

Chapter 9

Elogio della pigrizia

9.1 Lagrangiana

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

$$M\ddot{x} = -\nabla U \quad (9.1.1)$$

dove $x : \mathbb{R} \rightarrow \mathbb{R}^d$, 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

$$\mathcal{L}(x, y) = \frac{1}{2} \langle y, My \rangle - U(x). \quad (9.1.2)$$

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

$$\frac{d}{dt} \partial_y \mathcal{L}(x(t), \dot{x}(t)) - \partial_x \mathcal{L}(x(t), \dot{x}(t)) = 0. \quad (9.1.3)$$

Tali equazioni sono dette *equazioni di Lagrange*.

Esercizio 9.1.1. *Si mostri che le equazioni di Lagrange associate alla Lagrangiana (9.1.2) sono esattamente le equazioni di Newton (9.1.1).*

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 9.1.2. *Si consideri il cambio di coordinate $x = \Xi(q)$ allora le equazioni di Newton (9.1.1) nelle coordinate q coincidono con le equazioni di Lagrange per la Lagrangiana $\tilde{\mathcal{L}}(\xi, \eta) = \mathcal{L}(\Xi(\xi), D\Xi(\xi)\eta)$.¹*

¹Normalmente si abusano le notazioni e la nuova funzione $\tilde{\mathcal{L}}$ 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 $\partial_\xi \tilde{\mathcal{L}}(\xi, \eta)$ e $\partial_\eta \tilde{\mathcal{L}}(\xi, \eta)$) sebbene tali espressioni, per quanto intuitive, non abbiano molto senso. Nel seguito mi conformerò a questa discutibile, ma conveniente, tradizione.

Proof. Prima di tutto occorre calcolare come sono fatte le equazioni (9.1.1) nelle nuove coordinate. Si noti che

$$\begin{aligned}\dot{x} &= D\Xi(q)\dot{q} \\ \ddot{x}_i &= \sum_{j,k} \partial_{q_j} \partial_{q_k} \Xi_i(q) \dot{q}_j \dot{q}_k + \sum_j \partial_{q_j} \Xi_i(q) \ddot{q}_j \\ \partial_{q_i}(U \circ \Xi)(q) &= \sum_j (\partial_{x_j} U) \circ \Xi(q) \cdot \partial_{q_i} \Xi_j(q).\end{aligned}\tag{9.1.4}$$

Poichè Ξ è un diffeomorfismo $D\Xi$ deve essere invertibile, quindi le equazioni (9.1.1) sono equivalenti a

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

Ovvero, ponendo $\tilde{U} = U \circ \Xi$ e usando le (9.1.4) si ha

$$\sum_{l,j,k,r} \partial_{q_i} \Xi_l \cdot M_{l,r} \cdot \partial_{q_j} \partial_{q_k} \Xi_r \cdot \dot{q}_j \dot{q}_k + \sum_{l,j,r} \partial_{q_i} \Xi_l \cdot M_{l,r} \cdot \partial_{q_j} \Xi_r \ddot{q}_j = -\partial_{q_i} \tilde{U}. \tag{9.1.5}$$

D'altro canto

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

e inserendo queste espressioni nelle equazioni di Lagrange (9.1.3) si ottengono esattamente le (9.1.5), come annunciato. $\square$

Il vantaggio di cambiare facilmente coordinate è particolarmente evidente quanto si considerano sistemi vincolati.

9.2 Vincoli e principio dei lavori virtuali

Se il moto avviene sotto un vincolo Φ , ovvero, detto $x(t)$ il moto soddisfa $\Phi(x(t)) = 0$ per ogni t , allora devono essere in azione delle forze che costringono il moto a rispettare il vincolo. Queste forze vengono chiamate *reazioni vincolari*. Nella sezione 8 abbiamo visto una semplice discussione di come queste forze possano emergere da un modello più dettagliato del moto. Visto che le

reazioni vincolari dipendono in principio la come è realizzato il vincolo, essere sono incognite. Tuttavia, per *vincoli ideali* si assume che esse non facciano lavoro. Questa assunzione è una vera e propria aggiunta alle equazioni di Newton che è verificata sperimentalmente in moltissimi casi, ma non è necessariamente vera in tutte le situazioni.

Si ricordi che il lavoro fatto da una forza F mentre un corpo compie lo spostamento $x(t)$ dal tempo t_0 al tempo t_1 è definito come

$$L(t_0, t_1) = \int_{t_0}^{t_1} \langle F(x(t)), \dot{x}(t) \rangle dt.$$

Ne segue che se le reazioni vincolari non devono fare lavoro per qualunque spostamento compatibile col vincolo, allora deve essere

$$\int_t^{t_1} \langle R(x(s)), \dot{x}(s) \rangle dt = 0.$$

Derivando rispetto a t_1 al tempo t si ha

$$\langle R(x(t)), \dot{x}(t) \rangle = 0.$$

Poichè sia $x(t)$ che $\dot{x}(t)$ sono arbitrari, a parte il fatto che $\Phi(x(t)) = 0$, che, differenziando, implica $D\Phi\dot{x}(t) = 0$, abbiamo ottenuto il seguente fatto.

Definizione 9.2.1 (Principio dei lavori virtuali). *Dato un vincolo $\Phi \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^{d-m})$, $m < d$, il principio dei lavori virtuali asserisce che per ogni $x \in \mathbb{R}^d$ e $v \in \mathbb{R}^d$ tali che $\Phi(x) = 0$ e $D_x\Phi v = 0$, la reazione vincolare R soddisfa l'equazione*

$$\langle R(x), v \rangle = 0.$$

9.3 Equazioni del moto per sistemi vincolati

Dato un sistema vincolato sotto l'azione di un potenziale esterno, il moto può essere espresso dalle seguenti *equazioni cardinali della meccanica*:

$$\begin{aligned} M\ddot{x}(t) &= -\nabla U(x) + R(t) \\ \Phi(x) &= 0 \\ \langle R(t), v \rangle &= 0 \text{ per tutti i } v \in \mathbb{R}^d \text{ tali che } D\Phi(x(t))v = 0. \end{aligned} \tag{9.3.1}$$

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.

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. Ovvero $\Phi \circ \Xi \equiv 0$, che, differenziando,

implica $D\Phi D\Xi = 0$. Infine, poichè il kernel di Φ deve essere m dimensionale, tutti i vettori tali $D\Phi v = 0$ si possono esprimere 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

$$D\Xi(q(t))^T R(t) = 0. \quad (9.3.2)$$

Possiamo quindi moltiplicare la prima delle (9.3.1) per $D\Xi^T$ e ottenere

$$D\Xi^T M \frac{d}{dt} (D\Xi \dot{q}) = -D\Xi^T \nabla U(\Xi q). \quad (9.3.3)$$

Queste sono $2m$ equazioni differenziali per le $2m$ funzioni incognite $q(t), \dot{q}(t)$. 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 9.3.1. Si usino le (9.1.4) per scrivere esplicitamente le (9.3.3) in termini delle funzioni q e si verifichi che coincidono con le equazioni di Lagrange per la Lagrangiana $\mathcal{L}(q, \dot{q}) = \frac{1}{2} \langle D\Xi(q)\dot{q}, MD\Xi(q)\dot{q} \rangle - U(\Xi(q))$.

Commento 9.3.2. 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.

9.4 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 $S : \mathcal{M}_{t_0, t_1}(A, B) \rightarrow \mathbb{R}$ come

$$S(q) = \int_{t_0}^{t_1} \mathcal{L}(q(s), \dot{q}(s)) ds. \quad (9.4.1)$$

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

²Ovvero tutte le curve (i moti) che, partendo al tempo t_0 , nel tempo $t_1 - t_0$ vanno dal punto A al punto B .

Un attimo di riflessione mostra che per rispondere a tali domande occorre prima di tutto chiarirne il significato. Per esempio, una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ è continua se solo se per ogni successione convergente x_n si ha $\lim_{n \rightarrow \infty} f(x_n) = f(\lim_{n \rightarrow \infty} x)$. 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

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

Esercizio 9.4.1. *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.*

Esercizio 9.4.2. *Si mostri che $d : \mathcal{M}_{t_0, t_1}(A, B)^2 \rightarrow \mathbb{R}_+$ definita da*

$$d(q, \bar{q}) = \sup_{s \in [t_0, t_1]} \|q(s) - \bar{q}(s)\| + \sup_{s \in [t_0, t_1]} \|\dot{q}(s) - \dot{\bar{q}}(s)\|$$

è una metrica. Quindi $(\mathcal{M}_{t_0, t_1}(A, B), d)$ è uno spazio metrico completo.

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 $\lim_{n \rightarrow \infty} S(q_n) = S(\bar{q})$.

Esercizio 9.4.3. *Si mostri che la definizione di continuità di cui sopra è equivalente alla seguente: S è continua in $\bar{q}$ se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che $|S(q) - S(\bar{q})| \leq \varepsilon$ per ogni $d(q, \bar{q}) < \delta$.*

Esercizio 9.4.4. *Si mostri che se $\mathcal{L} \in \mathcal{C}^0$ allora S è continua.*

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

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

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

³Infatti 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 $\partial_{x_i} f$.

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

Esercizio 9.4.7. 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.

Definizione 9.4.8. Uno spazio vettoriale $\mathbb{V}$ con una norma $\|\cdot\|$ si dice uno spazio normato. Se lo spazio è completo, ovvero ogni successione di Cauchy è convergente, allora lo spazio si chiama di Banach.

Definizione 9.4.9. Diremo dunque che S è differenziabile in $q \in \mathcal{M}_{t_0, t_1}(A, B)$ se esiste una funzione lineare continua $L_q : \mathcal{M}_{t_0, t_1}(0, 0) \rightarrow \mathbb{R}$ tale che, per ogni successione $\{h_n\} \subset \mathcal{M}_{t_0, t_1}(0, 0)$ tale che $\lim_{n \rightarrow \infty} \|h_n\|_{\mathcal{C}^1} = 0$ si ha

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

Oppure, più brevemente,

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

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

Il lettore potrebbe pensare che la richiesta che L_q sia continua è superflua in quanto le funzioni lineari sono continue. Sfortunatamente questo è vero solo in spazi finito dimensionali, mentre è in generale falso per spazi di Banach.⁵ Ovviamente non gradiamo che l'approssimazione della differenza tra $S(q + h) - S(q)$ non sia continua in h , ed è quindi naturale richiedere la continuità. Si noti che dato uno spazio di Banach $\mathbb{V}$ e un operatore lineare $L : \mathbb{V} \rightarrow \mathbb{R}$, è naturale definire la norma

$$\|L\| = \sup_{\|x\| \leq 1} |L(x)|.$$

Se $\|L\| < \infty$ allora L si dice *limitato*. Poichè su spazi di Banach esistono operatori lineari non limitati, verrebbe da pensare che sia meglio richiedere anche che il differenziale sia limitato, tuttavia il prossimo Lemma mostra che questo è inutile.

⁴Si ricordi che, dato uno spazio vettoriale $\mathbb{V}$ sul campo dei reali, una funzione $\|\cdot\| : \mathbb{V} \rightarrow \mathbb{R}_+$ 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 $x, y \in \mathbb{V}$.

⁵Si noti comunque che l'esibizione di tali esempi non è costruttiva in quanto utilizza l'assioma della scelta.

Lemma 9.4.10. *Dato uno spazio di Banach $\mathbb{V}$ una funzione lineare $L : \mathbb{V} \rightarrow \mathbb{R}$ è continua se e solo se è limitata.*

Proof. Se la funzione è limitata allora

$$|L(x) - L(y)| = |L(x - y)| \leq |L((x - y)\|x - y\|^{-1})|\|x - y\| \leq \|L\|\|x - y\|$$

da cui segue la continuità. D'altro canto che la funzione è continua allora, per la definizione di continuità, esiste δ tale che $|L(x)| \leq 1$ per ogni $\|x\| \leq \delta$. Dunque per ogni $x \in \mathbb{V}$ tale che $\|x\| \leq 1$,

$$|L(x)| = \delta^{-1}|L(\delta x)| \leq \delta^{-1}$$

dunque $\|L\| \leq \delta^{-1}$, ovvero L è limitato. $\square$

Sperabilmente le considerazioni di cui sopra avranno convinto il lettore che la definizione 9.4.9 è l'estensione naturale del caso finito dimensionale.

Lemma 9.4.11. *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

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

Dunque

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

Il Lemma segue immediatamente dall'ultima equazione. $\square$

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 9.4.12. *Se $\mathcal{L} \in \mathcal{C}^2$ e $q \in \mathcal{M}_{t_0, t_1}(A, B) \cap \mathcal{C}^2$ 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

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

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 Lagrange non fossero soddisfatte, allora esisterebbe un i e un intervallo di tempi in cui $\partial_{q_i} \mathcal{L}(q(s), \dot{q}(s)) - \frac{d}{dt} \partial_{\dot{q}_i} \mathcal{L}(q(s), \dot{q}(s))$ 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 $L_q(h) = 0$. $\square$

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 9.4.13. *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 $L_q(h) = 0$. D'altro canto

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

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

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

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 $x = \dot{q}(s)$ e $\dot{q} \in \mathcal{C}^1$. $\square$

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 9.4.14. *Se $\mathcal{L} \in \mathcal{C}^3(\mathbb{R}^{2d}, \mathbb{R})$, $d = 1$ e $\inf_{x,y} \partial_y^2 \mathcal{L}(x, y) \geq m$ per qualche $m > 0$, 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 9.4.11, ma questa volta sviluppiamo la Lagrangiana fino al secondo ordine⁶

$$\begin{aligned} \mathcal{L}(q(s) + h(s), \dot{q}(s) + \dot{h}(s)) &= \mathcal{L}(q(s), \dot{q}(s)) + \partial_q \mathcal{L}(q(s), \dot{q}(s))h(s) \\ &+ \partial_q \mathcal{L}(q(s), \dot{q}(s))\dot{h}(s) + \frac{1}{2} \partial_q \partial_q \mathcal{L}(q(s), \dot{q}(s))h(s)^2 \\ &+ \partial_q \partial_{\dot{q}} \mathcal{L}(q(s), \dot{q}(s))h(s)\dot{h}(s) + \frac{1}{2} \partial_{\dot{q}} \partial_{\dot{q}} \mathcal{L}(q(s), \dot{q}(s))(\dot{h}(s))^2 \\ &+ o(h(s)^2 + (\dot{h}(s))^2). \end{aligned}$$

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

$$\begin{aligned} \int_{t_0}^{t_1} \partial_q \partial_{\dot{q}} \mathcal{L}(q(s), \dot{q}(s))h(s)\dot{h}(s) &= \frac{1}{2} \int_{t_0}^{t_1} \partial_q \partial_{\dot{q}} \mathcal{L}(q(s), \dot{q}(s)) \frac{d}{ds} (h(s)^2) \\ &= -\frac{1}{2} \int_{t_0}^{t_1} \frac{d}{ds} \partial_q \partial_{\dot{q}} \mathcal{L}(q(s), \dot{q}(s)) \cdot h(s)^2. \end{aligned} \quad (9.4.2)$$

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

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

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

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

$$\begin{aligned}
S(q+h) - S(q) &= \frac{1}{2} \int_{t_0}^{t_1} \left[\partial_q \partial_q \mathcal{L}(q(s), \dot{q}(s)) h(s)^2 + \partial_{\dot{q}} \partial_{\dot{q}} \mathcal{L}(q(s), \dot{q}(s)) \dot{h}(s)^2 \right] ds \\
&\quad - \frac{1}{2} \int_{t_0}^{t_1} \partial_q \left[\frac{d}{ds} \partial_{\dot{q}} \mathcal{L}(q(s), \dot{q}(s)) \right] \cdot h(s)^2 ds + o \left(\int_{t_0}^{t_1} h(s)^2 + (\dot{h}(s))^2 \right) \\
&= \frac{1}{2} \int_{t_0}^{t_1} \partial_{\dot{q}} \partial_{\dot{q}} \mathcal{L}(q(s), \dot{q}(s)) (\dot{h}(s))^2 ds + o \left(\int_{t_0}^{t_1} h(s)^2 + (\dot{h}(s))^2 \right) \\
&\geq \frac{m}{2} \int_{t_0}^{t_1} (\dot{h}(s))^2 ds + o \left(\int_{t_0}^{t_1} h(s)^2 + (\dot{h}(s))^2 \right).
\end{aligned} \tag{9.4.3}$$

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

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

Questo significa che se $\|h\|_{C^1}$ è sufficientemente piccolo, l'errore in (9.4.3) 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. $\square$

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 (9.4.2) è sostanzialmente unidimensionale. Non è quindi strano che in più dimensioni il risultato sia leggermente più debole.

Teorema 9.4.15. *Se $\mathcal{L} \in C^3(\mathbb{R}^{2d}, \mathbb{R})$, $\inf_{x,y} \partial_y^2 \mathcal{L}(x,y) \geq m$ per qualche $m > 0$, 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

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

In analogia con quanto fatto prima possiamo scrivere

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

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,

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

Per trattare il termine misto si noti che⁷

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

Da cui segue che

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

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

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.

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

Esercizio 9.4.16. 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.)

9.5 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 $v \in \mathbb{R}^d$,

$$\inf_{x, y} \langle v, \partial_y^2 \mathcal{L}(x, y) v \rangle \geq m \|v\|^2. \quad (9.5.1)$$

Esercizio 9.5.1. Si mostri che (9.5.1) implica che la matrice con elementi $\partial_{q_i} \partial_{q_j} \mathcal{L}$ è invertibile

Si noti che le equazioni di Lagrange hanno la forma

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

per una qualche funzione G . Dunque possiamo scrivere

$$\ddot{q} = [\partial_q^2 \mathcal{L}]^{-1} G(q, \dot{q}) = F(q, \dot{q}). \quad (9.5.2)$$

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 9.5.2. 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 $\bar{F}(z, \lambda) = (p, \lambda F(q, p))$, $\lambda \in [0, 1]$, allora l'equazione che ci interessa può essere scritta come⁸

$$\dot{z} = \bar{F}(z, \lambda).$$

La ragione per cui abbiamo introdotto il parametro λ è che l'equazione $\dot{z} = \bar{F}(z, 0)$, $z(t_0) = A$, $z(t_1) = B$ si può facilmente risolvere esplicitamente e la sua

⁸Questo 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.

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 i problemi di Cauchy

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

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

$$q(t_1, v, \lambda) - B = 0. \quad (9.5.3)$$

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 (9.5.3) 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

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

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

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

Commento 9.5.3. 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

$$\|D\| = \sup_{v \in \mathbb{R}^d \setminus \{0\}} \frac{\|Dv\|}{\|v\|}. \quad (9.5.4)$$

Esercizio 9.5.4. Si mostri che la (9.5.4) 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 9.5.5. *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. $\square$

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 9.5.6. *Data una funzione a valore nello spazio delle matrici $M(q)$ tale che $M(q) \geq \nu \mathbb{1}$, $\nu > 0$, si mostri che $\mathcal{L}(q, \dot{q}) = \left[\frac{1}{2} \langle \dot{q}, M(q) \dot{q} \rangle \right]^{\frac{1}{2}}$ è costante lungo le soluzioni delle equazioni di Eulero-Lagrange.*

Esercizio 9.5.7. *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 9.5.8. *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 9.5.9. *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.*

9.6 Teoria globale

Si sarà notato che tutto quello che è stato detto finora è solo locale, visto che non sempre esistono coordinate globali. Per spiegare meglio questo fatto in

questa sezione discutiamo il moto di un punto libero vincolato a muoversi sulla superficie di una sfera. Ovvero il moto avviene su $S^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$.

La superficie di una sfera può essere descritta quindi localmente con due coordinate, ma la cosa non può essere globale, ovvero non può esistere $\psi : \mathbb{R}^2 \rightarrow S^2$ continua, suriettiva e invertibile (un omomorfismo). Infatti, un omomorfismo mappa compatti in compatti mentre ψ^{-1} mapperebbe S^2 (che è compatto) in $\mathbb{R}^2$ (che compatto non è).

Questo problema sta alla nascita della Geometria Differenziale a partire, almeno dal lavoro di Gauss e, soprattutto Riemann [5].

La traiettoria della particella è descritta da una curva $x \in \mathcal{C}^1([0, T], S^2)$, e la Lagrangiana ha la forma $\mathcal{L}(x, v) = \frac{1}{2}\|v\|^2$. Quindi l'azione è una funzione $S \in \mathcal{C}^0(\mathcal{C}^1([0, T], S^2), \mathbb{R})$ definita come

$$S(x) = \int_0^T \|\dot{x}(s)\|^2 ds.$$

In analogia con quanto fatto prima si considerino due punti $A, B \in S^2$ e lo spazio $\mathcal{M}_{0,T}(A, B) = \{x \in \mathcal{C}^0(\mathcal{C}^1([0, T], S^2), \mathbb{R}) : x(0) = A, x(T) = B\}$, e possiamo quindi restringere il dominio di S a $\mathcal{M}_{0,T}(A, B)$.

Si noti che fino ad ora non è stato necessario introdurre le coordinate lagrangiane.

Successivamente, vorremmo capire se S è differenziabile. Il primo ostacolo è capire su quale spazio lineare il differenziale dovrebbe agire. Si noti che se $x \in \mathcal{M}_{0,T}(A, B)$ e $h \in \mathcal{C}^1([0, T], \mathbb{R}^2)$ con $h(0) = h(T) = 0$, allora, per ogni $\lambda \in \mathbb{R}$, $\|x + \lambda h\| = 1$ se e solo se

$$\langle x, \|h\|^{-1}h \rangle = -\frac{\lambda}{2}\|h\|$$

Questo significa che, nel limite di piccolissime deformazioni della traiettoria, deve essere $\langle x, h \rangle = 0$, ovvero il cambiamento deve essere tangente alla sfera. Ne segue che, come ogni punto della sfera ha un diverso piano tangente, così ogni traiettoria in $\mathcal{M}_{0,T}(A, B)$ ha un diverso spazio tangente. Più precisamente, per ogni $x \in \mathcal{M}_{0,T}(A, B)$ possiamo definire

$$\mathcal{T}_x \mathcal{M}_{0,T}(0, 0) = \{h \in \mathcal{C}^1([0, T], \mathbb{R}^3) : h(0) = h(T) = 0, \langle x(t), h(t) \rangle = 0 \forall t \in [0, T]\}$$

Esercizio 9.6.1. Si mostri che $\mathcal{T}_x \mathcal{M}_{0,T}(0, 0)$ è uno spazio vettoriale.

Esercizio 9.6.2. Si mostri che $\|h\|_{\mathcal{C}^1} = \sup_{t \in [0, T]} \|h(t)\| + \sup_{t \in [0, T]} \|h'(t)\|$ è una norma e che con questa norma $\mathcal{T}_x \mathcal{M}_{0,T}(A, B)$ è uno spazio di Banach.

A questo punto vorremmo introdurre le coordinate Lagrangiane. Si noti che la Lagrangiana deve essere definita su $TS = \{(x, v) \in S \times \mathbb{R}^3 : \langle x, v \rangle = 0\}$, ovvero ad ogni punto $x \in S$ è associato il suo spazio tangente. L'insieme TS si chiama *fibrato tangente* e $\mathcal{L}(x, v) = \frac{1}{2}\|v\|^2$.

Si noti che TS si può naturalmente equipaggiare con la topologia indotta da $\mathbb{R}^6$, ed è quindi uno spazio topologico. Le coordinate Lagrangiane sarebbero delle funzioni continue e invertibili $\Psi : TS \rightarrow \mathbb{R}^4$. Sfortunatamente tali

coordinante non esistono occorre quindi lavorare localmente. Ovvero occorre un ricoprimento finito $\{U_i\} \subset TS$ di aperti tali che per ogni U_i esiste una funzione continua invertibile e suriettiva $\Psi_i : U_i \rightarrow V_i$, dove V_i è un aperto di $\mathbb{R}^6$. Chiaramente, $(x, v) = \Psi_i^{-1}(q, \dot{q})$ forniscono delle coordinate Lagrangiane locali, le *carte locali*. Quindi localmente la Lagrangiana è data da $\mathcal{L}(q, \dot{q}) = \mathcal{L} \circ \Psi_i^{-1}(q, \dot{q})$.

Siccome vogliamo essere in grado di fare derivate in maniera consistente è necessario chiedere che derivate fatte in sistemi di coordinate diverse siano compatibili e questo richiede che $\Psi_j \circ \Psi_i^{-1} : \mathbb{R}^6 \rightarrow \mathbb{R}^6$ sia differenziabile per tutti i punti per cui è definito (i posti dove la carta i e la carta j si sovrappongono).

Esercizio 9.6.3. *Si mostri che, lavorando localmente, si possono riprodurre tutti i risultati delle sezioni precedenti.*