Appunti del corso di Calcolo delle Variazioni

Anno Accademico 2014 – 2015

Franco Rampazzo

Note a cura di Giorgio Stefani

Premessa. Le note che seguono sono gli appunti sulla prima parte del corso “Calcolo delle Variazioni” tenuto dal Professor Franco Rampazzo nel secondo semestre dell’anno accademico 2014 – 2015. Il testo di riferimento utilizzato per questi argomenti durante il corso è *Functional Analysis, Calculus of Variations and Optimal Control* di F. Clarke, [**2**]. Sono molto graditi suggerimenti sull’organizzazione del materiale didattico contenuto in queste note e segnalazioni di eventuali errori. (Versione del 16 maggio 2015).

Indice

CAPITOLO 1

Introduzione

Il Calcolo delle Variazioni è un campo dell'Analisi matematica che si occupa della ricerca dei punti estremali (massimi e minimi) dei cosiddetti funzionali — ovvero funzioni il cui dominio è a sua volta un insieme di funzioni — e delle loro proprietà. Tali funzionali possono per esempio essere formulati come integrali che coinvolgono una funzione incognita e le sue derivate. L'interesse principale è rivolto allo studio delle funzioni estremali, cioè quelle funzioni che rendono massimo o minimo il valore del funzionale.

Alcuni problemi classici sulle curve erano posti in questa forma. Un esempio è la curva detta *brachistocrona*, cioè il percorso, da un punto A ad un punto B non allineati verticalmente, lungo il quale una particella sottoposta alla sola gravità scenderebbe nel minor tempo: si deve quindi minimizzare la funzione che rappresenta il tempo fra tutte le curve da A a B. (Per un approccio intuitivo al problema della brachistocrona e per una breve introduzione al procedimento formale tipico della matematica del Settecento e, nello specifico, degli inizi del Calcolo delle Variazioni, si può consultare [**3**, Capitolo 7, §10]).

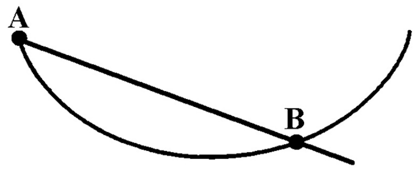

Figura 1.1. Brachistocrona (cicloide) e segmento.

Il teorema chiave del Calcolo delle Variazioni classico è l'equazione di Eulero–Lagrange, che corrisponde a una condizione di stazionarietà per il funzionale: come nel caso della ricerca dei massimi e dei minimi di una funzione, l'analisi delle piccole variazioni attorno a una presunta soluzione porta a una condizione del primo ordine. Non è possibile dire direttamente se è stato trovato un massimo, un minimo, o nessuno dei due: per poter avere qualche informazione in più è necessario studiare qualche condizione del secondo ordine, fornita ad esempio dai teoremi di Legendre e di Jacobi.

1. Il caso finito dimensionale

Supponiamo di voler minimizzare o massimizzare una funzione definita su un aperto di $\mathbb{R}^n$. Il seguente risultato fornisce una condizione necessaria affinché un punto x_* sia estremale.

Teorema 1.1 (Principio di Fermat). *Sia Ω un aperto di $\mathbb{R}^n$ e sia $f\colon \Omega \to \mathbb{R}$ una funzione. Supponiamo che $x_* \in \Omega$ sia un punto di estremo locale di f. Se f è differenziabile nel punto x_*, allora $Df(x_*) = 0$.*

Il motivo per cui vale il Teorema 1.1 è il seguente. Se $x_* \in \Omega$ è un punto di estremo locale per f, ad esempio un minimo locale, allora, fissato un vettore $w \in \mathbb{R}^n$, per ogni $\lambda \in \mathbb{R}$ sufficientemente piccolo deve essere

$$f(x_* + \lambda w) \geq f(x_*)$$

da cui

$$\frac{f(x_* + \lambda w) - f(x_*)}{\lambda} \geq 0$$

per $\lambda > 0$, e

$$\frac{f(x_* + \lambda w) - f(x_*)}{\lambda} \leq 0$$

per $\lambda < 0$. Quindi, passando al limite per $\lambda \to 0$, la derivata direzionale deve essere nulla, $D_w f(x_*) = 0$. Siccome $w \in \mathbb{R}^n$ è un vettore arbitrario, segue che $Df(x_*) = 0$.

In conclusione, se lo spazio ha dimensione finita e se f è una funzione differenziabile, la condizione $Df = 0$ deve essere verificata in ogni estremante locale per f. Questa semplice proprietà ci permette pertanto di individuare i punti del dominio di f candidati ad essere estremanti locali.

Cosa accade, invece, se lo spazio ha dimensione infinita?

2. Il caso infinito dimensionale

Nel caso di dimensione infinita, siamo interessati alla ricerca di una curva $x \in C^2([a,b];\mathbb{R})$ con estremi assegnati $x(a) = A$ e $x(b) = B$ che minimizza (o massimizza) una particolare quantità, come ad esempio la lunghezza, il tempo di percorrenza, l'energia... Facciamo qualche esempio.

Esempio 1.2 (Minima lunghezza). Supponiamo di voler individuare quale sia, tra tutte le curve $x \in C^2([a,b];\mathbb{R})$ con estremi assegnati $x(a) = A$ e $x(b) = B$, il percorso dal punto A al punto B che abbia minima lunghezza.

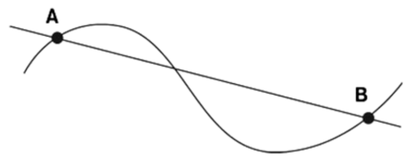

Figura 1.2. Il segmento come curva di minima lunghezza.

Ricordiamo che la lunghezza della curva x è data dall'integrale

$$\int_a^b \sqrt{1 + \dot{x}(t)^2}\,dt,$$

quindi il nostro problema è quello di minimizzare il *funzionale lunghezza*

$$J(x) = \int_a^b \sqrt{1 + \dot{x}(t)^2}\,dt$$

tra le curve $x \in C^2([a,b];\mathbb{R})$ con estremi assegnati $x(a) = A$ e $x(b) = B$. Naturalmente, come l'esperienza comune ci suggerisce, la curva di lunghezza minima che stiamo cercando è proprio il segmento di retta passante per A e B, cioè

$$x(t) = A + \frac{t-a}{b-a}(B - A)$$

e tale curva è pertanto il minimo (globale, in questo caso) per il funzionale lunghezza.[1]

Esempio 1.3 (Minima superficie). Supponiamo di voler individuare quale sia, tra tutte le curve $x \in C^2([a,b];\mathbb{R})$ con estremi assegnati $x(a) = A$ e $x(b) = B$, il profilo di minima superficie di rivoluzione dal punto A al punto B.

Ricordiamo che l'area della superficie di rivoluzione generata dalla curva x attorno all'asse delle ascisse è data dall'integrale

$$\int_a^b x(t)\sqrt{1 + \dot{x}(t)^2}\,dt,$$

quindi il nostro problema è quello di minimizzare il *funzionale di superficie*

$$J(x) = \int_a^b x(t)\sqrt{1 + \dot{x}(t)^2}\,dt$$

[1] Nella geometria Riemanniana le curve di minima lunghezza sono le *geodetiche*, che in generale non sono più segmenti di rette — si pensi ad esempio alla sfera in $\mathbb{R}^3$.

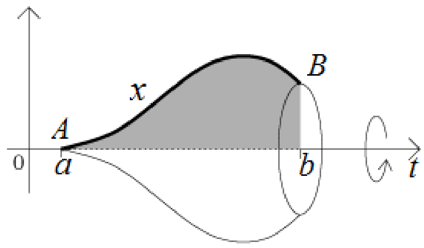

Figura 1.3. Superficie di rotazione generata dalla curva x.

tra le curve $x \in C^2([a,b];\mathbb{R})$ con estremi assegnati $x(a) = A$ e $x(b) = B$. Non è affatto chiaro, a differenza di quanto visto nell'esempio precedente, se esista un minimo anche solo locale per questo funzionale. Vedremo in seguito come risolvere questo problema.

Esempio 1.4 (Pendolo). Un esempio meccanico di minimizzazione di un funzionale è il seguente. Se indichiamo con $T = \frac{1}{2}a(q)\dot{q}^2$ l'energia cinetica e con $V = V(q)$ l'energia potenziale, allora il *moto effettivo* della particella materiale di coordinate meccaniche q minimizza il *funzionale d'azione*

$$\int_a^b (T - V)dt.$$

Un esempio classico è dato dal moto del pendolo.

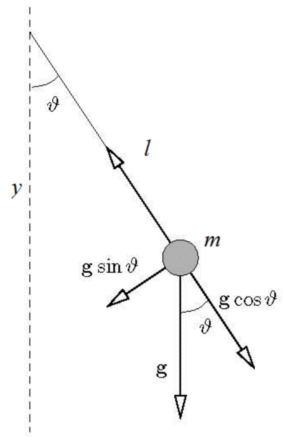

Figura 1.4. Schema delle forze nel moto del pendolo.

In riferimento alla Figura 1.4, abbiamo

$$y = l(1 - \cos\theta)$$

da cui si ricava

$$V = V(\theta) = mgy = mgl(1 - \cos\theta)$$

e

$$T = T(\theta) = \frac{1}{2}mv^2 = \frac{1}{2}ml^2\dot{\theta}^2.$$

Allora il moto oscillatorio del pendolo $t \mapsto \theta(t)$ minimizza il funzionale d'azione

$$\int_a^b \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1 - \cos\theta)dt.$$

CAPITOLO 2

L'equazione di Eulero–Lagrange

Nel Capitolo 1 abbiamo visto come, in dimensione finita, sia possibile dare condizioni necessarie (il Principio di Fermat, incontrato nel Teorema 1.1) affinché un punto sia un estremo (locale) per una funzione differenziabile. Vorremmo ora estendere questi risultati al caso infinito dimensionale.

1. L'equazione di Eulero–Lagrange come condizione del primo ordine

Supponiamo che la curva $x \in C^2([a,b];\mathbb{R})$ abbia estremi assegnati $x(a) = A$ e $x(b) = B$ e supponiamo che sia dato il funzionale

$$J(x) = \int_a^b L(t, x(t), \dot{x}(t))dt \tag{1.1}$$

dove $L = L(t,x,v)$ (tempo, posizione e velocità[1]) è una funzione di classe C^2 definita su un opportuno sottoinsieme di $\mathbb{R}^3$ ed è detta *Lagrangiana.* Il problema basilare che ci proponiamo è di minimizzare il funzionale J tra le funzioni di classe C^2 che soddisfano i vincoli assegnati, cioè

$$\min \ J(x) \ : \ x \in C^2([a,b];\mathbb{R}), \ x(a) = A, \ x(b) = B \tag{P}$$

(e analogamente se vogliamo invece massimizzare il funzionale assegnato). Diamo ora alcune definizioni.

Definizione 2.1. Il funzionale $J(x)$ alle volte è detto *costo* corrispondente alla funzione x. Una funzione x si dice *ammissibile* se x soddisfa le ipotesi di regolarità e i vincoli assegnati nel problema (P). Una *soluzione* x_* al problema (P) è una funzione ammissibile tale che

$$J(x_*) \leq J(x)$$

per tutte le funzioni x ammissibili — in questo caso diciamo anche che x_* è un *minimo* per il problema (P).

Teorema 2.2 (Equazione di Eulero–Lagrange). *Sia L di classe C^2 e sia $x_* \in C^2([a,b];\mathbb{R})$ un minimo, ovvero una soluzione per il problema* (P). *Allora x_* soddisfa l'equazione*[2]

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} = 0. \tag{Eulero-Lagrange}$$

Remark 2.3. La (Eulero-Lagrange) è un'equazione differenziale implicita e del secondo ordine, quindi la regolarità C^2 assunta per ipotesi è indispensabile affinché l'equazione e i calcoli necessari per la sua dimostrazione abbiano senso.

Dimostrazione. Fissiamo una *variazione* compatibile con il problema (P), cioè una funzione $w \in C^2([a,b];\mathbb{R})$ tale che

$$w(a) = w(b) = 0. \tag{1.2}$$

Osserviamo che, per ogni $\lambda \in \mathbb{R}$, la funzione $x_* + \lambda w$ è una funzione ammissibile per il problema (P). Consideriamo allora la seguente funzione (di una sola variabile)

$$g(\lambda) = J(x_* + \lambda w) = \int_a^b L(t, x_* + \lambda w, \dot{x}_* + \lambda \dot{w})dt, \tag{1.3}$$

[1] Talvolta, seguendo la notazione dei fisici, indicheremo la variabile velocità con $\dot{x}$, che in tal caso dovrà intendersi come un simbolo, non come una funzione del tempo.

[2] L'equazione di Eulero–Lagrange venne scoperta attorno al 1754 da Leonard Euler e, contemporaneamente, da Joseph-Louis Lagrange in connessione con lo studio del problema della brachistocrona (si veda il Capitolo 1). La dimostrazione originale di Eulero della (Eulero-Lagrange) passava attraverso un argomento di discretizzazione; la dimostrazione che presentiamo qui è dovuta a Lagrange.

dove la dipendenza dal tempo degli argomenti della Lagrangiana è stata omessa per semplicità. Per ipotesi x_* è un minimo per il funzionale J, per cui il punto $\lambda = 0$ è un minimo per la funzione g. Siccome per ipotesi sia L che x_* e w sono di classe almeno C^2, per il Teorema di derivazione sotto il segno di integrale (si veda ad esempio [**4**, Capitolo 6]) la funzione g è differenziabile. Per il Principio di Fermat (Teorema 1.1) si ha pertanto che

$$\frac{d}{d\lambda} g(\lambda)\Big|_{\lambda=0} = 0.$$

Ma, d'altra parte,

$$\begin{aligned}
\frac{d}{d\lambda} g(\lambda)\Big|_{\lambda=0} &= \int_a^b \frac{d}{d\lambda}\left(L(t, x_* + \lambda w, \dot{x}_* + \lambda \dot{w})\right)\Big|_{\lambda=0} dt = \\
&= \int_a^b \frac{\partial L}{\partial x}(t, x_* + \lambda w, \dot{x}_* + \lambda \dot{w}) \cdot w + \frac{\partial L}{\partial v}(t, x_* + \lambda w, \dot{x}_* + \lambda \dot{w}) \cdot \dot{w}\Big|_{\lambda=0} dt = \\
&= \int_a^b \frac{\partial L}{\partial x}(t, x_*, \dot{x}_*) \cdot w + \frac{\partial L}{\partial v}(t, x_*, \dot{x}_*) \cdot \dot{w}\ dt = \quad \text{(int. per parti)} \\
&= \int_a^b \frac{\partial L}{\partial x}(t, x_*, \dot{x}_*) \cdot w\ dt + \underbrace{\frac{\partial L}{\partial v}(t, x_*, \dot{x}_*) \cdot w\Big|_{t=a}^{t=b}}_{=0 \text{ per la } (1.2)} - \int_a^b \frac{d}{dt}\frac{\partial L}{\partial v}(t, x_*, \dot{x}_*) \cdot w\ dt = \\
&= \int_a^b \left(\frac{\partial L}{\partial x}(t, x_*, \dot{x}_*) - \frac{d}{dt}\frac{\partial L}{\partial v}(t, x_*, \dot{x}_*)\right) \cdot w\ dt
\end{aligned}$$

e quindi, per ogni variazione $w \in C^2([a,b];\mathbb{R})$ tale che $w(a) = w(b) = 0$ si deve avere che

$$\int_a^b \left(\frac{\partial L}{\partial x}(t, x_*, \dot{x}_*) - \frac{d}{dt}\frac{\partial L}{\partial v}(t, x_*, \dot{x}_*)\right) \cdot w\ dt = 0.$$

Per il Lemma 2.4 qui sotto si conclude che $\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial v} = 0$, che è proprio la (Eulero-Lagrange). ☐

LEMMA 2.4. *Sia $f\colon [a,b] \to \mathbb{R}$ una funzione continua. Supponiamo che*

$$\int_a^b f(t)w(t)\ dt = 0$$

per ogni funzione $w \in C^2([a,b];\mathbb{R})$ tale che $w(a) = w(b) = 0$. Allora $f \equiv 0$ in $[a,b]$.

DIMOSTRAZIONE. Per assurdo, supponiamo esista $\bar{t} \in (a,b)$ tale che $f(\bar{t}) \neq 0$, ad esempio $f(\bar{t}) > r$, per un qualche $r > 0$. Allora, per continuità, esiste $\delta > 0$ tale che $f(t) > r/2$ per $|t - \bar{t}| < \delta$ (e, pertanto, non è restrittivo supporre che $\bar{t}$ sia un punto interno all'intervallo). Sia allora

$$w(t) = \frac{\eta(\delta^2 - |t-\bar{t}|^2)}{\eta(\delta^2 - |t-\bar{t}|^2) + \eta(|t-\bar{t}|^2 - \delta^2/4)} \tag{1.4}$$

dove

$$\eta(t) = \begin{cases} e^{-1/t} & t \geq 0 \\ 0 & t < 0. \end{cases}$$

Abbiamo che $w \in C^2([a,b];\mathbb{R})$, perché $\eta \in C^\infty(\mathbb{R})$, ed inoltre $w(a) = w(b) = 0$. In particolare, $w \equiv 1$ per $|t - \bar{t}| < \delta/2$, $w \equiv 0$ per $|t - \bar{t}| \geq \delta$ e $w \geq 0$ su $[a,b]$. Allora, per tale w, si ha

$$\int_a^b f(t)w(t)\ dt = \int_{\bar{t}-\delta}^{\bar{t}+\delta} f(t)w(t)\ dt \geq \int_{\bar{t}-\delta/2}^{\bar{t}+\delta/2} f(t)w(t)\ dt > \frac{r\delta}{2} > 0,$$

una contraddizione. ☐

REMARK 2.5. Il Lemma 2.4 è una semplice variante del seguente risultato, detto Lemma fondamentale del Calcolo delle Variazioni: *se $f \in C^k([a,b])$ è tale che $\int_a^b f(t)w(t)\ dt = 0$ per ogni funzione $w \in C^k([a,b])$ tale che $w(a) = w(b) = 0$, allora $f \equiv 0$ in $[a,b]$.*

Si noti che, nella dimostrazione del Lemma 2.4, la funzione w definita nella (1.4) è di classe C^∞. In effetti, il Lemma fondamentale del Calcolo delle Variazioni può essere generalizzato nel seguente risultato, noto come Lemma di du Bois-Reymond: *se Ω è un aperto in $\mathbb{R}^d$ e se $u \in L^1_{loc}(\Omega)$ è tale che $\int_\Omega u\phi\, dx = 0$ per ogni funzione $\phi \in C_c^\infty(\Omega)$, allora $u = 0$ q.o. in Ω.* Per una dimostrazione si veda [**1**, Corollario 4.24].

Esempio 2.6 (Legge di Newton). Ricordiamo che, nel caso meccanico dell'Esempio 1.4, avevamo

$$L = \frac{1}{2}mv^2 - V(x).$$

Per tale Lagrangiana, la (Eulero-Lagrange) diviene

$$m\ddot{x} - V'(x) = 0,$$

cioè la legge di Newton.

Esempio 2.7. Nell'Esempio 1.2 abbiamo intuito che, tra le curve $x \in C^2([a,b];\mathbb{R})$ con estremi assegnati $x(a) = A$ e $x(b) = B$, il percorso di lunghezza minima è il segmento di retta passante per A e B. Con la (Eulero-Lagrange) possiamo ora provare rigorosamente questa affermazione. Infatti, per il funzionale di minima lunghezza, avevamo

$$L = \sqrt{1 + v^2}$$

e per tale Lagrangiana la (Eulero-Lagrange) diventa

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} = \frac{d}{dt}\left(\frac{v}{\sqrt{1+v^2}}\right) = 0$$

da cui $\dfrac{\dot{x}}{\sqrt{1+\dot{x}^2}}$ = costante. Ma allora deve essere $\dot{x}_*$ = costante, e quindi integrando e imponendo i vincoli agli estremi si ritrova

$$x(t) = A + \frac{t-a}{b-a}(B - A).$$

2. Minimi locali deboli e forti

Abbiamo dimostrato la (Eulero-Lagrange) supponendo che x_* fosse un minimo globale per il funzionale J. In realtà, nella dimostrazione non abbiamo mai usato il fatto che questo minimo fosse globale: è sufficiente, infatti, che x_* sia un minimo locale (ovvero che, per la funzione g definita in (1.3), si possa scegliere λ in un opportuno intorno aperto dell'origine).

Ad una più attenta analisi della dimostrazione della (Eulero-Lagrange), emerge la seguente osservazione: la perturbazione $x_* + \lambda w$ del minimo locale x_* tramite la variazione w è "vicina" al minimo x_* nel senso della topologia data dalla norma C^1. Abbiamo infatti che le funzioni $x_* + \lambda w$ e x_* sono vicine[3]

- in norma, perché $\|(x_* + \lambda w) - x_*\| = |\lambda|\|w\| \xrightarrow{\lambda\to 0} 0$;
- in norma per le derivate, perché $\|(\dot{x}_* + \lambda\dot{w}) - \dot{x}_*\| = |\lambda|\|\dot{w}\| \xrightarrow{\lambda\to 0} 0$.

Dunque, la soluzione x_* del problema (P) a cui abbiamo fatto riferimento nella dimostrazione della (Eulero-Lagrange), è un minimo locale per il funzionale J in un senso ben preciso, cioè secondo la topologia dello spazio delle funzioni di classe C^1.

Definizione 2.8 (Minimo forte e minimo debole). Diciamo che x_* è

(1) un *minimo locale debole* se esiste $\varepsilon > 0$ tale che $J(x_*) \le J(x)$ per ogni x ammissibile tale che $\|x - x_*\| < \varepsilon$ e $\|\dot{x} - \dot{x}_*\| < \varepsilon$ (topologia dello spazio C^1);
(2) un *minimo locale forte* se esiste $\varepsilon > 0$ tale che $J(x_*) \le J(x)$ per ogni x ammissibile tale che $\|x - x_*\| < \varepsilon$ (topologia dello spazio C^0).

(Analoga definizione per il caso in cui x_* sia un massimo locale).

[3] In questi appunti, con $\|\cdot\|$ indichiamo la norma dello spazio $L^\infty([a,b];\mathbb{R})$, cioè l'insieme delle funzioni definite q.o. sull'intervallo $[a,b]$ che sono essenzialmente limitate. Ricordiamo che, per le funzioni continue sull'intervallo $[a,b]$, la norma $\|\cdot\|$ diventa semplicemente la norma della convergenza uniforme (o "norma del sup"), cioè si ha che

$$\|f\| = \sup_{[a,b]}|f(t)| = \max_{[a,b]}|f(t)|$$

per ogni funzione $f\colon [a,b] \to \mathbb{R}$ continua.

Con questa nuova terminologia, per quanto osservato sopra, il Teorema 2.2 si può rienunciare nella forma seguente:

Sia L di classe C^2 e sia $x_* \in C^2([a,b];\mathbb{R})$ un minimo locale debole per il funzionale J. Allora x_* soddisfa la (Eulero-Lagrange).

REMARK 2.9. Nella Definizione 2.8, la differenza tra minimo locale debole e minimo locale forte è sottile ma cruciale. *Un minimo locale forte è sempre un minimo locale debole.* Infatti, quando diciamo che x_* è un minimo locale forte, stiamo confrontando la funzione x_* con un insieme molto grande di funzioni ammissibili x, perché stiamo usando come metro di giudizio una topologia molto grezza, quella dello spazio C^0. Quando invece diciamo che x_* è un minimo locale debole, stiamo confrontando la funzione x_* con un insieme più piccolo di funzioni ammissibili x, perché in questo caso stiamo usando come metro di giudizio una topologia molto più fine, cioè più precisa, della precedente, quella dello spazio C^1, in cui anche le derivate devono essere "vicine". Pertanto, se x_* è un minimo locale forte, allora è minimo locale in confronto ad un insieme molto grande di funzioni ammissibili (spazio C^0), e quindi, in particolare, è minimo locale in confronto ad un sottoinsieme di questo insieme (spazio C^1), cioè è difatti anche un minimo locale debole.

3. La condizione di Erdmann

Il risultato che segue è una diretta conseguenza della (Eulero-Lagrange) nel caso in cui la Lagrangiana L sia indipendente dal tempo, $L = L(x,v)$, e costituisce un utilissimo punto di partenza per identificare le curve estremali per il problema (P).

TEOREMA 2.10 (Condizione di Erdmann, 1877). *Sia $x_* \in C^2([a,b];\mathbb{R})$ un minimo locale debole per il problema* (P)*. Se la Lagrangiana L è di classe C^2 e non dipende dal tempo, allora esiste una costante $k \in \mathbb{R}$ tale che*

$$\frac{\partial L}{\partial v}(x_*,\dot{x}_*)\cdot\dot{x}_* - L(x_*,\dot{x}_*) = k. \tag{Erdmann}$$

DIMOSTRAZIONE. Abbiamo infatti che

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial L}{\partial v}(x_*,\dot{x}_*)\cdot\dot{x}_* - L(x_*,\dot{x}_*)\right) &= \frac{d}{dt}\left(\frac{\partial L}{\partial v}\right)\dot{x}_* + \frac{\partial L}{\partial v}\ddot{x}_* - \frac{\partial L}{\partial x}\dot{x}_* - \frac{\partial L}{\partial v}\ddot{x}_* = \\ &= \dot{x}_*\left(\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x}\right) = 0\end{aligned}$$

perché la funzione L è indipendente dal tempo e perché x_* soddisfa la (Eulero-Lagrange). □

ESEMPIO 2.11. Nel caso meccanico la (Erdmann) diventa proprio l'energia totale associata al sistema. Infatti era

$$L = \frac{1}{2}mv^2 - V(x)$$

(si veda l'Esempio (1.4)), per cui

$$\frac{\partial L}{\partial v}\cdot v - L = mv\cdot v - \frac{1}{2}mv^2 + V(x) = \frac{1}{2}mv^2 + V(x) \equiv E_{totale}.$$

4. La catenaria: analisi del problema della minima superficie

Nell'Esempio 1.3 abbiamo introdotto il problema della minima area per una superficie di rotazione. A differenza del segmento di retta, non è assolutamente facile intuire se, anzitutto, esista una curva di minimo per questo problema e, se questo è il caso, che forma debba avere.

Armati dei nuovi strumenti sviluppati nei Teoremi 2.2 e 2.10, possiamo affrontare questo problema analiticamente, studiando le condizioni necessarie che una ipotetica curva di minimo x_* debba soddisfare.

Per il problema della minima superficie, la Lagrangiana era della forma

$$L = x\sqrt{1+v^2}.$$

Con un semplice calcolo di derivazione di funzione composta, si trova

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} = \frac{-\dot{x}^2 + x\ddot{x} - 1}{(1+\dot{x}^2)^{\frac{3}{2}}}.$$

Assumiamo che $x_*(t) > 0\ \forall t$. La (Eulero-Lagrange) diventa $-\dot{x}_*^2 + x_*\ddot{x}_* - 1 = 0$, per cui ricaviamo che

$$\ddot{x}_* = \frac{1+\dot{x}_*^2}{x_*}.$$

Pertanto $\ddot{x}_*(t) > 0\ \forall t$, quindi $\dot{x}_*$ è strettamente crescente e x_* è strettamente convessa.

Siccome L non dipende dal tempo, possiamo applicare il Teorema 2.10. La (Erdmann) diventa allora

$$\dot{x}\frac{\partial L}{\partial v} - L = -\frac{x}{\sqrt{1+\dot{x}^2}} = k$$

da cui

$$-x = k\sqrt{1+\dot{x}^2}$$

e cioè

$$\dot{x}^2 = \frac{x^2}{k^2} - 1 \tag{4.1}$$

per una opportuna costante $k \in \mathbb{R}$. Distinguiamo ora tre casi possibili.

Caso 1. Sia $\dot{x}_* > 0$ in $[a,b]$. Allora, separando le variabili nella (4.1), si ha

$$\frac{k\ dx}{\sqrt{x^2-k^2}} = dt.$$

Il membro a sinistra ha primitiva $k\cosh^{-1}(x/k) + c$, quindi x_* è della forma

$$x_*(t) = k\cosh\left(\frac{t+c}{k}\right).$$

Una curva di questo tipo è detta *catenaria*. Le costanti c e k sono determinate dai vincoli iniziali.[4]

Caso 2. Sia $\dot{x}_* < 0$ in $[a,b]$. Allora, in modo simile al caso precedente, si trova

$$x_*(t) = \tilde{k}\cosh\left(\frac{t+\tilde{c}}{\tilde{k}}\right)$$

per opportune costanti $\tilde{c}$ e $\tilde{k}$.

Caso 3. Rimane la possibilità che $\dot{x}_* < 0$ in $[a,r)$ e $\dot{x}_* > 0$ in $(r,b]$ per un certo $a < r < b$. Si ha allora, ragionando in modo analogo ai due casi precedenti sugli intervalli corrispondenti,

$$x_*(t) = k\cosh\left(\frac{t+c}{k}\right) \text{ in } [a,r) \quad \text{e} \quad x_*(t) = \tilde{k}\cosh\left(\frac{t+\tilde{c}}{\tilde{k}}\right) \text{ in } (r,b]$$

da cui, ricordando che x_* deve essere di classe C^2, $c = \tilde{c}$ e $k = \tilde{k}$, ovvero x_* è ancora un (unico) arco di catenaria.

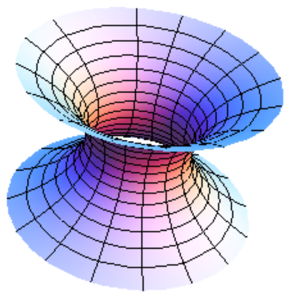

Figura 2.1. Superficie generata dalla catenaria.

[4]Attenzione: non tutte le scelte dei vincoli $x(a) = A$ e $x(b) = B$ definisco una catenaria. Non ci occuperemo di approfondire questo fatto: il lettore interessato può consultare [**2**].

Remark 2.12. I calcoli che abbiamo svolto dimostrano che *se esiste un minimo locale debole* $x_* \in C^2([a,b];\mathbb{R})$ *per il funzionale di minima superficie che è positivo su* $[a,b]$ *allora* x_* *è un arco di catenaria.* Questo non dimostra in alcun modo che questo minimo x_* esista. Per dimostrarne l'esistenza serve un argomento *ad hoc* specifico per questo problema di minimizzazione, che noi non trattiamo.[5].

[5]Il lettore intraprendente potrebbe sperimentare "a mano" la veridicità dei calcoli svolti in questa sezione: si potrebbe tracciare sulla carta il profilo di una piccola catenella sospesa, oppure, con l'aiuto di un po' di acqua e sapone e di fil di ferro, si potrebbe costruire una superficie di bolla di sapone tra due circonferenze. Alcuni di questi esperimenti sono spiegati in [**3**, Capitolo 7, §11]

CAPITOLO 3

La condizione di Legendre

Nel Capitolo 2 abbiamo sviluppato una teoria del primo ordine per il funzionale J definito in (1.1), individuando una condizione necessaria, la (Eulero-Lagrange), affinché un curva ammissibile fosse un estremante locale per il problema fondamentale (P). In modo analogo al Principio di Fermat incontrato nel Teorema 1.1, la (Eulero-Lagrange) non distingue tra minimo e massimo. In questo Capitolo cercheremo di individuare nuove condizioni per determinare, in maniera simile al caso finito dimensionale, di che natura sia l'estremante locale per il problema (P).

1. Il caso finito dimensionale

Il Principio di Fermat (Teorema (1.1)) fornisce solamente una condizione necessaria affinché un punto sia un estremante locale per una funzione differenziabile. I seguenti due risultati forniscono condizioni necessarie e sufficienti affinché un punto sia un minimo locale di una funzione di classe C^2 (risultati analoghi valgono per i massimi locali).

Teorema 3.1 (Condizioni necessarie di estremalità). *Sia Ω un aperto di $\mathbb{R}^n$ e sia $f\colon \Omega \to \mathbb{R}$ una funzione di classe C^2. Sia $x_* \in \Omega$ un punto di minimo locale della funzione f. Allora valgono*

(i) $Df(x_*) = 0$;
(ii) $Hf(x_*) \geq 0$, *cioè la matrice Hessiana della funzione f è semidefinita positiva.*

Teorema 3.2 (Condizioni sufficienti di estremalità). *Sia Ω un aperto di $\mathbb{R}^n$ e sia $f\colon \Omega \to \mathbb{R}$ una funzione di classe C^2. Sia $x_* \in \Omega$. Supponiamo che*

(i) $Df(x_*) = 0$;
(ii) $Hf(x_*) > 0$, *cioè la matrice Hessiana della funzione f è definita positiva.*

Allora $x_* \in \Omega$ è un punto di minimo locale della funzione f.

La dimostrazione di questi risultati è una conseguenza dello sviluppo in serie di Taylor per le funzioni di classe C^2. Per $n = 1$, la matrice Hessiana della funzione f nei due risultati precedenti è semplicemente la derivata seconda f''. Pertanto, quando $n = 1$, se $f'(x_*) = 0$, allora i Teoremi 3.1 e 3.2 affermano semplicemente che

$$x_* \text{ è un minimo locale} \implies f''(x_*) \geq 0$$

e

$$f''(x_*) > 0 \implies x_* \text{ è un minimo locale.}$$

Notare che, nella seconda implicazione, la stretta positività è necessaria: si consideri ad esempio la funzione $f(x) = x^3$ nel punto $x = 0$.

Vediamo ora come estendere queste condizioni al caso infinito dimensionale.

2. La condizione necessaria di Legendre

Questa sezione è dedicata alla dimostrazione del seguente risultato.

Teorema 3.3 (Condizione di Legendre, 1786). *Supponiamo che L sia di classe C^3, e sia $x_* \in C^2([a,b];\mathbb{R})$ un minimo locale debole per il problema* (P). *Allora*

$$\text{(Legendre)} \qquad \frac{\partial^2 L}{\partial v^2}(t, x_*(t), \dot{x}_*(t)) \geq 0 \quad \forall t \in [a,b].$$

Dimostrazione. Consideriamo la funzione g definita in (1.3). Siccome ora L è di classe C^3, g' è differenziabile. Per ipotesi x_* è un minimo locale per il funzionale J, quindi $\lambda = 0$ è un minimo locale per la funzione g. Per il Teorema 3.1 abbiamo che

$$\left.\frac{d^2}{d\lambda^2} g(\lambda)\right|_{\lambda=0} \geq 0.$$

Dai calcoli svolti nella dimostrazione del Teorema 2.2, abbiamo che, per il Teorema di Schwarz,

$$\begin{aligned}
\left.\frac{d^2}{d\lambda^2} g(\lambda)\right|_{\lambda=0} &= \left.\frac{d}{d\lambda} \int_a^b \frac{\partial L}{\partial x}(t, x_* + \lambda w, \dot{x}_* + \lambda \dot{w}) \cdot w + \frac{\partial L}{\partial v}(t, x_* + \lambda w, \dot{x}_* + \lambda \dot{w}) \cdot \dot{w} \; dt\right|_{\lambda=0} = \\
&= \int_a^b \frac{\partial^2 L}{\partial x^2}(t, x_*, \dot{x}_*) \cdot w^2 + 2\frac{\partial^2 L}{\partial v \partial x}(t, x_*, \dot{x}_*) \cdot w\dot{w} + \frac{\partial^2 L}{\partial v^2}(t, x_*, \dot{x}_*) \cdot \dot{w}^2 \; dt = \text{ (int. per parti)} \\
&= \int_a^b \frac{\partial^2 L}{\partial x^2}(t, x_*, \dot{x}_*) \cdot w^2 + \frac{\partial^2 L}{\partial v^2}(t, x_*, \dot{x}_*) \cdot \dot{w}^2 \; dt + \\
&\quad \underbrace{+ 2\frac{\partial^2 L}{\partial v \partial x}(t, x_*, \dot{x}_*) \cdot \left.\frac{w^2}{2}\right|_{t=a}^{t=b}}_{=0 \text{ per la (1.2)}} - \int_a^b \frac{d}{dt}\left(\frac{\partial^2 L}{\partial v \partial x}\right)(t, x_*, \dot{x}_*) \cdot w^2 \; dt = \\
&= \int_a^b P(t)\dot{w}^2 + Q(t) w^2 \; dt
\end{aligned}$$

dove abbiamo posto

$$P(t) = \frac{\partial^2 L}{\partial v^2}(t, x_*, \dot{x}_*) \tag{2.1}$$

e

$$Q(t) = \frac{\partial^2 L}{\partial x^2}(t, x_*, \dot{x}_*) - \frac{d}{dt}\left(\frac{\partial^2 L}{\partial v \partial x}\right)(t, x_*, \dot{x}_*). \tag{2.2}$$

Abbiamo dunque che, per ogni variazione $w \in C^2([a, b]; \mathbb{R})$ tale che $w(a) = w(b) = 0$, si deve avere

$$\int_a^b P(t)\dot{w}^2 + Q(t) w^2 \; dt \geq 0.$$

Fissiamo $\varepsilon > 0$ e un intervallo $[c, d] \subseteq [a, b]$. Scegliamo una variazione w Lipschitziana su $[a, b]$ tale che $|\dot{w}(t)| = 1$ q.o. su $[c, d]$, $\|w\| \leq \varepsilon$ e $w \equiv 0$ fuori dell'intervallo scelto $[c, d]$.

Figura 3.1. La variazione Lipschitziana w.

Allora, per tale variazione w,

$$0 \leq \int_a^b P(t)\dot{w}^2 + Q(t) w^2 \; dt \leq \int_c^d P(t) \cdot 1 + Q(t)\varepsilon^2 \; dt \tag{2.3}$$

e passando al limite per $\varepsilon \to 0$ si trova $\int_c^d P(t) \; dt \geq 0$. Siccome l'intervallo $[c, d]$ era arbitrario, deve essere $P \geq 0$ su $[a, b]$, cioè

$$\frac{\partial^2 L}{\partial v^2}(t, x_*(t), \dot{x}_*(t)) \geq 0$$

per ogni $t \in [a, b]$. Per mostrare la (Legendre), dobbiamo mostrare che la scelta di variazioni Lipschitziane (non necessariamente C^1) non influenza il nostro ragionamento. È chiaro che i conti fatti sopra per arrivare alla (2.3) valgono anche nel caso in cui w sia solamente Lipschitziana, perché le funzioni

Lipschitziane sono assolutamente continue e quindi permettono di integrare per parti. Per concludere rimane da dimostrare solamente che, se x_* è un minimo locale debole relativamente alle funzioni C^2, allora x_* è anche minimo locale debole relativamente alle funzioni Lipschitziane. La dimostrazione di questa affermazione è riportata nella Appendice A (per un riferimento, si veda [**2**, Esercizio 21.13]). $\square$

3. Condizione sufficiente per intervalli piccoli. La disuguaglianza di Wirtinger

Nella sezione precedente abbiamo dimostrato una condizione necessaria di minimalità, la (Legendre), del tutto analoga a quella incontrata nel caso finito dimensionale nel Teorema 3.1. Guidati dall'analogia tra le disuguaglianze, ci poniamo la seguente domanda: *è vero che, sotto l'ipotesi che L sia di classe C^3 e x_* sia soluzione della* (Eulero-Lagrange)*, la disuguaglianza*

$$\frac{\partial^2 L}{\partial v^2}(t, x_*(t), \dot{x}_*(t)) > 0 \quad \forall t \in [a, b]$$

è una condizione sufficiente affinché la funzione x_* sia un minimo locale debole per il problema (P)?

Sappiamo che il Teorema 3.2 deriva dallo sviluppo in serie di Taylor: proviamo a rispondere alla domanda che ci siamo posti cercando di applicare questo sviluppo alla funzione g definita nella (1.3).

Osserviamo che, per un certo $\bar{\lambda} \in (0, 1)$, per lo sviluppo di Taylor si ha

$$J(x_* + w) - J(x_*) = g(1) - g(0) = \frac{1}{2} g''(\bar{\lambda})$$

perché $g'(0) = 0$, essendo x_* una soluzione della (Eulero-Lagrange). Avevamo già calcolato g'' in precedenza, e possiamo scrivere

$$\begin{aligned}
g''(\bar{\lambda}) = & \int_a^b \frac{\partial^2 L}{\partial x^2}(t, x_* + \bar{\lambda} w, \dot{x}_* + \bar{\lambda} \dot{w}) \cdot w^2 + 2 \frac{\partial^2 L}{\partial v \partial x}(t, x_* + \bar{\lambda} w, \dot{x}_* + \bar{\lambda} \dot{w}) \cdot w \dot{w} + \\
& + \frac{\partial^2 L}{\partial v^2}(t, x_* + \bar{\lambda} w, \dot{x}_* + \bar{\lambda} \dot{w}) \cdot \dot{w}^2 \, dt = \\
= & \int_a^b \frac{\partial^2 L}{\partial x^2}(t, x_*, \dot{x}_*) \cdot w^2 + 2 \frac{\partial^2 L}{\partial v \partial x}(t, x_*, \dot{x}_*) \cdot w \dot{w} + \frac{\partial^2 L}{\partial v^2}(t, x_*, \dot{x}_*) \cdot \dot{w}^2 \, dt + \Delta(L) = \\
= & \int_a^b P(t) \dot{w}^2 + Q(t) w^2 \, dt + \Delta(L)
\end{aligned}$$

dove $\Delta(L)$ è il termine di errore dato da

$$\begin{aligned}
\Delta(L) :=\ & g''(\bar{\lambda}) - g''(0) = \\
= & \int_a^b \left(\frac{\partial^2 L}{\partial x^2}(t, x_* + \bar{\lambda} w, \dot{x}_* + \bar{\lambda} \dot{w}) - \frac{\partial^2 L}{\partial x^2}(t, x_*, \dot{x}_*) \right) \cdot w^2 + \\
& + 2 \left(\frac{\partial^2 L}{\partial v \partial x}(t, x_* + \bar{\lambda} w, \dot{x}_* + \bar{\lambda} \dot{w}) - \frac{\partial^2 L}{\partial v \partial x}(t, x_*, \dot{x}_*) \right) \cdot w \dot{w} + \\
& + \left(\frac{\partial^2 L}{\partial v^2}(t, x_* + \bar{\lambda} w, \dot{x}_* + \bar{\lambda} \dot{w}) - \frac{\partial^2 L}{\partial v^2}(t, x_*, \dot{x}_*) \right) \cdot \dot{w}^2 \, dt
\end{aligned}$$

(abbiamo semplicemente aggiunto e tolto la quantità $g''(0)$).

Stimiamo il termine di errore $\Delta(L)$. Siccome la funzione L è di classe C^3 per ipotesi, le derivate seconde di L sono di classe C^1, e pertanto sono funzioni localmente Lipschitziane. Fissiamo dunque una variazione w sufficientemente piccola in norma C^1, ad esempio tale che $\max\{ \|w\|, \|\dot{w}\| \} \leq 1$. Per le entrate delle derivate seconde di L abbiamo chiaramente la stima

$$\max \left\{ \left| x_* + \bar{\lambda} w \right|, \quad \left| \dot{x}_* + \bar{\lambda} \dot{w} \right| \right\} \leq \|x_*\| + \|\dot{x}_*\| + 1$$

e ovviamente $|t| \leq \max\{ |a|, |b| \}$ per la variabile di integrazione. Sia allora

$$R := \max\{ |a|, |b| \} + \|x_*\| + \|\dot{x}_*\| + 1$$

e sia $K := \bar{B}(R)$ la palla chiusa di centro l'origine e raggio R in $\mathbb{R}^3$. Sia M una costante di Lipschitz sul compatto K per le derivate seconde della funzione L (si può scegliere il massimo tra le costanti di

Lipschitz su K delle funzioni $\frac{\partial^2 L}{\partial x^2}$, $\frac{\partial^2 L}{\partial v \partial x}$ e $\frac{\partial^2 L}{\partial v^2}$). Allora, per ogni $t \in [a,b]$, si ha

$$\begin{aligned}|\Delta(L)| &\leq \int_a^b \left|\frac{\partial^2 L}{\partial x^2}[\bar\lambda] - \frac{\partial^2 L}{\partial x^2}[0]\right| \cdot w^2 + 2\left|\frac{\partial^2 L}{\partial v \partial x}[\bar\lambda] - \frac{\partial^2 L}{\partial v \partial x}[0]\right| \cdot |w||\dot w| + \left|\frac{\partial^2 L}{\partial v^2}[\bar\lambda] - \frac{\partial^2 L}{\partial v^2}[0]\right| \cdot \dot w^2 \; dt \leq \\ &\leq M \int_a^b \big(\left|x_*(t) + \bar\lambda w(t) - x_*(t)\right| + \left|\dot x_*(t) + \bar\lambda \dot w(t) - \dot x_*(t)\right|\big) \cdot \big(w^2(t) + 2|w(t)||\dot w(t)| + \dot w^2(t)\big) \; dt \leq \\ &\leq M\left(\|w\| + \|\dot w\|\right) \int_a^b w^2(t) + 2|w(t)||\dot w(t)| + \dot w^2(t) \; dt.\end{aligned}$$

Affermiamo che esiste una costante $C > 0$, dipendente solo dall'intervallo $[a,b]$, tale che

$$\int_a^b w^2(t) + 2|w(t)||\dot w(t)| + \dot w^2(t) \; dt \leq C \int_a^b \dot w^2(t) \; dt.$$

Infatti, ricordato che $2\alpha\beta \leq \alpha^2 + \beta^2$ per ogni $\alpha, \beta \in \mathbb{R}$ e che $w(t) = \int_a^b \dot w(s) \; ds$, essendo $w(a) = 0$, si può stimare (in modo non ottimale)

$$\begin{aligned}\int_a^b w^2(t) + 2|w(t)||\dot w(t)| + \dot w^2(t) \; dt &\leq 2\int_a^b w^2(t) + \dot w^2(t) \; dt \leq \\ &\leq 2\int_a^b \left(\int_a^t \dot w(s) \; ds\right)^2 + \dot w^2(t) \; dt \leq \quad \text{(dis. di Cauchy–Schwarz)} \\ &\leq 2\int_a^b (t-a)\left(\int_a^t \dot w^2(s) \; ds\right) + \dot w^2(t) \; dt \leq \\ &\leq 2\int_a^b (b-a)\left(\int_a^b \dot w^2(s) \; ds\right) + \dot w^2(t) \; dt = \\ &= 2\left((b-a)^2 + 1\right)\int_a^b \dot w^2(t) \; dt\end{aligned}$$

da cui $C = 2\left((b-a)^2 + 1\right)$.

Pertanto abbiamo che

$$|\Delta(L)| \leq CM\left(\|w\| + \|\dot w\|\right) \int_a^b \dot w^2(t) \; dt.$$

Poniamo

$$\delta := CM\left(\|w\| + \|\dot w\|\right). \tag{3.1}$$

Si noti che δ è sufficientemente piccolo, pur di scegliere la variazione w molto piccola in norma C^1.

In conclusione, abbiamo che

$$\begin{aligned}J(x_* + w) - J(x_*) = \frac{1}{2}g''(\bar\lambda) = \frac{1}{2}\int_a^b P(t)\dot w^2 + Q(t)w^2 \; dt + \Delta(L) &\geq \\ \geq \frac{1}{2}\int_a^b P(t)\dot w^2 + Q(t)w^2 \; dt - \delta\int_a^b \dot w^2(t) \; dt &= \\ = \frac{1}{2}\int_a^b (P(t) - \delta)\dot w^2 + Q(t)w^2 \; dt.&\end{aligned}$$

Adesso supponiamo che $\frac{\partial^2 L}{\partial v^2}(t, x_*, \dot x_*) > 0$ sull'intervallo $[a,b]$, come abbiamo ipotizzato nella domanda che ci siamo posti all'inizio di questa sezione. Possiamo allora scegliere δ sufficientemente piccolo in modo che $P(t) - \delta > 0$ su $[a,b]$ (questo è possibile, come abbiamo osservato in occasione della (3.1), pur di scegliere la variazione w molto vicina alla funzione nulla in norma C^1).

Scelto δ in questo modo, se raccogliamo a fattore il termine positivo $(P(t) - \delta)$ e riscriviamo ciò che rimane nell'ultimo integrale della disuguaglianza sopra come un quadrato, allora abbiamo mostrato che $J(x_* + w) - J(x_*) \geq 0$ per ogni variazione w sufficientemente vicina alla funzione nulla in norma C^1, cioè che x_* è un minimo locale debole per il funzionale J, come volevamo.

Per far comparire il quadrato, usiamo la seguente idea ingegnosa, dovuta a Legendre stesso. Sia ψ una qualunque funzione derivabile definita sull'intervallo $[a,b]$. Siccome la variazione w si annulla agli estremi dell'intervallo, per il Teorema fondamentale del calcolo integrale si ha

$$\int_a^b (\psi w^2)' \, dt = \psi w^2 \Big|_{t=a}^{t=b} = 0.$$

Allora si può scrivere

$$\begin{aligned}\int_a^b (P(t)-\delta)\dot{w}^2 + Q(t)w^2 \, dt &= \int_a^b (P(t)-\delta)\dot{w}^2 + Q(t)w^2 + (\psi w^2)' \, dt = \\ &= \int_a^b (P(t)-\delta)\dot{w}^2 + Q(t)w^2 + \dot{\psi}w^2 + 2\psi w\dot{w} \, dt = \\ &= \int_a^b (P(t)-\delta) \underbrace{\left(\dot{w}^2 + \frac{Q(t)+\dot{\psi}}{P(t)-\delta} w^2 + \frac{2\psi}{P(t)-\delta} w\dot{w} \right)}_{(\star)} dt.\end{aligned}$$

L'espressione $(\star)$ tra parentesi tonde diventa un quadrato di binomio se e solo se

$$\frac{Q(t)+\dot{\psi}}{P(t)-\delta} = \left(\frac{\psi}{P(t)-\delta} \right)^2$$

ovvero

$$\dot{\psi} = \frac{\psi^2}{P(t)-\delta} - Q(t) \tag{3.2}$$

per ogni $t \in [a,b]$.

A quanto pare, siamo arrivati molto vicini alla soluzione del problema che ci siamo posti all'inizio di questa sezione: se esiste un funzione ψ sull'intervallo $[a,b]$ che risolve l'equazione differenziale (3.2), allora x_* è un minimo locale debole per il funzionale J. Più precisamente, i calcoli e le considerazioni sviluppati fino a qui dimostrano il seguente risultato.

Teorema 3.4. *Supponiamo che la funzione L sia di classe C^3, e sia $x_* \in C^2([a,b];\mathbb{R})$ soluzione della* (Eulero-Lagrange)*, cioè un estremo locale per il problema* (P)*. Se x_* è tale che*

$$\frac{\partial^2 L}{\partial v^2}(t, x_*(t), \dot{x}_*(t)) > 0 \quad \forall t \in [a,b] \tag{Legendre stretta}$$

e se esiste una funzione $\psi\colon [a,b] \to \mathbb{R}$ tale che

$$\psi'(t) = \frac{\psi^2(t)}{P(t)-\delta} - Q(t) \quad \forall t \in [a,b], \tag{D}$$

con P e Q definite come in (2.1) *e in* (2.2) *e δ sufficientemente piccolo, allora x_* è un minimo locale debole per il funzionale J.*

Rimane aperto un problema, un ultimo ostacolo prima della completa estensione, al caso infinito dimensionale, del Teorema 3.2: esiste sempre una soluzione ψ all'equazione differenziale (D) *definita su tutto intervallo* $[a,b]$? La risposta è no, non sempre esiste una tale funzione.[1]

Per capire meglio il ruolo giocato dall'equazione (D) nel Teorema 3.4, ricordiamo il seguente classico risultato sulle equazioni differenziali ordinarie (si veda ad esempio [**4**, Capitolo 11], oppure [**5**, Teorema 5.2, pag. 68]).

Teorema 3.5 (Cauchy–Lipschitz, Picard–Lindelöf). *Siano $\Omega \subseteq \mathbb{R}_t \times \mathbb{R}^d_u$ un insieme aperto, $(t_0,u_0) \in \Omega$ e sia $f \in C^0(\Omega;\mathbb{R}^d)$ una funzione localmente di Lipschitz nella variabile u. Allora esiste $\varepsilon > 0$ tale che il problema*

$$\begin{cases} u'(t) = f(t,u(t)) \\ u(t_0) = u_0 \end{cases} \tag{PC}$$

[1] Legendre stesso incappò nell'errore di ritenere che una soluzione globale all'equazione (D) esistesse sempre, e venne per questo duramente criticato da Lagrange.

ha una soluzione unica $u \in C^1([x_0-\varepsilon, x_0+\varepsilon];\mathbb{R}^d)$.

Osserviamo che, nell'equazione differenziale (D), è

$$f(t,u) = \frac{u^2}{P(t)-\delta} - Q(t), \tag{3.3}$$

funzione localmente Lipschitziana nella variabile u che dipende in modo continuo dal parametro δ. Allora, per il Teorema 3.5, fissato un valore a piacere per la funzione ψ nell'estremo a dell'intervallo e fissato δ, esiste $\varepsilon > 0$ tale che esiste un'unica soluzione $\psi \equiv \psi_\delta$ all'equazione (D) di classe C^1 definita sull'intervallo $[a, a+\varepsilon]$. Siccome la funzione f nella (3.3), come visto, dipende in modo continuo da δ, se sappiamo che l'equazione (D) con la scelta di $\delta = 0$ ammette una soluzione globale $\psi \equiv \psi_0$, allora[2] esiste $\hat{\delta} > 0$ tale che esiste un'unica soluzione $\psi \equiv \psi_\delta$ all'equazione (D) di classe C^1 definita sull'intervallo $[a, a+\varepsilon]$ per ogni scelta di $\delta \in [0, \hat{\delta}]$. *Quindi, in altre parole, pur di prendere l'intervallo* $[a,b]$ *di piccole dimensioni, una soluzione* ψ *dell'equazione* (D) *esiste globalmente uniformemente per piccoli valori del parametro* δ. Possiamo riassumere quanto abbiamo visto nel seguente risultato.

Corollario 3.6 (Condizione sufficiente di Legendre per intervalli piccoli). *Supponiamo che la funzione* L *sia di classe* C^3*, e sia* $x_* \in C^2([a,b];\mathbb{R})$ *soluzione della* (Eulero-Lagrange)*, cioè un estremo locale per il problema* (P)*. Se* x_* *è tale che*

$$\frac{\partial^2 L}{\partial v^2}(t, x_*(t), \dot{x}_*(t)) > 0 \quad \forall t \in [a,b]$$

e se esiste una funzione $\psi\colon [a,b] \to \mathbb{R}$ *tale che*

$$\psi'(t) = \frac{\psi^2(t)}{P(t)} - Q(t) \quad \forall t \in [a,b],$$

con P *e* Q *definite come in* (2.1) *e in* (2.2)*, allora* x_* *è un minimo locale debole per il funzionale* J.

Dimostrazione. Si segue il ragionamento che abbiamo sviluppato sopra per arrivare al Teorema 3.4. Per concludere, è sufficiente scegliere un $\hat{\delta} > 0$ sufficientemente piccolo in modo che

(1) $P(t) - \delta > 0$ per ogni $t \in [a,b]$ e per ogni $\delta \in [0, \hat{\delta}]$;
(2) la soluzione $\psi \equiv \psi_\delta$ del problema (D) esista su tutto l'intervallo $[a,b]$ per ogni $\delta \in [0,\hat{\delta}]$.

L'esistenza di un tale $\hat{\delta} > 0$ è assicurata dal Teorema 3.5. □

Il Corollario 3.6 rappresenta l'analogo del Teorema 3.2 nel caso infinito dimensionale ed è, per le possibilità degli strumenti che abbiamo a nostra disposizione, il miglior risultato possibile in questa direzione. L'esistenza soltanto locale di una soluzione ψ all'equazione (D), inoltre, porta con sé una nuova idea, realistica sotto il profilo delle applicazioni pratiche: *in generale, gli estremanti possono essere ottimali solo a breve termine, cioè senza essere ottimali globalmente sull'intero intervallo di definizione.* Vogliamo concludere questa sezione mostrando qualche esempio in cui si verifica questa situazione.

Esempio 3.7 (La disuguaglianza di Wirtinger, [**2**, Esercizio 14.10]). Sia $T > 0$ fissato. Vogliamo studiare la disuguaglianza

$$\int_0^T x^2(t)\, dt \le \int_0^T \dot{x}^2(t)\, dt \tag{Wirtinger}$$

per le funzioni $x \in C^2([0,T])$ tali che $x(0) = x(T) = 0$.

La funzione identicamente nulla soddisfa chiaramente la (Wirtinger), ed ha la proprietà di annullare la quantità $\int_0^T \dot{x}^2(t) - x^2(t)\, dt$. Possiamo allora riformulare il problema dato dalla (Wirtinger) nel senso seguente: *per quali valori del parametro* $T > 0$ *la funzione identicamente nulla* $x_* \equiv 0$ *è soluzione del problema*

$$\min\ J(x) = \int_0^T \dot{x}^2(t) - x^2(t)\, dt \ :\ x \in C^2([0,T];\mathbb{R}),\ x(0) = 0,\ x(T) = 0\ ? \tag{PW}$$

[2] Questa risultato di esistenza uniforme rispetto al parametro δ della soluzione ψ all'equazione (D) è una conseguenza della stabilità per convergenza uniforme del Teorema 3.5: piccole variazioni della funzione f (ed eventualmente del dato iniziale (t_0, u_0)) comportano piccole variazioni sulla soluzione u, cioè u dipende in modo continuo dai dati iniziali del problema (PC).

Studiamo il problema (PW) distinguendo due casi.

Caso 1. Se $T \leq 1$, allora $x_* \equiv 0$ è un minimo globale debole per il problema (PW). Sia infatti $T \leq 1$. Allora, osservato che

$$x(t) = \int_0^t \dot{x}(s)\ ds$$

per ogni $t \in [0, T]$, essendo $x(0) = 0$, si trova

$$x^2(t) = \left(\int_0^t \dot{x}(s)\ ds \right)^2 \leq t \int_0^t \dot{x}^2(s)\ ds \leq T \int_0^T \dot{x}^2(s)\ ds$$

per la disuguaglianza di Cauchy–Schwarz. Allora, integrando tra 0 e T, troviamo

$$\int_0^T x^2(t)\ dt \leq T^2 \int_0^T \dot{x}^2(t)\ dt \leq \int_0^T \dot{x}^2(t)\ dt$$

perché $T \leq 1$, e quindi $J(x) \geq 0 = J(x_*)$ per ogni funzione x ammissibile, cioè la funzione nulla è un minimo globale debole per il funzionale J.

Caso 2. Se $T \geq 4$, allora $x_* \equiv 0$ non è un minimo globale debole per il problema (PW). Per mostrare questo fatto osserviamo preliminarmente che il funzionale J soddisfa la proprietà

$$J(\lambda x) = \lambda^2 J(x) \quad \forall \lambda \in \mathbb{R} \tag{3.4}$$

per ogni funzione $x \in C^1([0, T]; \mathbb{R})$. Infatti

$$J(\lambda x) = \int_0^T (\lambda \dot{x})^2 - (\lambda x)^2\ dt = \lambda^2 \int_0^T \dot{x}^2 - x^2\ dt = \lambda^2 J(x).$$

Dalla proprietà (3.4) segue che

$$x_* \equiv 0 \text{ è minimio locale debole} \iff x_* \equiv 0 \text{ è minimio globale debole.} \tag{3.5}$$

Mostriamo l'implicazione interessante. Supponiamo per assurdo che $x_* \equiv 0$ non sia un minimo globale. Allora esiste una funzione ammissibile y tale che $J(y) < J(x_*) = 0$. Siccome $x_* \equiv 0$ è un minimo locale debole, esiste $\varepsilon > 0$ tale che $J(x_*) \leq J(x)$ per ogni funzione x ammissibile tale che $\|x - x_*\| < \varepsilon$ e $\|\dot{x} - \dot{x}_*\| < \varepsilon$. Scegliamo allora $\lambda > 0$ tale che $\|\lambda y - x_*\| = \lambda \|y\| < \varepsilon$ e $\|\lambda \dot{y} - \dot{x}_*\| = \lambda \|\dot{y}\| < \varepsilon$ (è sufficiente prendere $\lambda < \min\{ \frac{\varepsilon}{\|y\|}, \frac{\varepsilon}{\|\dot{y}\|} \}$). Allora, con questa scelta, abbiamo che $0 = J(x_*) \leq J(\lambda y) = \lambda^2 J(y)$, da cui $J(y) \geq 0$, una contraddizione.

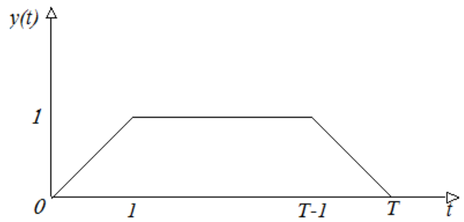

Figura 3.2. La funzione Lipschitziana y.

Grazie all'equivalenza (3.5), per concludere ci basta mostrare che esiste una funzione $x \in C^2([0, T]; \mathbb{R})$ ammissibile tale che $J(x) < 0$. In riferimento alla Figura 3.2, consideriamo infatti la funzione Lipschitziana y tale che $y(0) = 0$ e

$$y'(t) = \begin{cases} 1 & 0 \leq t \leq 1 \\ 0 & 1 < t < T - 1 \\ -1 & T - 1 \leq t \leq T, \end{cases}$$

ovvero

$$y(t) = \begin{cases} t & 0 \leq t \leq 1 \\ 1 & 1 < t < T - 1 \\ T - t & T - 1 \leq t \leq T. \end{cases}$$

Allora, per calcolo diretto,

$$J(y) = \int_0^T \dot{y}^2\ dt - \int_0^T y^2\ dt = 2 - \left(2\int_0^1 t^2\ dt + \int_1^{T-1} 1\ dt\right) = \frac{10}{3} - T < 0.$$

Osserviamo poi che, fissato $\varepsilon > 0$, per il Teorema A.4 è possibile trovare un polinomio p con $p(0) = p(T) = 0$ tale che $\|y - p\| < \varepsilon$ e $|J(y) - J(p)| < \varepsilon$. Allora, pur di scegliere $\varepsilon > 0$ sufficientemente piccolo, esiste una funzione x (un polinomio) ammissibile per il problema (PW) tale che $J(x) < 0$. In questo caso, pertanto, $x_* \equiv 0$ non è un minimo globale debole.

Per quanto abbiamo visto, la soluzione nulla $x_* \equiv 0$ è un minimo debole globale per il funzionale J definito in (PW) finché $T < \bar{T}$, per un certo $\bar{T} \in (1, 4)$. Quindi, come mostra questo esempio, gli estremanti possono essere ottimali solo a breve termine, cioè senza essere ottimali globalmente sull'intero intervallo di definizione. Questo può accadere proprio perché l'esistenza solamente locale della soluzione dell'equazione (D) nel Teorema 3.4 è un problema cruciale per l'efficacia della (Legendre stretta) come condizione sufficiente per la minimalità locale. Osserviamo infatti che, per ogni $T \geq 0$, la funzione nulla $x_* \equiv 0$ è un estremante locale, perché la Lagrangiana è

$$L(t, x, v) \equiv L(x, v) = v^2 - x^2,$$

per cui la (Eulero-Lagrange) per il problema (PW) è

$$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} = \frac{d}{dt}(2v) - (-2x) = 2\ddot{x} + 2x$$

ovvero $\ddot{x} + x = 0$, che è soddisfatta dalla funzione $x_* \equiv 0$. Inoltre, abbiamo che

$$\frac{\partial^2 L}{\partial v^2}(x_*, \dot{x}_*) = 2 > 0,$$

e quindi la (Legendre stretta) è soddisfatta dalla funzione $x_* \equiv 0$ per ogni $T \geq 0$. Ma, come abbiamo appena visto, questo non implica che $x_* \equiv 0$ sia un minimo debole locale (ovvero globale, per la (3.5)) per ogni $T \geq 0$. Infatti, il controesempio trovato quando $T \geq 4$ mostra che la tesi del Teorema 3.4 è falsa se non si conosce, a priori, l'esistenza di una soluzione per la (D). Precisamente, in questo caso, abbiamo che

$$P(t) = \frac{\partial^2 L}{\partial v^2}(x_*, \dot{x}_*) = 2$$

e

$$Q(t) = \frac{\partial^2 L}{\partial x^2}(x_*, \dot{x}_*) - \frac{d}{dt}\left(\frac{\partial^2 L}{\partial v \partial x}\right)(x_*, \dot{x}_*) = -2$$

per ogni $t \in [0, T]$, per cui la (D), posto $\delta = 0$ in riferimento al Corollario 3.6, porta al problema di determinare una soluzione per l'equazione

$$\psi' = \frac{\psi^2}{2} + 2.$$

Si verifica immediatamente che una soluzione è $\psi(t) = 2\tan(t)$, che risulta essere ben definita sull'intervallo $[0, \frac{\pi}{2})$. Dunque il Corollario 3.6 dimostra che la funzione $x_* \equiv 0$ è un minimo locale debole per il problema (PW) finché $T < \frac{\pi}{2}$. In particolare, otteniamo la stima $\bar{T} \in [\frac{\pi}{2}, 4)$, leggermente migliore della precedente.

Esempio 3.8 ([**2**, Esempio 14.9]). Sia $T > 0$ fissato. Studiamo il problema (P) per il funzionale

$$J(x) = \int_0^T \frac{1}{2}\dot{x}^2 + t\dot{x}\sin(x)\ dt \tag{3.6}$$

sull'intervallo $[0, T]$ con vincoli $x(0) = x(T) = 0$ al variare del parametro T.

Per la (Eulero-Lagrange) abbiamo che

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} &= \frac{d}{dt}(\dot{x} + t\sin(x)) - t\dot{x}\cos(x) = \\ &= \ddot{x} + \sin(x) + t\dot{x}\cos(x) - t\dot{x}\cos(x) = \\ &= \ddot{x} + \sin(x)\end{aligned}$$

da cui $\ddot{x}+\sin(x) = 0$. Osserviamo che la funzione nulla $x_* \equiv 0$ soddisfa questa equazione ed è ammissibile: studiamo questo estremante particolare.

Osserviamo che la (Legendre stretta) è soddisfatta, perché $\frac{\partial^2 L}{\partial v^2}(x_*, \dot{x}_*) = 1 > 0$ per ogni $T > 0$. Tuttavia, da questo non possiamo concludere che $x_* \equiv 0$ sia un minimo debole locale: vogliamo applicare il Corollario 3.6, perciò dobbiamo di studiare l'equazione (D) con $\delta = 0$. Osserviamo che

$$P(t) = \frac{\partial^2 L}{\partial v^2}(t, x_*, \dot{x}_*) = 1,$$

mentre

$$\begin{aligned} Q(t) &= \frac{\partial^2 L}{\partial x^2}(t, x_*, \dot{x}_*) - \frac{d}{dt}\left(\frac{\partial^2 L}{\partial v \partial x}\right)(t, x_*, \dot{x}_*) = \\ &= \frac{\partial}{\partial x}(t\dot{x}\cos(x)) - \frac{d}{dt}(t\cos(x))\Big|_{x=x_*} = \\ &= -t\dot{x}\sin(x) - \cos(x) + t\dot{x}\sin(x)\big|_{x=x_*} = \\ &= -\cos(x_*) = -1 \end{aligned}$$

essendo $x_* \equiv 0$. Quindi la (D) porta all'equazione

$$\psi' = \psi^2 + 1$$

che, similmente a quanto visto nell'Esempio 3.7, ha come soluzione la funzione $\psi(t) = \tan(t)$, ben definita sull'intervallo $[0, \frac{\pi}{2})$. Per il Corollario 3.6 possiamo allora concludere che $x_* \equiv 0$ è un minimo locale debole per il funzionale J definito nella (3.6) finché $T < \frac{\pi}{2}$. Per $T \geq \frac{\pi}{2}$ non possiamo dire nulla.

CAPITOLO 4

L'equazione di Jacobi

Nel Capitolo 3 abbiamo visto che la condizione di Legendre è una caratterizzazione discretamente efficace della minimalità locale. La (Legendre stretta), tuttavia, presenta l'inconveniente di essere applicabile solo per intervalli $[a, b]$ molto piccoli, a causa dell'esistenza solamente locale della soluzione dell'equazione (D), soluzione indispensabile nella dimostrazione del Teorema 3.4.

In particolare, nell'Esempio 3.7 abbiamo mostrato che gli estremanti possono essere ottimali solo a breve termine, cioè senza essere ottimali globalmente su tutto l'intervallo $[a, b]$ di definizione del problema (P). In altre parole, la (Legendre stretta), in generale, assicura l'esistenza di un $\varepsilon > 0$ tale che il candidato estremante locale x_* sia un minimo se si restringe il problema (P) al subintervallo $[a, a + \varepsilon)$.

La domanda più naturale che possiamo porci, a questo punto, è la seguente: il subintervallo $[a, a+\varepsilon)$ è ottimale in tal senso? ovvero, *esiste un punto $c \in (a, b)$ tale che l'intervallo $[a, c)$ sia il più grande subintervallo di $[a, b]$ con la proprietà che il candidato estremante locale x_* sia un minimo se si restringe il problema* (P) *all'intervallo* $[a, c)$? Una risposta affermativa a questa domanda sarebbe interessante dal punto di vista delle applicazioni: ad esempio, permetterebbe di stabilire con esattezza per quali intervalli la (Wirtinger) sia ottimale. In questo Capitolo vogliamo studiare la teoria, sviluppata da Jacobi mezzo secolo dopo i risultati di Legendre, per risolvere questo tipo di problemi.

1. Equazione di Jacobi. Punti coniugati

Individuato un candidato estremante locale x_* per il problema (P), il nodo centrale della dimostrazione del Teorema 3.4 è rappresentato dall'equazione differenziale (D), che è lineare del primo ordine e coinvolge le funzioni $P(t)$ e $Q(t)$ definite rispettivamente nella (2.1) e nella (2.2). Risulta dunque interessante ottenere qualche informazione in più su questa equazione differenziale. L'idea di Jacobi è stata quella di studiare l'equazione differenziale lineare del secondo ordine definita ponendo

$$-\frac{d}{dt}\Big(P(t)u'(t)\Big) + Q(t)u(t) = 0 \qquad \text{(Jacobi)}$$

al variare di $t \in [a, b]$, dove $u \in C^2([a, b]; \mathbb{R})$ è una funzione incognita. (La forma un po' insolita in cui questa equazione è scritta — non abbiamo sviluppato la derivata nel tempo — è in realtà il modo tradizionale in cui è presentato il problema di Sturm–Liouville: si veda ad esempio [**1**, Sezione 8.6]). Notiamo che la funzione nulla $u \equiv 0$ è una soluzione della (Jacobi) qualsiasi siano le funzioni P e Q assegnate. Per questo, nel seguito ci riferiamo a questa soluzione della (Jacobi) come alla soluzione *banale*; tutte le altre soluzioni della (Jacobi) sono dette *non banali.*

Definizione 4.1 (Punti coniugati). Un punto $\tau \in [a, b]$ si dice *coniugato ad a* (relativamente al candidato estremante locale x_*) se esiste una soluzione non banale u dell'equazione (Jacobi) tale che $u(a) = u(\tau) = 0$.

Facciamo ora due osservazioni molto importanti sulla Definizione 4.1.

Osservazione 4.2. *Se u è una soluzione non banale dell'equazione* (Jacobi) *tale che $u(a) = 0$, allora* $u'(a) \neq 0$. Questa è una conseguenza immediata del Teorema 3.5: infatti, per ogni $\alpha \in \mathbb{R}$, il problema di Cauchy del secondo ordine

$$\begin{cases} -\dfrac{d}{dt}\Big(P(t)u'(t)\Big) + Q(t)u(t) = 0 \\ u(a) = 0 \\ u'(a) = \alpha \end{cases}$$

è equivalente al problema di Cauchy del primo ordine

$$\begin{cases} -\dfrac{d}{dt}\Big(P(t)v(t)\Big) + Q(t)u(t) = 0 \\ v(t) = u'(t) \\ u(a) = 0 \\ v(a) = \alpha \end{cases}$$

che, per il Teorema 3.5, ha un'unica soluzione locale in un intorno (destro) del punto a. Pertanto, u è soluzione della (Jacobi) tale che $u(a) = 0$ e $u'(a) = 0$ se e solo se $u \equiv 0$.

Osservazione 4.3. *Se u_1 e u_2 sono due soluzioni non banali della* (Jacobi) *tali che $u_1(a) = u_2(a) = 0$, allora u_1 e u_2 sono linearmente dipendenti su $\mathbb{R}$.* Infatti, per l'Osservazione 4.2, abbiamo che $u_1'(a) \neq 0$ e $u_2'(a) \neq 0$. Allora esistono $c, d \in \mathbb{R}$ non nulli tali che

$$cu_1'(a) + du_2'(a) = 0$$

(ad esempio $c = -u_2'(a)$ e $d = u_1'(a)$). Siccome la (Jacobi) è lineare, la funzione $u := cu_1 + du_2$ è una soluzione della (Jacobi) tale che $u(a) = u'(a) = 0$. Ma allora, ancora per l'Osservazione 4.2, deve essere $u \equiv 0$, cioè u_1 e u_2 sono linearmente dipendenti.

In particolare, u_1 e u_2 hanno gli stessi zeri nell'intervallo $[a, b]$. Questo significa che la Definizione 4.1 di punto coniugato *non dipende dalla soluzione non banale scelta.*

2. Condizioni necessaria e sufficiente di Jacobi

La nozione di punto coniugato indotta dalla (Jacobi) permette, finalmente, di completare l'estensione al caso infinito dimensionale dei Teoremi 3.1 e 3.2 in modo molto più preciso di quanto visto nel Capitolo 3.

Teorema 4.4 (Jacobi, 1838). *Supponiamo che L sia di classe C^3, e sia $x_* \in C^2([a, b]; \mathbb{R})$ un estremante locale ammissibile per il problema* (P) *che soddisfa la* (Legendre stretta). *Valgono le seguenti affermazioni.*

(i) (Condizione necessaria) *Se x_* è un minimo locale debole per il problema* (P)*, allora non esiste alcun punto coniugato ad a nell'intervallo (a, b).*
(ii) (Condizione sufficiente) *Se non esiste alcun punto coniugato ad a nell'intervallo $(a, b]$, allora x_* è un minimo locale debole per il problema* (P)*.*

Dimostrazione. Mostriamo solo la condizione sufficiente. Sia dunque $x_* \in C^2([a, b]; \mathbb{R})$ un estremante locale ammissibile per il problema (P) che soddisfa la (Legendre stretta) e supponiamo che non esista alcun punto coniugato ad a nell'intervallo $(a, b]$.

Affermiamo allora che la (Jacobi) ammette una soluzione $\bar{u}$ mai nulla nell'intervallo $[a, b]$. Non diamo la dimostrazione di questo fatto in dettaglio (il lettore interessato veda [**2**, pag. 300]). Osserviamo solamente che questa è una conseguenza della prolungabilità e della stabilità rispetto ai dati iniziali delle soluzioni locali date dal Teorema 3.5 ("imbedding theorem") ed è in questo punto che si usano sia la non esistenza di punti coniugati ad a che la (Legendre stretta), ovvero è fondamentale che si abbia $P(t) > 0$ sull'intero intervallo $[a, b]$.

Sia pertanto $\bar{u}$ una soluzione mai nulla nell'intervallo $[a, b]$ per la (Jacobi). Risulta allora ben definita la funzione [1]

$$\psi(t) := -\frac{P(t)\bar{u}'(t)}{\bar{u}(t)}$$

al variare di $t \in [a, b]$. Allora

$$\bar{u}'(t) = -\frac{\psi(t)\bar{u}(t)}{P(t)}$$

[1] Si tratta sostanzialmente di un brillante cambio di variabili nell'equazione (D) che è stato scoperto da Jacobi.

da cui

$$
\begin{aligned}
\psi'(t) &= -\frac{\frac{d}{dt}\Big(P(t)\bar{u}'(t)\Big)\bar{u}(t) - P(t)(\bar{u}'(t))^2}{\bar{u}^2(t)} = \\
&= \frac{-Q(t)\bar{u}^2(t) + \dfrac{\psi^2(t)\bar{u}^2(t)}{P(t)}}{\bar{u}^2(t)} = \\
&= \frac{\psi^2(t)}{P(t)} - Q(t)
\end{aligned}
$$

per ogni $t \in [a, b]$. Quindi, per il Corollario 3.6, x_* è un minimo locale debole per il problema (P). □

Corollario 4.5. *Supponiamo che L sia di classe C^3, e sia $x_* \in C^2([a, b]; \mathbb{R})$ un estremante locale ammissibile per il problema* (P) *che soddisfa la* (Legendre stretta). *Se esiste una soluzione u alla* (Jacobi) *che è mai nulla sull'intervallo $[a, b]$ allora x_* è un minimo locale debole per il problema* (P).

Dimostrazione. Segue immediatamente dall'argomento usato per dimostrare la condizione sufficiente nel Teorema 4.4. □

Concludiamo con una applicazione del Teorema 4.4.

Esempio 4.6 (La disuguaglianza di Wirtinger, continuazione, [**2**, Esercizio 14.17]). Vogliamo applicare il Teorema 4.4 alla disuguaglianza di Wirtinger incontrata nell'Esempio 3.7. Rimaneva da determinare il valore critico $\bar{T} \in [\frac{\pi}{2}, 4]$: vogliamo mostrare che $\bar{T} = \pi$. Più precisamente, vale il seguente risultato.

Proposizione 4.7. *Sia $T \in [0, \pi]$. Allora, per ogni $x \in C^2([0, T]; \mathbb{R})$ tale che $x(0) = x(T) = 0$, si ha*

$$\int_0^T x^2(t)\ dt \le \int_0^T \dot{x}^2(t)\ dt.$$

Questa disuguaglianza è falsa per $T > \pi$.

Dimostrazione. Ricordiamo che $P \equiv 2$ e $Q \equiv -2$, quindi la (Jacobi) in tal caso diventa $u'' + u = 0$. Se imponiamo la condizione $u(0) = 0$, questa equazione ha come soluzione la funzione $u(t) = \sin(t)$. Pertanto $\tau = \pi$ è il più piccolo punto (positivo) coniugato all'estremo $a = 0$. Per il Teorema 4.4, allora, $x_* \equiv 0$ è un minimo globale debole per il problema (PW) se e solo se $T \le \pi$. In particolare, quando $T > \pi$, $x_* \equiv 0$ non è un minimo per il problema (PW), cioè la disuguaglianza è falsa in questo caso. □

La Proposizione 4.7 può essere migliorata nel senso seguente.

Proposizione 4.8 (Disuguaglianza di Wirtinger). *Sia $f \in \mathrm{Lip}([a, b]; \mathbb{R})$ tale che $f(a) = f(b) = 0$. Allora*

$$\int_a^b f(t)^2\ dt \le \frac{(b-a)^2}{\pi^2} \int_a^b f'(t)^2\ dt,$$

con l'uguaglianza se e solo se $f(t) = c \sin\left(\dfrac{t-a}{b-a}\,\pi\right)$ per qualche $c \in \mathbb{R}$.

Lasciamo al lettore la dimostrazione di questo risultato (si veda [**2**, Esercizio 21.14]).

APPENDICE A

Sugli archi Lipschitziani

1. Un teorema di approssimazione per archi Lipschitziani

Sia $h\colon [a,b] \to \mathbb{R}$ una funzione Lipschitziana. Nel seguito proviamo il seguente risultato.

TEOREMA A.1. *Per ogni $\varepsilon > 0$, esiste un polinomio $g(t)$ con $g(a) = h(a)$ e $g(b) = h(b)$ tale che*

$$\|g'\| \leq \|h'\| + \varepsilon, \quad \|g - h\| + \|g' - h'\|_{L^1} < \varepsilon.$$

Per dimostrare questo risultato abbiamo bisogno di richiamare i seguenti due teoremi classici.

TEOREMA A.2 (Lusin). *Sia Ω un insieme aperto limitato in $\mathbb{R}^n$ e sia $\phi\colon \Omega \to \mathbb{R}$ una funzione misurabile tale che $|\phi(x)| \leq M$ q.o. Per ogni $\varepsilon > 0$ esiste $g\colon \Omega \to \mathbb{R}$, continua con supporto compatto, tale che*

$$\sup_{\Omega}|g| \leq M, \quad \operatorname{meas}\{\, x \in \Omega : \phi(x) \neq g(x) \,\} < \varepsilon.$$

TEOREMA A.3 (Approssimazione polinomiale di Weierstrass). *Sia $f\colon [a,b] \to \mathbb{R}$ una funzione continua. Per ogni $\varepsilon > 0$, esiste un polinomio p tale che*

$$\|f - p\| < \varepsilon.$$

Per le dimostrazioni di questi risultati si veda [**6**]. Diamo ora una dimostrazione del Teorema A.1.

DIMOSTRAZIONE. Seguiamo i suggerimenti dati in [**2**, Esercizio 21.13]. Fissiamo $\varepsilon > 0$. Siccome h è una funzione Lipschitziana, la derivata h' esiste q.o. e definisce una funzione misurabile limitata $h'\colon [a,b] \to \mathbb{R}$. Pertanto, per il Teorema di Lusin A.2, esiste una funzione continua $f\colon [a,b] \to \mathbb{R}$ tale che

$$\sup_{[a,b]}|f| \leq \|h'\|, \quad \operatorname{meas}\{\, x \in \Omega : h'(x) \neq f(x) \,\} < \varepsilon. \tag{1.1}$$

Siccome f è continua sull'intervallo compatto $[a,b]$, per il Teorema di Weierstrass A.3 esiste un polinomio $p(t)$ tale che

$$\|p - f\| < \varepsilon. \tag{1.2}$$

Ora poniamo

$$\phi(t) = h(a) + \int_a^t p(s)ds$$

e

$$g(t) = \phi(t) + c(t-a)$$

per ogni $t \in [a,b]$, dove $c \in \mathbb{R}$ è una costante da fissare. Chiaramente $g(a) = h(a)$ per ogni $c \in \mathbb{R}$, mentre

$$g(b) = \phi(b) + c(b-a) = h(a) + \int_a^b p(s)ds + c(b-a)$$

è uguale a $h(b)$ se e solo se

$$c = \frac{1}{b-a}\left(h(b) - h(a) - \int_a^b p(s)ds\right).$$

Ricordiamo che, siccome h è Lipschitziana, h è assolutamente continua e perciò

$$h(t) = h(a) + \int_a^t h'(s)ds \quad \forall t \in [a,b].$$

Possiamo allora stimare la costante c come

$$\begin{aligned} |c| &= \frac{1}{b-a}\left|h(b)-h(a)-\int_a^b p(s)ds\right| = \\ &= \frac{1}{b-a}\left|\int_a^b h'(s)-p(s)ds\right| \le \\ &\le \frac{1}{b-a}\int_a^b |h'(s)-f(s)|+|f(s)-p(s)|ds \le \\ &\le \frac{b-a+2\|h'\|}{b-a}\varepsilon \end{aligned} \tag{1.3}$$

per la (1.2), dato che, per la (1.1),

$$\begin{aligned} \int_a^b |h'(s)-f(s)|ds &= \int_{\{\,h'=f\,\}} |h'(s)-f(s)|ds + \int_{\{\,h'\neq f\,\}} |h'(s)-f(s)|ds = \\ &= \int_{\{\,h'\neq f\,\}} |h'(s)-f(s)|ds \le \\ &\le 2\|h'\| \int_{\{\,h'\neq f\,\}} ds < 2\|h'\|\varepsilon. \end{aligned}$$

Inoltre

$$g'(t) = \phi'(t) + c = p(t) + c \quad \forall t \in (a,b),$$

quindi

$$\|g'\| \le \|p\| + |c| \le \|f\| + \varepsilon + |c| \le \|h'\| + \varepsilon + |c|.$$

Infine, per ogni $t \in [a,b]$,

$$\begin{aligned} |g(t)-h(t)| &= \left|h(a)+\int_a^t p(s)ds + c(t-a) - h(t)\right| \le \\ &\le \left|\int_a^t h'(s)-p(s)ds\right| + |c|(t-a) \le \\ &\le (b-a+2\|h'\|)\,\varepsilon + |c|(b-a) \end{aligned}$$

e, stimando gli integrali come sopra,

$$\begin{aligned} \int_a^b |g'(t)-h'(t)|dt &\le \int_a^b |p(t)-h'(t)|dt + |c|(b-a) \le \\ &\le \int_a^b |p(t)-f(t)|+|f(t)-h'(t)|dt + |c|(b-a) \le \\ &\le (b-a)\varepsilon + 2\|h'\|\varepsilon + |c|(b-a) = (b-a+2\|h'\|)\,\varepsilon + |c|(b-a). \end{aligned}$$

Pertanto

$$\max\{\,\|g-h\|, \|g'-h'\|_{L^1}\,\} \le (b-a+2\|h'\|)\,\varepsilon + |c|(b-a).$$

In conclusione, la funzione g soddisfa

$$g(a)=h(a), \quad g(b)=h(b)$$

e

$$\|g'\| \le \|h'\| + \varepsilon + |c| \le \|h'\| + \frac{b-a+2\|h'\|}{b-a}\varepsilon$$

e

$$\begin{aligned} \|g-h\| + \|g'-h'\|_{L^1} &\le 2\max\{\,\|g-h\|, \|g'-h'\|_{L^1}\,\} \le \\ &\le 2\,(b-a+2\|h'\|)\,\varepsilon + 2|c|(b-a) \le \\ &\le 4\,(b-a+2\|h'\|)\,\varepsilon \end{aligned}$$

entrambi per la (1.3). Questo conclude la dimostrazione. □

2. Un teorema per funzioni estremanti Lipschitziane

Possiamo ora applicare il Teorema A.1 allo studio del funzionale

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

dove L è una funzione localmente Lipschitziana.

COROLLARIO A.4 ("fairing theorem"). *Per ogni $x \in \mathrm{Lip}([a,b];\mathbb{R})$ e per ogni $\varepsilon > 0$, esiste un polinomio y, avente gli stessi valori di x in a e b, tale che*

$$\|x - y\| < \varepsilon, \quad |J(x) - J(y)| < \varepsilon.$$

DIMOSTRAZIONE. Per il Teorema A.1, esiste un polinomio y che ha gli stessi valori di x in a e b e che soddisfa

$$\|\dot{y}\| \leq \|\dot{x}\| + \varepsilon, \quad \|y - x\| + \|\dot{y} - \dot{x}\|_{L^1} < \varepsilon.$$

Definiamo il compatto

$$K = [a,b] \times \bar{B}(\|x\| + \varepsilon) \times \bar{B}(\|\dot{x}\| + \varepsilon)$$

dove le palle sono chiuse e centrate nell'origine di $\mathbb{R}^n$. Sia $M > 0$ la costante di Lipschitz della funzione L su K. Allora possiamo stimare

$$\begin{aligned}|J(x) - J(y)| \leq \int_a^b &|L(t, x(t), \dot{x}(t)) - L(t, y(t), \dot{y}(t))|dt \leq \\ &\leq M \int_a^b |x(t) - y(t)| + |\dot{x}(t) - \dot{y}(t)|dt \\ &\leq M\left((b-a)\|x - y\| + \|\dot{x} - \dot{y}\|_{L^1}\right) < \\ &< M(b - a + 1)\varepsilon\end{aligned}$$

e la dimostrazione è completa. □

Facciamo ora la seguente osservazione.

OSSERVAZIONE A.5. Sia $X \subset \mathrm{Lip}([a,b];\mathbb{R})$ e siano $r, R \in (0, \infty]$, e consideriamo il problema di minimizzazione

$$\min J(x) \quad : \quad x \in X, \quad x(a) = A, \quad x(b) = B, \quad \|x\| < r, \quad \|\dot{x}\| < R. \tag{2.1}$$

Allora l'estremo inferiore in questo problema è lo stesso per ogni insieme X che contiene tutti i polinomi, i.e. per ogni $X_1, X_2 \subset \mathrm{Lip}([a,b];\mathbb{R})$ abbiamo

$$X_1 \cap X_2 \supset P([a,b];\mathbb{R}) \implies \inf_{X_1} J(x) = \inf_{X_2} J(x).$$

Ragioniamo per assurdo. Assumiamo, per contraddizione, che $\inf_{X_1} J(x) < \inf_{X_2} J(x)$ e fissiamo $\varepsilon > 0$. Allora, per definzione di inf, possiamo trovare $x_\varepsilon \in X_1$ tale che

$$J(x_\varepsilon) + \varepsilon \leq J(x) \quad \forall x \in X_2.$$

Siccome nel problema (2.1) ogni arco Lipschitziano ammissibile è limitato e ha derivata limitata, possiamo applicare il Corollario A.4 e trovare un polinomio y, ammissibile per il problema (2.1) in virtù del Teorema A.1 (eventualmente scegliendo un ε più piccolo), tale che

$$\|x_\varepsilon - y\| < \varepsilon, \quad |J(x_\varepsilon) - J(y)| < \varepsilon.$$

Ma allora

$$J(y) = J(y) - J(x_\varepsilon) + J(x_\varepsilon) \leq |J(y) - J(x_\varepsilon)| + J(x_\varepsilon) < \varepsilon + J(x) - \varepsilon = J(x) \quad \forall x \in X_2$$

e quindi, in particolare, $J(y) < J(y)$, perché X_2 contiene tutti i polinomi. Ma questo è assurdo.

Con questa semplice osservazione, possiamo concludere dimostrando il seguente risultato.

TEOREMA A.6. *Sia L una funzione localmente Lipschitziana e assumiamo che $x_* \in C^2([a,b];\mathbb{R})$ sia un minimo locale debole per il problema*

$$\text{(P)} \qquad \min J(x) \quad : \quad x \in C^2([a,b];\mathbb{R}), \quad x(a) = A, \quad x(b) = B.$$

Allora x_* è anche un minimo locale debole relativamente allo spazio $\mathrm{Lip}([a,b];\mathbb{R})$. *Inoltre, lo stesso enunciato vale per minimi locali forti.*

DIMOSTRAZIONE. Se $x_* \in C^2([a,b];\mathbb{R})$ è un minimo locale debole per (P), allora esiste $\varepsilon > 0$ tale che $J(x_*) \leq J(x)$ per ogni funzione ammissibile $x \in C^2([a,b];\mathbb{R})$ tale che $\|x - x_*\| < \varepsilon$ e $\|\dot{x} - \dot{x}_*\| < \varepsilon$. Definiamo $r = \|x_*\| + \varepsilon$ e $R = \|\dot{x}_*\| + \varepsilon$. Allora il problema (P) è esattamente il problema (2.1) che abbiamo considerato nell'Osservazione A.5 sopra, e possiamo concludere che x_* è anche un minimo locale debole relativamente allo spazio $\mathrm{Lip}([a,b];\mathbb{R})$ semplicemente per le inclusioni $P([a,b];\mathbb{R}) \subset C^2([a,b];\mathbb{R}) \subset \mathrm{Lip}([a,b];\mathbb{R})$.

Se invece $x_* \in C^2([a,b];\mathbb{R})$ è un minimo locale forte per (P), allora esiste $\varepsilon > 0$ tale che $J(x_*) \leq J(x)$ per ogni funzione ammissibile $x \in C^2([a,b];\mathbb{R})$ tale che $\|x - x_*\| < \varepsilon$. Questa volta definiamo $r = \|x_*\| + \varepsilon$ e $R = +\infty$. Allora il problema (P) è esattamente il problema (2.1) che abbiamo considerato nell'Osservazione A.5 sopra, e concludiamo che x_* è anche un minimo locale forte relativamente allo spazio $\mathrm{Lip}([a,b];\mathbb{R})$ come nel caso precedente. $\square$

Riferimenti

[1] H. Brézis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Universitext Series, Springer, 2011.
[2] F. Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*, GTM Series, Vol. 264, Springer, 2013.
[3] R. Courant and H. Robbins, *Che cos'è la matematica?*, 2nd ed., Bollati Boringhieri, 2000.
[4] G. De Marco, *Analisi Due. Teoria ed esercizi*, 2nd ed., Decibel–Zanichelli, 1999.
[5] R. Monti, *Analisi 2. Appunti del corso*, 2012.
[6] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 1966.