ELOGIO DELLA PIGRIZIA:
IL PRINCIPIO DI MINIMA AZIONE

CARLANGELO LIVERANI

1. Lagrangiana

Dato un sistema di N particelle soggette a forze conservative, possiamo scrivere le equazioni di Newton come

$$M\ddot{x} = -\nabla U \tag{1.1}$$

dove $x : \mathbb{R} \to \mathbb{R}^d$, M è una matrice $d \times d$ definita positiva e $U \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ una funzione detta *potenziale*. Accade che tali equazioni possono essere alternativamente descritte attraverso una funzione $\mathcal{L} \in \mathcal{C}^2(\mathbb{R}^{2d}, \mathbb{R})$ definita da

$$\mathcal{L}(x, y) = \frac{1}{2}\langle y, My\rangle - U(x). \tag{1.2}$$

Tale funzione $\mathcal{L}$ è detta *Lagrangiana* e determina delle equazioni differenziali, per la funzione incognita $x \in \mathcal{C}^2(\mathbb{R}, \mathbb{R}^d)$, definite da

$$\frac{d}{dt}\partial_y \mathcal{L}(x(t), \dot{x}(t)) - \partial_x \mathcal{L}(x(t), \dot{x}(t)) = 0. \tag{1.3}$$

Tali equazioni sondo dette *equazioni di Lagrange.*

Esercizio 1.1. *Si mostri che le equazioni di Lagrange associate alla Lagrangiana* (1.2) *sono esattamente le equazioni di Newton* (1.1).

Lo scrivere le equazioni del moto in forma lagrangiana ha svariati vantaggi, il primo è una maggiore facilità nel cambio di coordinate. Infatti, sia $\Xi \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$ un diffeomorfismo allora si ha il seguente fondamentale fatto.

Teorema 1.2. *Si consideri il cambio di coordinate $x = \Xi(q)$ allora the equazioni di Newton* (1.1) *nelle coordinate q coincidono con le equazioni di Lagrange per la Lagrangiana $\widetilde{\mathcal{L}}(\xi, \eta) = \mathcal{L}(\Xi(\xi), D\Xi(\xi)\eta)$.*[1]

Proof. Prima di tutto occorre calcolare come sono fatte le equazioni (1.1) nelle nuove coordinate. Si noti che

$$\begin{aligned}
&\dot{x} = D\Xi(q)\dot{q}\\
&\ddot{x}_i = \sum_{j,k} \partial_{q_j}\partial_{q_k}\Xi_i(q)\dot{q}_j\dot{q}_k + \sum_j \partial_{q_j}\Xi_i(q)\ddot{q}_j\\
&\partial_{q_i}(U \circ \Xi)(q) = \sum_j (\partial_{x_j}U) \circ \Xi(q) \cdot \partial_{q_i}\Xi_j(q).
\end{aligned} \tag{1.4}$$

Date: Versione di May 7, 2021.

[1]Normalmente si abusano le notazioni e la nuova funzione $\widetilde{\mathcal{L}}$ viene chiamata nuovamente $\mathcal{L}$, sebbene sia ovviamente una funzione differente. Inoltre si tende a scrivere $\partial_q\mathcal{L}(q, \dot{q})$ e $\partial_{\dot{q}}\mathcal{L}(q, \dot{q})$ (invece di $\partial_\xi\widetilde{\mathcal{L}}(\xi, \eta)$ e $\partial_\eta\widetilde{\mathcal{L}}(\xi, \eta)$) sebbene tali espressioni, per quanto intuitive, non abbiano molto senso. Nel seguito mi conformerò a questa discutibile, ma conveniente, tradizione.

Poichè Ξ è un diffeomorfismo $D\Xi$ deve essere invertible, quindi le equazioni (1.1) sono equivalenti a

$$D\Xi^T M\ddot{x} = -D\Xi^T \nabla U(x) = -\left[\nabla(U \circ \Xi)\right] \circ \Xi^{-1}(x).$$

Ovvero, ponendo $\widetilde{U} = U \circ \Xi$ e usando le (1.4) si ha

$$\sum_{l,j,k,r} \partial_{q_i}\Xi_l \cdot M_{l,r} \cdot \partial_{q_j}\partial_{q_k}\Xi_r \cdot \dot{q}_j\dot{q}_k + \sum_{l,j,r} \partial_{q_i}\Xi_l \cdot M_{l,r} \cdot \partial_{q_j}\Xi_r \ddot{q}_j = -\partial_{q_i}\widetilde{U}. \tag{1.5}$$

D'altro canto

$$\begin{aligned}
&\mathcal{L}(q,\dot{q}) = \frac{1}{2}\langle \dot{q}, D\Xi^T(q) M D\Xi(q)\dot{q}\rangle - \tilde{U}(q)\\
&\partial_{q_i}\mathcal{L} = \sum_{j,k,l} \partial_{q_r}\Xi_j M_{j,k} \partial_{q_i}\partial_{q_l}\Xi_k \dot{q}_l\dot{q}_r - \partial_{q_i}\widetilde{U}\\
&\partial_{\dot{q}_i}\mathcal{L} = \sum_{l,r,j} \partial_{q_i}\Xi_l \cdot M_{l,r} \cdot \partial_{q_j}\Xi_r \dot{q}_j\\
&\frac{d}{dt}\partial_{\dot{q}_i}\mathcal{L} = \sum_{l,r,j,k} \partial_{q_k}\partial_{q_i}\Xi_l \cdot M_{l,r} \cdot \partial_{q_k}\partial_{q_j}\Xi_r \dot{q}_j\dot{q}_k + \sum_{l,r,j} \partial_{q_i}\Xi_l \cdot M_{l,r} \cdot \partial_{q_j}\Xi_r \ddot{q}_j
\end{aligned}$$

e inserendo queste espressione nelle equazioni di Lagrange (1.3) si ottengono esattamente le (1.5), come annunciato. □

Il vantaggio di cambiare facilmente coordinate è particolarmente evidente quanto si considerano sistemi vincolati. Si consideri il caso

$$\begin{aligned}
&M\ddot{x}(t) = -\nabla U(x) + R(t)\\
&\Phi(x) = 0\\
&\langle R(t), v\rangle = 0 \text{ per tutti i } v \in \mathbb{R}^d \text{ tali che } D\Phi(x(t))v = 0.
\end{aligned} \tag{1.6}$$

Dove R sono le reazioni vincolari incognite. $\Phi \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^{d-m})$, $m < d$, esprime il vincolo e si assume che $D\Phi$ abbia rango massimo ad ogni punto. Infine, l'ultima riga esprime il principio dei lavori virtuali ovvero (parlando un poco impropriamente) che le forze vincolari al tempo t non compiono lavoro su qualunque moto compatibile col vincolo nel punto $x(t)$.

Poichè il rango di $D\Phi$ è massimo, il teorema della funzione implicita ci dice che localmente (ovvero nell'intorno di qualunque punto assegnato) esiste una funzione $\Xi \in \mathcal{C}^2(V, \mathbb{R}^d)$, dove V è un aperto di $\mathbb{R}^m$, tale che per ogni $q \in V$ si ha $\Phi(\Xi(q)) = 0$. Dunque se consideriamo solo moti del tipo $x(t) = \Xi(q(t))$, questi soddisfano automaticamente il vincolo. Inoltre, differenziando, si ha $D\Phi D\Xi = 0$. Infine, poichè il kernel di Φ deve essere m dimensionale, tutti i vettori tali $D\phi\, v = 0$ si possono esprime come $v = D\Xi w$ per qualche $w \in \mathbb{R}^m$. Questo significa che, per ogni $w \in \mathbb{R}^m$ deve essere

$$\langle R(t), D\Xi(q(t))w\rangle = 0$$

ovvero il principio dei lavori virtuali è equivalente a

$$D\Xi(q(t))^T R(t) = 0. \tag{1.7}$$

Possiamo quindi moltiplicare la prima delle (1.6) per $D\Xi^T$ e ottenere

$$D\Xi^T M\ddot{x} = -D\Xi^T \nabla U(x). \tag{1.8}$$

Queste sono m equazioni differenziali per le m funzioni incognite $q(t)$. Ci siamo dunque ridotti ad un problema in cui da un lato sono state eliminate le reazioni vincolari e che ha il giusto numero di equazioni per determinare le rimanenti incognite.

Esercizio 1.3. *Si usino le (1.4) per scrivere esplicitamente le (1.8) in termini delle funzioni q e si verifichi che coincidono con le equazioni di Lagrange per la Lagrangiana* $\mathcal{L}(q,\dot{q}) = \frac{1}{2}\langle \dot{q}, D\Xi^T(q)MD\Xi(q)\dot{q}\rangle - U(\Xi(q))$.

Remark 1.4. *Abbiamo dunque visto, che dato un sistema vincolato con vincoli ideali (ovvero che soddisfano il principio dei lavori virtuali), le equazioni del moto possono essere facilmente derivate prima cambiando variabile nella Lagrangiana e poi usando le equazioni di Lagrange per la Lagrangiana così ottenuta.*

2. Azione

Un altro importante uso della Lagrangiana è quello di definire un'altra sorprendente funzione: l'*azione*. Questa è una funzione definita su un insieme un poco strano: un insieme di funzioni! Vediamo di essere più precisi: per ogni $t_0, t_1 \in \mathbb{R}$ e $A, B \in \mathbb{R}^d$ si consideri l'insieme (detto *insieme dei moti*) $\mathcal{M}_{t_0,t_1}(A,B) = \{q \in \mathcal{C}^1([t_0,t_1],\mathbb{R}^d) \;:\; q(t_0) = A,\, q(t_1) = B\}$.[2] Possiamo dunque definire la funzione $S : \mathcal{M}_{t_0,t_1}(A,B) \to \mathbb{R}$ come

$$S(q) = \int_{t_0}^{t_1} \mathcal{L}(q(s),\dot{q}(s))ds. \tag{2.1}$$

Siccome una funzione di funzioni può sembrare alquanto inusuale, viene spontaneo chiedersi quali proprietà abbia, per esempio: è continua? è differenziabile? è convessa ? etc.

Un attimo di riflessione mostra che per rispondere a tali domande occorre prima di tutto chiarire che significano. Per esempio, una funzione $f : \mathbb{R}^n \to \mathbb{R}$ è continua se solo se per ogni successione convergente x_n si ha $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x)$. Ci si può chiedere se la stessa cosa sia vera per S, ma per rispondere occorre prima spiegare che significa *convergere* per una successione di elementi di $\mathcal{M}_{t_0,t_1}(A,B)$. A questa domanda si possono dare molte risposte e approfondire questo tema ci porterebbe allo studio delle possibili topologie per uno spazio infinito dimensionale. Questo è al di fuori dei nostri scopi attuali, ci accontentiamo quindi della seguente definizione che, sperabilmente, il lettore troverà naturale: diremo che una successione $\{q_n\} \subset \mathcal{M}_{t_0,t_1}(A,B)$ converge a $\bar{q} \in \mathcal{M}_{t_0,t_1}(A,B)$ se

$$\lim_{n\to\infty} \sup_{s\in[t_0,t_1]} \|q_n(s) - \bar{q}(s)\| + \sup_{s\in[t_0,t_1]} \|\dot{q}_n(s) - \dot{\bar{q}}(s)\| = 0.$$

Esercizio 2.1. *Si mostri che con la nozione di convergenza appena introdotta* $\mathcal{M}_{t_0,t_1}(A,B)$ *è uno spazio completo, ovvero: ogni successione di Cauchy è convergente.*

In analogia col caso finito dimensionale diremo dunque che S è continua in $\bar{q}$ se, per ogni $\{q_n\}$ che converge a $\bar{q}$ si ha $\lim_{n\to\infty} S(q_n) = S(\bar{q})$.

Esercizio 2.2. *Si mostri che se $\mathcal{L} \in \mathcal{C}^0$ allora S è continua.*

---

[2] Ovvero tutti i moti che, patendo al tempo t_0, nel tempo $t_1 - t_0$ vanno dal punto A al punto B.

Che dire della differenziabiltà? Una funzione da $f : \mathbb{R}^n \to \mathbb{R}$ è differenziabile in x se esiste una funzione lineare $L : \mathbb{R}^n \to \mathbb{R}^n$ tale che

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - L(h)\|}{\|h\|} = 0.$$

Tale funzione L si chiama derivata e si scrive $D_x f \cdot h = \langle \nabla_x f, h\rangle = L(h)$.[3]

Dunque per parlare di differenziabilità occorre introdurre una funzione lineare. Sfortunatamente $\mathcal{M}_{t_0,t_1}(A,B)$ non è uno spazio lineare e quindi non ha senso parlare di funzioni lineari su $\mathcal{M}_{t_0,t_1}(A,B)$.

Esercizio 2.3. *Si mostri che se $q, \tilde{q} \in \mathcal{M}_{t_0,t_1}(A,B)$ allora $q - \tilde{q} \in \mathcal{M}_{t_0,t_1}(0,0)$. Inoltre si mosti che $\mathcal{M}_{t_0,t_1}(0,0)$ è uno spazio vettoriale sul campo dei reali.*

In analogia con il caso finito dimensionale possiamo dunque studiare $S(q+h) - S(q)$ dove $q \in \mathcal{M}_{t_0,t_1}(A,B)$ e $h \in \mathcal{M}_{t_0,t_1}(0,0)$. Inoltre su $\mathcal{M}_{t_0,t_1}(0,0)$ possiamo introdurre la norma

$$\|h\|_{\mathcal{C}^1} = \sup_{s\in[t_0,t_1]} \|h(s)\| + \sup_{s\in[t_0,t_1]} \|\dot{h}(s)\|.$$

Esercizio 2.4. *Si mostri che $\|\cdot\|_{\mathcal{C}^1}$ è una norma su $\mathcal{M}_{t_0,t_1}(0,0)$.*[4]

Esercizio 2.5. *Si mostri che $\mathcal{M}_{t_0,t_1}(0,0)$ è uno spazio completo rispetto alla norma $\|\cdot\|_{\mathcal{C}^1}$, ovvero che ogni successione di Cauchy è convergente.*[5]

Diremo dunque che S è differenziabile in $q \in \mathcal{M}_{t_0,t_1}(A,B)$ se esiste una funzione lineare $L_q : \mathcal{M}_{t_0,t_1}(0,0) \to \mathbb{R}$ tale che, per ogni successione $\{h_n\} \subset \mathcal{M}_{t_0,t_1}(0,0)$ tale che $\lim_{n\to\infty} \|h_n\|_{\mathcal{C}^1} = 0$ si ha

$$\lim_{n\to\infty} \frac{|S(q+h_n) - S(q) - L_q(h_n)|}{\|h_n\|_{\mathcal{C}^1}} = 0.$$

Oppure, più brevemente,

$$\lim_{h\to 0} \frac{|S(q+h) - S(q) - L_q(h)|}{\|h\|_{\mathcal{C}^1}} = 0.$$

Naturalmente diremo che L_q è il differenziale di S in q (per essere precisi si chiama *differenziale di Fréchet*).

Lemma 2.6. *Se $\mathcal{L} \in \mathcal{C}^1$ allora la funzione S è differenziabile in ogni $q \in \mathcal{M}_{t_0,t_1}(A,B)$ e si ha che*

$$L_q(h) = \int_{t_0}^{t_1} \sum_i \left\{ \partial_{\dot{q}_i} \mathcal{L}(q(s), \dot{q}(s)) \dot{h}_i + \partial_{q_i} \mathcal{L}(q(s), \dot{q}(s)) h_i(s) \right\} ds.$$

Proof. Usando la formula di Taylor al primo ordine si ha

$$\begin{aligned}\mathcal{L}(q(s) + h(s), \dot{q}(s) + \dot{h}(s)) =& \mathcal{L}(q(s), \dot{q}(s)) + \sum_i \partial_{q_i} \mathcal{L}(q(s), \dot{q}(s)) h_i(s) \\ &+ \sum_i \partial_{\dot{q}_i} \mathcal{L}(q(s), \dot{q}(s)) \dot{h}_i(s) + o(\|h\|_{\mathcal{C}^1}).\end{aligned}$$

---

[3] Infatti una funzione lineare da $\mathbb{R}^n$ a $\mathbb{R}$ è identificata univocamente da un vettore, tale vettore è appunto $\nabla_x f$ i cui elementi sono le derivate parziali della f, ovvero $\partial_{x_i} f$.

[4] Si ricordi che, dato uno spazio vettoriale $\mathbb{V}$ sul campo dei reali, una funzione $\|\cdot\| : \mathbb{V} \to \mathbb{R}_+$ è una norma se: 1) $\|x\| > 0$ per ogni $x \in \mathbb{V}$, $x \neq 0$; 2) $\|\lambda x\| = |\lambda| \|x\|$ per ogni $x \in \mathbb{V}$ e $\lambda \in \mathbb{R}$; 3) $\|x+y\| \leq \|x\| + \|y\|$ per ogni $x, y \in \mathbb{V}$.

[5] Gli spazi vettoriali normati e completi si chiamano *spazi di Banach*.

Dunque

$$
\begin{aligned}
S(q+h)-S(q) = &\int_{t_0}^{t_1} \sum_i \partial_{q_i}\mathcal{L}(q(s),\dot q(s))h_i(s)ds \\
&+ \int_{t_0}^{t_1} \sum_i \partial_{\dot q_i}\mathcal{L}(q(s),\dot q(s))\dot h_i(s)ds + o(\|h\|_{\mathcal{C}^1}).
\end{aligned}
$$

Il Lemma segue immediatamente dall'ultima equazione. □

Visto che abbiamo il differenziale, è naturale studiare i *punti stazionari*, ovvero i moti q per cui $L_q = 0$. Risulta che tali moti sono molto speciali, cominciamone lo studio.

Come al solito, se assumiamo un poco di più otteniamo un risultato più preciso.

Lemma 2.7. *Se $\mathcal{L} \in \mathcal{C}^2$ e $q \in \mathcal{C}^2$ allora si ha che $L_q(h) = 0$ per ogni $h \in \mathcal{M}_{t_0,t_1}(0,0)$ se e solo se q soddisfa le equazioni di Lagrange con condizioni $q(t_0) = A$, $q(t_1) = B$.*

Proof. Per la necessità basta notare che

$$
\begin{aligned}
L_q(h) = &\int_{t_0}^{t_1} \sum_i \partial_{q_i}\mathcal{L}(q(s),\dot q(s))h_i(s)ds \\
&- \int_{t_0}^{t_1} \sum_i \frac{d}{dt}\partial_{\dot q_i}\mathcal{L}(q(s),\dot q(s)) \cdot h_i(s)ds
\end{aligned}
$$

dove abbiamo integrato per parti e ricordato che $h(t_0) = h(t_1) = 0$ e che, per ipotesi, $\partial_{q_i}\mathcal{L}(q(\cdot),\dot q(\cdot)) \in \mathcal{C}^1$. Se infatti le equazioni di Langrange non fossero soddisfatte, allora esisterebbe un i e un intervallo di tempi in cui $\partial_{q_i}\mathcal{L}(q(s),\dot q(s)) - \frac{d}{dt}\partial_{q_i}\mathcal{L}(q(s),\dot q(s))$ non si annulla mai. Moltiplicando per una h supportata in tale intervallo e con $h_i \geq 0$ si avrebbe quindi un integrale non nullo contrariamente alle ipotesi.

La sufficienza segue dal fatto che se q soddisfa le equazioni di Lagrange allora $L_q(h) = 0$. □

È interessante notare che l'ipotesi $q \in \mathcal{C}^2$ in realtà è molto meno forte di quanto possa sembrare a prima vista.

Lemma 2.8. *Se $\mathcal{L} \in \mathcal{C}^2$ e $q \in \mathcal{M}_{t_0,t_1}(A,B)$ è stazionaria e la matrice $\partial_{\dot q_i}\partial_{\dot q_j}\mathcal{L}(q,\dot q)$ è invertibile, allora $q \in \mathcal{C}^2((t_0,t_1),\mathbb{R}^d)$.*

Proof. Il fatto che q sia stazionario significa che per ogni $h \in \mathcal{M}_{t_0,t_1}(0,0)$ si ha $L_q(h) = 0$. Quindi

$$
\begin{aligned}
L_q(h) &= \int_{t_0}^{t_1} \sum_i \partial_{q_i}\mathcal{L}(q(s),\dot q(s)) \int_{t_0}^{s} \dot h_i(s_1)ds_1 ds \\
&\quad + \int_{t_0}^{t_1} \sum_i \partial_{\dot q_i}\mathcal{L}(q(s),\dot q(s)) \cdot \dot h_i(s)ds \\
&= \int_{t_0}^{t_1} \sum_i \left\{ \int_{s}^{t_1} \partial_{q_i}\mathcal{L}(q(s_1),\dot q(s_1))ds_1 + \partial_{\dot q_i}\mathcal{L}(q(s),\dot q(s)) \right\} \dot h_i(s)ds.
\end{aligned}
$$

Ora si noti che se $g \in \mathcal{C}^0([t_0,t_1],\mathbb{R}^d)$, allora $h(t) = \int_{t_0}^t g(s)ds - \frac{t-t_0}{t_1-t_0}\int_{t_0}^{t_1} g(s)ds \in \mathcal{M}_{t_0,t_1}(0,0)$. Poichè $\dot h = g - \frac{1}{t_1-t_0}\int_{t_0}^{t_1} g(s)ds =: g - \alpha_g$, segue che se $G \in \mathcal{C}^1$ e $\int_{t_0}^{t_1}\langle G, \dot h\rangle = 0$ per ogni $h \in \mathcal{M}_{t_0,t_1}(0,0)$, allora

$$0 = \int_{t_0}^{t_1} \langle G(s), g(s) - \alpha_g\rangle ds = \int_{t_0}^{t_1} \langle \left[G(s) - \frac{1}{t_1 - t_0}\int_{t_0}^{t_1} G(s_1)ds_1\right], g(s)\rangle ds,$$

per ogni $g \in \mathcal{C}^0([t_0,t_1],\mathbb{R}^d)$. Questo implica che, per ogni $s \in [t_0,t_1]$,

$$\int_s^{t_1} \partial_{q_i}\mathcal{L}(q(s_1),\dot q(s_1))ds_1 + \partial_{\dot q_i}\mathcal{L}(q(s),\dot q(s)) - c_i = 0$$

per opportune costanti c_i. A questo punto si noti che la condizione di invertibilità della matrice $\partial_{\dot q_i}\partial_{\dot q_j}\mathcal{L}(q,\dot q)$ è esattamente quello che occorre per applicare il teorema della funzione implicita ed esprimere $\dot q$ come una funzione $\mathcal{C}^1$. Infatti $H(s) = \int_s^{t_1}\partial_q\mathcal{L}(q(s_1),\dot q(s_1))ds_1$ e $G(s,x) = \partial_q\mathcal{L}(q(s),x)$ sono $\mathcal{C}^1$ per ipotesi. Dunque, ponendo $F(s,x) = H(s) + G(s,x) - c$, il teorema della funzione implicita afferma che $F(s,x) = 0$ ha una unica soluzione $\dot q(s)$ e $\dot q \in \mathcal{C}^1$. □

Abbiamo dunque il fatto notevole che S ha derivata nulla sulle soluzioni delle equazioni di Lagrange. Certo rimane la cosa un poco strana che invece di avere un problema di Cauchy (ovvero specificare $q(t_0), \dot q(t_0)$) si specificano i valori delle sole posizioni agli estremi dell'intervallo temporale considerato. Viene naturale domandarsi che relazione ci sia tra questo problema "ai bordi" e il problema di Cauchy. Prima di affrontare questo problema mostriamo un'altra proprietà sorprendente dell'azione. In questo caso assumiamo che $\mathcal{L} \in \mathcal{C}^3(\mathbb{R}^{2d},\mathbb{R})$.

Teorema 2.9. *Se $d = 1$ e esiste $m > 0$ tale che $\inf_{x,y}\partial_y^2\mathcal{L}(x,y) \geq m$, allora i moti che soddisfano le equazioni di Lagrange con condizioni ai bordi $q(t_0) = A$, $q(t_1) = B$ sono minimi locali dell'azione.*

Proof. Argomentiamo come nel Lemma 2.6, ma questa volta sviluppiamo la Lagrangiana fino al secondo ordine[6]

$$\begin{aligned}\mathcal{L}(q(s)+h(s),\dot q(s)+\dot h(s)) &= \mathcal{L}(q(s),\dot q(s)) + \partial_q\mathcal{L}(q(s),\dot q(s))h(s)\\ &+ \partial_q\mathcal{L}(q(s),\dot q(s))\dot h(s) + \frac12\partial_q\partial_q\mathcal{L}(q(s),\dot q(s))h(s)^2\\ &+ \partial_q\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))h(s)\dot h(s) + \frac12\partial_{\dot q}\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))(\dot h(s))^2\\ &+ o(h(s)^2 + (\dot h(s))^2).\end{aligned}$$

A questo punto si noti che, poichè le derivate miste sono uguali visto che $\mathcal{L} \in \mathcal{C}^2$,

$$\begin{aligned}\int_{t_0}^{t_1}\partial_q\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))h(s)\dot h(s) &= \frac12\int_{t_0}^{t_1}\partial_q\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\frac{d}{ds}(h(s)^2)\\ &= -\frac12\int_{t_0}^{t_1}\frac{d}{ds}\partial_q\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\cdot h(s)^2.\end{aligned} \tag{2.2}$$

[6]Visto che sappiamo dallo studio delle funzioni da $\mathbb{R}^n$ a $\mathbb{R}$ che i massimi e minimi dipendono dalle proprità della derivata seconda.

D'altro canto, poichè $\mathcal{L} \in \mathcal{C}^3$ e quindi anche l'ordine delle derivate terze non ha importanza,

$$\begin{aligned}\frac{d}{ds}\partial_q\partial_{\dot q}\mathcal{L}(q(s),\dot q(s)) =&\partial_q\partial_q\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\dot q + \partial_{\dot q}\partial_q\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\ddot q\\ &= \partial_q\left[\partial_q\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\dot q + \partial_{\dot q}\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\ddot q\right]\\ &= \partial_q\left[\frac{d}{ds}\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\right].\end{aligned}$$

Possiamo dunque usare le relazioni precedenti per scrivere, ricordando che le q soddisfano le equazioni di Lagrange,

(2.3)

$$\begin{aligned}S(q+h)-S(q) =&\frac{1}{2}\int_{t_0}^{t_1}\left[\partial_q\partial_q\mathcal{L}(q(s),\dot q(s))h(s)^2 + \partial_{\dot q}\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\dot h(s)^2\right]ds\\ &-\frac{1}{2}\int_{t_0}^{t_1}\partial_q\left[\frac{d}{ds}\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))\right]\cdot h(s)^2 ds + o\left(\int_{t_0}^{t_1} h(s)^2 + (\dot h(s))^2\right)\\ =&\frac{1}{2}\int_{t_0}^{t_1}\partial_{\dot q}\partial_{\dot q}\mathcal{L}(q(s),\dot q(s))(\dot h(s))^2 ds + o\left(\int_{t_0}^{t_1} h(s)^2 + (\dot h(s))^2\right)\\ \geq&\frac{m}{2}\int_{t_0}^{t_1}(\dot h(s))^2 ds + o\left(\int_{t_0}^{t_1} h(s)^2 + (\dot h(s))^2\right).\end{aligned}$$

Per concludere si noti che, usando la disuguaglianza di Schwarz,

$$\int_{t_0}^{t_1} h(s)^2 = \int_{t_0}^{t_1}\left(\int_{t_0}^{s}\dot h(\tau)d\tau\right)^2 \leq \int_{t_0}^{t_1}\int_{t_0}^{s}(\dot h(\tau))^2 d\tau \leq (t_1-t_0)\int_{t_0}^{t_1}(\dot h(s))^2 ds.$$

Questo significa che se $\|h\|_{\mathcal{C}^1}$ è sufficientemente piccolo, l'errore in (2.3) è più piccolo del primo termine (che è positivo). In altre parole, per qualunque moto sufficientemente vicino a q l'azione ha un valore più grande di $S(q)$, ovvero q è un minimo locale dell'azione. □

A questo punto viene naturale chiedersi che cosa accade se $d > 1$. Ovviamente si può procedere nello stesso modo anche se ora lo sviluppo di Taylor della Lagrangiana è più complesso. Tuttavia l'argomento usato nell'equazione (2.2) è sostanzialmente unidimensionale. Non è quindi strano che in più dimensioni il risultato sia leggermente più debole.

Teorema 2.10. *Se esiste $m > 0$ tale che $\inf_{x,y}\inf_v \frac{\langle v,\partial_y^2\mathcal{L}(x,y)v\rangle}{\|v\|^2} \geq m$ e se $t_1 - t_0$ è sufficientemente piccolo, allora i moti che soddisfano le equazioni di Lagrange con condizioni ai bordi $q(t_0) = A$, $q(t_1) = B$ sono minimi locali dell'azione.*

Proof. Nuovamente sviluppiamo la Lagrangiana fino al secondo ordine

$$\begin{aligned}&\mathcal{L}(q(s)+h(s),\dot q(s)+\dot h(s)) = \mathcal{L}(q(s),\dot q(s)) + \sum_i \partial_{q_i}\mathcal{L}(q(s),\dot q(s))h_i(s)\\ &+\sum_i \partial_{q_i}\mathcal{L}(q(s),\dot q(s))\dot h_i(s) + \frac{1}{2}\sum_{ij}\partial_{q_i}\partial_{q_j}\mathcal{L}(q(s),\dot q(s))h_i(s)h_j(s)\\ &+\sum_{ij}\partial_{q_i}\partial_{\dot q_j}\mathcal{L}(q(s),\dot q(s))h_i(s)\dot h_j(s) + \frac{1}{2}\sum_{ij}\partial_{\dot q_i}\partial_{\dot q_j}\mathcal{L}(q(s),\dot q(s))\dot h_i(s)\dot h_j(s)\\ &+o(\|h(s)\|^2 + \|\dot h(s)\|^2).\end{aligned}$$

In analogia con quanto fatto prima possiamo scrivere

$$\begin{aligned}
S(q+h)-S(q)=&\frac{1}{2}\int_{t_0}^{t_1}\sum_{i,j}\partial_{q_i}\partial_{q_j}\mathcal{L}(q(s),\dot q(s))h_i(s)h_j(s)\\
&+\frac{1}{2}\int_{t_0}^{t_1}\sum_{i,j}\partial_{\dot q_i}\partial_{\dot q_j}\mathcal{L}(q(s),\dot q(s))\dot h_i(s)\dot h_j(s)ds\\
&+\int_{t_0}^{t_1}\sum_{i,j}\partial_{q_i}\partial_{\dot q_j}\mathcal{L}(q(s),\dot q(s))\cdot h_i(s)\dot h_j(s)ds\\
&+o\left(\int_{t_0}^{t_1}\|h(s)\|^2+\|\dot h(s)\|^2\right)\\
\geq&\frac{m}{2}\int_{t_0}^{t_1}\|\dot h_i(s)\|^2ds-C\int_{t_0}^{t_1}\left[\|h(s)\|^2+\|\dot h(s)\|\|h(s)\|\right]ds\\
&+o\left(\int_{t_0}^{t_1}\|h(s)\|^2+\|\dot h(s)\|^2\right),
\end{aligned}$$

dove C è una costante determinata dal massimo del valore assoluto delle derivate seconde lungo la curva q e abbiamo usato la disuguaglianza di Schwartz nella forma

$$\left|\sum_{i=1}^d a_i\right|\leq\left[\sum_{i=1}^d a_i^2\right]^{\frac{1}{2}}\sqrt{d}.$$

D'altro canto, in maniera simile a quanto fatto nel caso unidimensionale,

$$\int_{t_0}^{t_1}\|h(s)\|^2=\int_{t_0}^{t_1}\left\|\int_{t_0}^{s}\dot h(\tau)d\tau\right\|^2\leq\int_{t_0}^{t_1}\int_{t_0}^{s}\|\dot h(\tau)\|^2d\tau\leq(t_1-t_0)\int_{t_0}^{t_1}\|\dot h(s)\|^2ds.$$

Per trattare il termine misto si noti che[7]

$$2\int_{t_0}^{t_1}\|h(s)\|\|\dot h(s)\|=\int_{t_0}^{t_1}(t_1-t_0)^{-\frac{1}{2}}\|h(s)\|^2+(t_1-t_0)^{\frac{1}{2}}\|\dot h(s)\|^2ds.$$

Da cui segue che

$$S(q+h)-S(q)\geq\frac{1}{2}\left[1-C(t_1-t_0)^{\frac{1}{2}}\right]\int_{t_0}^{t_1}\|\dot h(s)\|^2ds+o\left(\|\dot h(s)\|^2\right).$$

Quindi nuovamente q è un minimo locale a patto che $C(t_1-t_0)^{\frac{1}{2}}<1$. □

Tuttavia si noti che non abbiamo mostrato che l'azione ha sempre dei minimi o, più in generale, dei punti stazionari. La cosa non è ovvia come mostra il seguente esempio.

Esercizio 2.11. *Si consideri la Lagrangiana* $\mathcal{L}(q,\dot q)=\frac{1}{2}\dot q^2-\frac{1}{2}q^2$, $q\in\mathbb{R}$, *e* $t_0=0$ *e* $t_1=2\pi$. *Si mostri che l'azione non ha punti stazionari in* $\mathcal{M}_{t_0,t_1}(0,1)$. *(Suggerimento: se avesse punti stazionari allora dovrebbero soddisfare le equazioni di Lagrange e* $q(0)=0$, $q(2\pi)=1$. *Si risolvano le equazioni di Lagrange e si veda che tali soluzioni non esistono.)*

---

[7]Qui stiamo usando la disuguaglianza $2|ab|\leq\delta^2a^2+\delta^{-2}b^2$ che si ottiene banalmente dal fatto che $(\delta a-\delta^{-1}b)^2\geq 0$ e $(\delta a+\delta^{-1}b)^2\geq 0$.

3. Esistenza dei minimi e equazioni differenziali con condizioni al bordo

Per concludere discutiamo un attimo la relazione tra il problema di Cauchy e il problema al bordo che si ottiene minimizzando l'azione. Come abbiamo già fatto nello studio dell'azione assumiamo che esista $m > 0$ tale che, per ogni $v \in \mathbb{R}^d$,

$$\inf_{x,y} \langle v, \partial_y^2 \mathcal{L}(x,y) v\rangle \geq m\|v\|^2. \tag{3.1}$$

Esercizio 3.1. *Si mostri che (3.1) implica che la matrice con elementi $\partial_{\dot{q}_i}\partial_{\dot{q}_j}\mathcal{L}$ è invertibile*

Si noti che le equazioni di Lagrange hanno la forma

$$\partial_{\dot{q}}^2 \mathcal{L}\ddot{q} = G(q, \dot{q})$$

per una qualche funzione G. Dunque possiamo scrivere

$$\ddot{q} = \left[\partial_{\dot{q}}^2 \mathcal{L}\right]^{-1} G(q,\dot{q}) = F(q,\dot{q}). \tag{3.2}$$

Ci siamo così ridotti a delle equazioni differenziali del secondo ordine nella forma solita. Si noti che se $\mathcal{L} \in \mathcal{C}^4$ allora $F \in \mathcal{C}^2$. Il risultato di base per chiarificare il problema che ci siamo posti è il seguente.

Teorema 3.2. *Data $F \in \mathcal{C}^2(\mathbb{R}^{2d}, \mathbb{R}^d)$, $t_0, t_1 \in \mathbb{R}$ e $A, B \in \mathbb{R}^d$ il problema*

$$\begin{aligned} &\ddot{q} = F(q,\dot{q}) \\ &q(t_0) = A, \quad q(t_1) = B \end{aligned}$$

ammette una unica soluzione a patto che $t_1 - t_0$ sia sufficientemente piccolo.

Proof. L'idea di base è quella di confrontare il problema con quello di Cauchy per cui abbiamo già molti risultati. A questo scopo è meglio trasformare l'equazione di partenza in una equazione del primo ordine: sia $z = (q,p) \in \mathbb{R}^{2d}$, $q, p \in \mathbb{R}^d$, e $\overline{F}(z,\lambda) = (p, \lambda F(q,p))$, $\lambda \in [0,1]$, allora l'equazione che ci interessa può essere scritta come[8]

$$\dot{z} = \overline{F}(z, 1).$$

La ragione per cui abbiamo introdotto il parametro λ è che l'equazione $\dot{z} = \overline{F}(z,0)$, $z(t_0) = A$, $z(t_1) = B$ si può facilmente risolvere esplicitamente e la sua unica soluzione è $\bar{z}(t) = (A + (B-A)\frac{t-t_0}{t_1-t_0}, (B-A)(t_1-t_0)^{-1})$. Possiamo quindi considerare il problema di Cauchy

$$\begin{aligned} &\dot{z} = \overline{F}(z,\lambda) \\ &z(0) = (A, v). \end{aligned}$$

Sappiamo che tale problema ha una unica soluzione $z(t,v,\lambda) = (q(t,v,\lambda), p(t,v,\lambda))$. Sembra quindi naturale esprimere tale soluzione come funzione di v, λ e cercare di risolvere l'equazione

$$q(t_1, v, \lambda) - B = 0. \tag{3.3}$$

Se è possibile (per $\lambda = 1$) questo fornisce una soluzione del nostro problema. L'approccio ovvio è quello di cercare di usare il teorema della funzione implicita. Questo ci dice che, in un intorno di $\lambda = 0$ e $v = (B-A)(t_1-t_0)^{-1}$ la equazione (3.3)

[8] Questo non è un sistema molto furbo di trasformare le equazioni di Lagrange in un sistema del primo ordine. Esiste un formalismo (il formalismo *Hamiltoniano*) che permette di farlo in maniera molto più efficiente e profonda. Tuttavia per i nostri scopi è sufficiente.

ha una unica soluzione a patto che la matrice $d \times d$ data da $\xi(t,v,\lambda) = \partial_v q(t,v,\lambda)$ sia invertibile per $t = t_1$. Il teorema sulla dipendenza differenziabile dalle condizioni iniziali e dai parametri implica che $z(t_1,v,\lambda)$ è una funzione differenziabile di v e λ e inoltre, ponendo $\eta(t,v,\lambda) = \partial_v p(t,v,\lambda)$ abbiamo

$$\begin{aligned}
&\dot{\xi} = \eta \\
&\dot{\eta} = \partial_q F(z)\xi + \partial_p F(z)\eta \\
&\xi(0) = 0 \qquad \eta(0) = \mathbb{1}.
\end{aligned}$$

Poichè le equazioni sono lineari sappiamo che esiste una costante $C \geq \|F\|_{\mathcal{C}^1}$, tale che $\|\xi(t)\| + \|\eta(t)\| \leq e^{C(t-t_0)}$. Dunque

$$\xi(t) = \int_{t_0}^{t} \eta(s)ds = \mathbb{1}(t-t_0) + \int_{t_0}^{t} ds \int_{t_0}^{s} \partial_q F(z(s_1))\xi(s_1) + \partial_p F(z(s_1))\eta(s_1)ds_1.$$

Remark 3.3. *Si noti che nelle righe sopra abbiamo introdotto la norma di una matrice $d \times d$. Questa non è definita in modo canonico, esiste tuttavia una scelta particolarmente felice: sia D una matrice $d \times d$ allora definiamo*

$$\|D\| = \sup_{v \in \mathbb{R}^d \setminus \{0\}} \frac{\|Dv\|}{\|v\|}. \tag{3.4}$$

Esercizio 3.4. *Si mostri che la (3.4) definisce una norma nello spazio della matrici. Inoltre si mostri che, date comunque due matrici $d \times d$ D, E, si ha che $\|DE\| \leq \|D\| \cdot \|E\|$.*

Quindi

$$\|\xi(t_1) - \mathbb{1}(t_1 - t_0)\| \leq \int_{t_0}^{t_1} ds \int_{t_0}^{s} 2Ce^{C(s_1-t_0)}ds_1 \leq Ce^{C(t_1-t_0)}(t_1-t_0)^2.$$

Possiamo quindi scrivere $\xi(t_1,v,\lambda) = (t_1 - t_0)\,[\mathbb{1} - D]$ dove $\|D\| \leq Ce^{C(t_1-t_0)}(t_1 - t_0)$.

Esercizio 3.5. *Si mostri che se $\|D\| < 1$ allora $\mathbb{1} - D$ è invertibile. (suggerimento: si mostri che la serie $\sum_{n=0}^{\infty} D^n$ converge in norma e che quindi definisce una matrice E. Si verifichi che $E(\mathbb{1} - D) = (\mathbb{1} - D)E = \mathbb{1}$.)*

Dunque se $Ce^{C(t_1-t_0)}(t_1 - t_0) < 1$, allora $\xi(t_1)$ è invertibile ed il teorema della funzione implicita si applica per ogni λ sufficientemente piccolo. Questo significa che per tali λ il nostro problema ha un'unica soluzione e questa dipende con continuità da λ. Questo implica che l'insieme dei λ per cui esiste la soluzione è chiuso. Sia λ_1 il suo supremo. Ma ora possiamo fare lo stesso argomento partendo da λ_1 invece che da zero. Per gli stessi esatti argomenti il teorema della funzione implicita si applicherà e otterremo una unica soluzione per tutti i λ in un intorno di λ_1, contrariamente all'ipotesi che λ_1 fosse il più grande λ per cui esisteva una soluzione. Quindi il nostro problema ha una soluzione per tutti i λ, ed in particolare per $\lambda = 1$, il che implica il teorema. □

Abbiamo così visto che il problema al contorno determinato dal minimizzare l'azione ha una unica soluzione (a patto che lo si consideri per tempi sufficientemente piccoli) e quindi da un lato è equivalente ad un qualche (più convenzionale) problema di Cauchy e dall'altro significa che l'azione deve avere un minimo. Si noti infine che sebbene la soluzione possa esistere per tempi più lunghi, essa può non essere unica, come si evince risolvendo i seguenti esercizi.

Esercizio 3.6. *Si consideri la Lagrangiana* $\mathcal{L}_1(x, \dot{x}) = \frac{1}{2}\|\dot{x}\|^2$, $q \in \mathbb{R}^3$ *e sia dato il vincolo* $\|x\| = 1$ *(ovvero il moto si svolge sulla superficie della sfera unitaria). Si scriva la Lagrangiana nelle coordinate*

$$x = (\cos\theta \sin\varphi, \sin\theta \sin\varphi, \cos\varphi)$$

dove $\theta \in [0, 2\pi]$ *e* $\varphi \in [0, \pi]$. *Si faccia lo stesso per la Lagrangiana* $\mathcal{L}_2 = \|\dot{x}\|$ *e si mostri che ha gli stessi punti stazionari di* $\mathcal{L}_1$.

Esercizio 3.7. *Si mostri che, per ogni* $T > 0$, *e* $\mathcal{M}_{0,T}(A, B)$ *con* $A = (0, 0, 1)$ *e* $B = (0, 0, -1)$, *l'azione associata alla Lagrangiana* $\mathcal{L}_1$ *ha infiniti punti stazionari.*

Esercizio 3.8. *Si noti che l'azione associata alla lagrangiana* $\mathcal{L}_2$ *non è altro che la lunghezza della curva su cui avviene il moto. Si usi questo fatto per dare una interpretazione geometrica degli esercizi precedenti.*

Carlangelo Liverani, Dipartimento di Matematica, II Università di Roma (Tor Vergata), Via della Ricerca Scientifica, 00133 Roma, Italy.

Email address: `liverani@mat.uniroma2.it`