(-\varepsilon) = e$ e $\gamma_h(h) = \delta_{1/h}(\gamma(t-\varepsilon)^{-1}\gamma(t+h))$; dal momento che $d_{CC}(\delta_{1/h}(\gamma(t-\varepsilon)^{-1}\gamma(t+h)), H) < r$, la proprietà (P1) ci dice allora che la geodetica γ_h è di classe \mathcal{C}^q nell'intervallo $[-\varepsilon, h]$.

Da questa informazione segue facilmente che γ è di classe \mathcal{C}^q in un intorno di t ; il teorema è dunque dimostrato. \square

7.4 Regolarità globale delle geodetiche

Diamo ora una condizione che, unitamente alla (P1), garantisce che le geodetiche sono globalmente di classe \mathcal{C}^q .

Teorema 7.2. *Sia \mathbb{G} un gruppo di Carnot per cui vale la proprietà (P1) come nel teorema 7.1; supponiamo inoltre che valga la seguente proprietà*

- (P2) *Se $c : [0, \infty] \rightarrow \mathbb{G}$ è una geodetica parametrizzata per lunghezza d'arco tale che $c(0) = e$ e $c|_{[s,t]}$ è minimale per ogni $s < t$, allora $c(t) = \exp(tv)$ per un certo $v \in V_1$ con $|v|_H = 1$.*

Allora ogni geodetica γ parametrizzata per lunghezza d'arco è di classe \mathcal{C}^q (dove q è come nella proprietà (P1)).

Dimostrazione. Seguendo la notazione introdotta nella dimostrazione del teorema 7.1 sarà sufficiente dimostrare che per ogni geodetica γ parametrizzata per lunghezza d'arco si ha $\Sigma = \emptyset$.

Procediamo per assurdo e supponiamo che esista $t \in \Sigma$; possiamo naturalmente supporre che $t = 0$. Inoltre, a meno di una traslazione a sinistra per $\gamma(0)^{-1}$, possiamo supporre $\gamma(0) = e$: è chiaro che questa operazione non influisce sulla regolarità di γ , in quanto le traslazioni a sinistra sono di classe \mathcal{C}^∞ . Per definizione di Σ abbiamo allora che esistono $r_0, \varepsilon > 0$ tali che

$$d_{CC}(\delta_{1/r}(\gamma(r)), H) \geq \varepsilon \quad \text{per ogni } |r| \leq r_0.$$

A meno di restringere ulteriormente r_0 possiamo anche supporre che $\gamma|_{[-r_0, r_0]}$ sia minimale. Per ogni $r \in [-r_0, r_0]$ definiamo

$$\begin{aligned} \gamma_r &: \left[-\frac{r_0}{r}, \frac{r_0}{r}\right] \rightarrow \mathbb{G} \\ s &\longmapsto \delta_{1/r}(\gamma(sr)) \end{aligned}$$

ed estendiamo γ_r a tutto \mathbb{R} , ponendo $\gamma_r(s) := \gamma_r(\text{sign}(s)\frac{r_0}{r})$ se $|s| > \frac{r_0}{r}$. Le γ_r sono equicontinue in quanto 1-Lipschitziane; inoltre $\gamma_r(0) = e$ e dunque esse risultano anche equilimitate sui compatti di \mathbb{R} . Ne segue che esiste una successione $r_j \rightarrow 0$ tale che γ_{r_j} converge uniformemente sui compatti di \mathbb{R} ad una certa curva $\bar{\gamma} : \mathbb{R} \rightarrow \mathbb{G}$.

Dimostriamo ora che per ogni $s < t$ la curva $\bar{\gamma}|_{[s, t]}$ è minimizzante. Innanzitutto notiamo che ogni γ_r è 1-Lipschitziana e quindi lo sarà anche $\bar{\gamma}$; il nostro obiettivo sarà dunque raggiunto se mostreremo che $d_{CC}(\bar{\gamma}(s), \bar{\gamma}(t)) = t - s$. Tuttavia la minimalità di γ in $[-r_0, r_0]$ implica che γ_r è minimizzante in $[-\frac{r_0}{r}, \frac{r_0}{r}]$, quindi per j sufficientemente grande (precisamente per $\frac{r_0}{r_j} > \sup\{|s|, |t|\}$) avremo che $d_{CC}(\gamma_{r_j}(s), \gamma_{r_j}(t)) = t - s$ e dunque

$$d_{CC}(\bar{\gamma}(s), \bar{\gamma}(t)) = \lim_{j \rightarrow +\infty} d_{CC}(\gamma_{r_j}(s), \gamma_{r_j}(t)) = t - s.$$

In particolare $\bar{\gamma}$ è anche parametrizzata per lunghezza d'arco; la (P2) ci dice allora che esiste $v \in V_1$, $|v|_H = 1$, tale che $\bar{\gamma}(s) = \exp(sv)$ e quindi $\bar{\gamma}(1) \in H$.

Tuttavia per $|r| < r_0$ abbiamo $d_{CC}(\gamma_r(1), H) \geq \varepsilon$, che ci darebbe $d_{CC}(\bar{\gamma}(1), H) \geq \varepsilon > 0$. L'assurdo nasce dall'aver supposto $\Sigma \neq \emptyset$, quindi γ è di classe \mathcal{C}^q su tutto l'intervallo di definizione. Il teorema è dunque dimostrato. \square

Le proprietà (P1) e (P2) sono verificate, ad esempio, nel gruppo di Heisenberg. Non è noto se esse siano vere in generale nei gruppi di Carnot: un'indagine in questo senso giocherebbe un ruolo fondamentale nella questione della regolarità delle geodetiche.

Appendice A

Funzioni assolutamente continue

I problemi variazionali che andremo a trattare saranno tipicamente ambientati nello spazio delle funzioni assolutamente continue. Andiamo dunque a presentarne brevemente le proprietà principali, seguendo l'approccio di [2], a cui rimandiamo per le dimostrazioni ed eventuali approfondimenti.

Sia dunque $I =]a, b[$ un intervallo aperto (limitato o meno) e sia $p \in [1, +\infty]$.

Definizione A.1. Lo *spazio di Sobolev* $W^{1,p}(I)$ è l'insieme delle funzioni $u \in L^p(I)$ tali che esiste $g \in L^p(I)$ per cui

$$\int_I u \dot{\varphi} = - \int_I g \varphi$$

per ogni $\varphi \in \mathcal{C}_c^\infty(\Omega)$.

Per $u \in W^{1,p}(I)$ scriveremo \dot{u} invece di g (che è la derivata nel senso delle distribuzioni): effettivamente si può dimostrare che tale funzione coincide quasi ovunque con la derivata (intesa nel senso classico di limite dei rapporti incrementali) del rappresentante continuo della classe di equivalenza di u . Vale anche di più: se $u \in L^p$ e $\dot{u} \in L^p$ (qui intendiamo la derivata usuale), allora $u \in W^{1,p}$, e la derivata distribuzionale coincide quasi ovunque con quella classica. In particolare $\mathcal{C}^1(\bar{I}) \subset W^{1,p}(I)$ se I è limitato.

Lo spazio $W^{1,p}(I)$ viene munito della norma

$$\|u\|_{W^{1,p}(I)} := \left(\|u\|_{L^p(I)}^p + \|\dot{u}\|_{L^p(I)}^p \right)^{1/p}$$

che lo rende completo e dunque uno spazio di Banach. Tale norma è poi equivalente alla

$$\|u\|_{L^p(I)} + \|\dot{u}\|_{L^p(\Omega)}.$$

Se $p = 2$ scriveremo $H^1(I)$ invece di $W^{1,2}(I)$: esso risulta uno spazio di Hilbert in quanto la norma $\|\cdot\|_{H^1}$ è indotta dal prodotto scalare

$$\langle u, v \rangle_{H^1} := \int_I uv + \int_I \dot{u}\dot{v}.$$

Proposizione A.1. *Lo spazio di Banach $W^{1,p}$ è riflessivo per $1 < p < \infty$ e separabile per $1 \leq p < \infty$. Lo spazio H^1 è uno spazio di Hilbert separabile.*

Fondamentalmente le funzioni di $W^{1,p}$ sono le primitive di funzioni dello spazio L^p ; vale infatti il

Teorema A.1. *Sia $u \in W^{1,p}(I)$; allora esiste un (unico) rappresentante continuo \tilde{u} della classe di equivalenza di u in L^p (ovvero secondo la relazione di uguaglianza quasi ovunque). Inoltre per ogni $x, y \in I$ si ha*

$$\tilde{u}(x) - \tilde{u}(y) = \int_y^x \dot{u}(t) dt.$$

Talvolta può essere utile approssimare le funzioni di $W^{1,p}$ con funzioni regolari. Ciò può essere fatto mediante il seguente

Teorema A.2. *Sia $u \in W^{1,p}(I)$ con $p \neq \infty$. Allora esiste una successione $\{u_j\} \subset C_c^\infty(\mathbb{R})$ tale che*

$$u_j|_I \xrightarrow{W^{1,p}(I)} u.$$

In altre parole, il sottospazio C_c^∞ è denso in $W^{1,p}$.

Ad esempio, grazie al teorema A.2 si dimostra che valgono le usuali formule di derivata di un prodotto e di integrazione per parti quando si ha a che fare con funzioni di $W^{1,p}$. Se I è limitato, $v \in C^1(\mathbb{R})$ e $u \in W^{1,p}(I)$ vale anche la nota formula di derivazione per la funzione composta $v \circ u$.

Indichiamo con $(W^{1,p}(I))'$ il duale di $W^{1,p}(I)$, ovvero lo spazio dei funzionali lineari e continui su $W^{1,p}(I)$. Diremo che una successione $\{u_j\} \subset W^{1,p}(I)$ **converge debolmente** a $u \in W^{1,p}(I)$ (e scriveremo $u_j \rightharpoonup u$) se

$$F(u_j) \xrightarrow{\mathbb{R}} F(u)$$

per ogni $F \in (W^{1,p}(I))'$. Indicando con p' l'esponente coniugato a p abbiamo

Teorema A.3. *Sia $F \in (W^{1,p}(I))'$, $p \neq \infty$. Allora esistono $f_0, f_1 \in L^{p'}(I)$, non necessariamente uniche, tali che*

$$F(u) = \int_I f_0 u + \int_I f_1 \dot{u}$$

per ogni $u \in W^{1,p}(I)$.

Quindi $u_j \rightharpoonup u$ se e solo se $\int_I u_j f \rightarrow \int_I u f$ e $\int_I \dot{u}_j f \rightarrow \int_I \dot{u} f$ per ogni $f \in L^{p'}$, ovvero se e solo se $u_j \rightharpoonup u$ e $\dot{u}_j \rightharpoonup \dot{u}$ in L^p .

Definizione A.2. Una funzione $u : I \rightarrow \mathbb{R}$ si dice **assolutamente continua** se per ogni $\epsilon > 0$ esiste $\delta > 0$ tale che

$$\sum_{i=1}^m (b_i - a_i) < \delta \implies \sum_{i=1}^m |u(b_i) - u(a_i)| < \epsilon$$

per ogni famiglia finita $\{]a_1, b_1[, \dots,]a_m, b_m[\}$ di segmenti disgiunti contenuti in I .

È immediato verificare che una funzione assolutamente continua è in particolare continua. Lo spazio delle funzioni assolutamente continue su I si indica con $AC(I)$. Il seguente teorema ci dice che tale spazio coincide con lo spazio di Sobolev $W^{1,1}(I, \mathbb{R})$; precisamente

Teorema A.4. Se $u \in AC(I)$, allora u ha quasi ovunque una derivata classica \dot{u} che appartiene ad $L^1(I)$ e che, vista come elemento di L^1 , è la derivata distribuzionale di u .

Viceversa, ogni $u \in W^{1,1}(I)$ ha un (unico) rappresentante che risulta una funzione assolutamente continua.

Infine, $u \in AC(I)$ se e solo se u è derivabile in senso classico quasi ovunque, con derivata \dot{u} in $L^1(I)$ e tale che vale il teorema fondamentale del calcolo, ovvero

$$u(x) - u(y) = \int_y^x \dot{u} dt$$

per ogni $x, y \in I$.

Infine possiamo estendere la definizione di spazi di Sobolev $W^{1,p}$ e di funzioni assolutamente continue alle funzioni a valori in \mathbb{R}^n , imponendo che ogni componente stia in $W^{1,p}$ (o in AC). Anche in questi spazi valgono le naturali estensioni della proposizione A.1 e dei teoremi A.1, A.2, A.3 e A.4. In particolare

$$\begin{aligned} u_j \rightharpoonup u \text{ in } AC(I, \mathbb{R}^n) &\iff \int_I u_j^i f \rightarrow \int_I u^i f \quad \text{e} \quad \int_I \dot{u}_j^i f \rightarrow \int_I \dot{u}^i f \\ &\text{per ogni } i = 1, \dots, n \text{ e ogni } f \in L^\infty(I) \\ &\iff u_j^i \rightharpoonup u^i \text{ e } \dot{u}_j^i \rightharpoonup \dot{u}^i \text{ in } L^1 \text{ per ogni } i = 1, \dots, n. \end{aligned}$$

La norma di AC (cioè quella di $W^{1,1}$) è semicontinua inferiormente per convergenza debole; il teorema di Dunford-Pettis 3.4 fornisce criteri per poter estrarre sottosuccessioni debolmente convergenti da successioni limitate in AC .

Bibliografia

- [1] J. M. Ball and V. J. Mizel. One-dimensional variational problems whose minimizers do not satisfy the Euler-Lagrange equation. *Arch. Ration. Mech. Anal.*, 90:325–388, 1985.
- [2] H. Brezis. *Analisi Funzionale: Teoria e Applicazioni*. Liguori Editore, Napoli, 1990.
- [3] G. Buttazzo. *Semicontinuity, relaxation and integral representation in the Calculus of Variations*. Longman House, Harlow, 1989.
- [4] G. Buttazzo, M. Giaquinta, and S. Hildebrandt. *One-dimensional Variational Problems: An Introduction*. Clarendon Press, Oxford, 1998.
- [5] C. Carathéodory. Untersuchungen über die grundlagen der thermodynamik. *Math. Ann.*, 67:355–386, 1909.
- [6] C. Castaing and M. Valadier. *Convex Analysis and Measurable Multifunction*. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1977.
- [7] W. L. Chow. Über systeme von linearen partiellen differentialgleichungen erster ordnung. *Math. Ann.*, 117:98–105, 1939.
- [8] F. H. Clarke and R. B. Vinter. Existence and regularity in the small in the Calculus of Variations. *J. Differ. Equations*, 59:336–354, 1985.
- [9] F. H. Clarke and R. B. Vinter. Regularity properties of solutions to the basic problem in the Calculus of Variations. *Trans. Am. Math. Soc.*, 289:73–98, 1985.
- [10] E. De Giorgi. Teoremi di semicontinuità nel Calcolo delle Variazioni, 1968-69. Note of a course held at the Istituto Nazionale di Alta Matematica, Roma.
- [11] C. Dellacherie and P. Meyer. *Probabilities and Potential*. North-Holland, Amsterdam, 1978.

- [12] W. H. Fleming and R. W. Rishel. *Deterministic and Stochastic Optimal Control*. Lecture Notes in Mathematics. Springer-Verlag, New York, 1975.
- [13] R. Hermann. Geodesics of singular riemannian metrics. *Bull. Am. Math. Soc.*, 79:780–782, 1973.
- [14] A. D. Ioffe. On lower semicontinuity of integral functionals I.II. *SIAM J. Control Optimization*, 15:521–538 and 991–1000, 1977.
- [15] E. Lanconelli. Stime subellittiche e metriche riemanniane singolari I. In *Seminario di Analisi Matematica*, pages VIII–1 VIII–17. Università di Bologna, 1983.
- [16] V. Magnani. *Elements of Geometric Measure Theory on sub-Riemannian Groups*. PhD thesis, Scuola Normale Superiore di Pisa, 2001.
- [17] R. Montgomery. Survey of singular geodesics. In A. Bellaïche and J.-J. Risler Editors, editors, *Sub-Riemannian Geometry*, Lecture Notes in Mathematics. Birkhäuser Verlag, Basel, 1996.
- [18] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. Mathematical Surveys and Monographs. American Mathematical Society, 2002.
- [19] R. Monti. *Distances, Boundaries and Surface Measures in Carnot-Carathéodory Spaces*. PhD thesis, Università degli Studi di Trento, 2001.
- [20] A. Nagel, E. M. Stein, and S. Wainger. Balls and metrics defined by vector fields I: Basic properties. *Acta Math.*, 155:103–147, 1985.
- [21] R. S. Strichartz. Sub-riemannian geometry. *J. Diff. Geom.*, 24:221–263, 1983.
- [22] R. S. Strichartz. Corrections to sub-riemannian geometry. *J. Diff. Geom.*, 30:595–596, 1989.
- [23] H. J. Sussmann. Orbits of families of vector fields and integrability of distributions. *Trans. Am. Math. Soc.*, 180:171–188, 1973.
- [24] H. J. Sussmann. A cornucopia of four-dimensional abnormal sub-riemannian minimizers. In A. Bellaïche and J.-J. Risler Editors, editors, *Sub-Riemannian Geometry*, Lecture Notes in Mathematics. Birkhäuser Verlag, Basel, 1996.
- [25] V. S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations*. Graduate Texts in Mathematics. Springer-Verlag, New York, 1974.