G. Benettin, L. Galgani, A. Giorgilli
Appunti di Meccanica Razionale

*anno accademico 1991 – 92*

# 4

# I PRINCIPI VARIAZIONALI<br>LE TRASFORMAZIONI CANONICHE<br>L'EQUAZIONE DI HAMILTON – JACOBI<br>INTRODUZIONE ALLA TEORIA DELLE PERTURBAZIONI

## 4.1 I principi variazionali della meccanica

La via che abbiamo seguito fino ad ora per studiare le leggi del moto di un sistema di punti materiali si potrebbe chiamare "via differenziale", nel senso che abbiamo assunto, a suo fondamento, l'equazione differenziale di Newton $m\mathbf{a} = \mathbf{F}$, e da essa abbiamo dedotto, ancora nella forma di equazioni differenziali, le leggi generali del moto di un sistema meccanico (equazioni di Lagrange e di Hamilton). L'idea centrale, sottostante tutta la trattazione che abbiamo sin qui svolta, è che le traiettorie vere, o naturali, di un sistema meccanico sono, tra tutte quelle a priori pensabili, quelle che soddisfano, punto per punto, una certa relazione differenziale caratteristica del sistema.

In questo paragrafo vedremo una via diversa (anche se di fatto equivalente) per formulare le leggi generali della meccanica, in cui la traiettoria vera si distingue, tra tutte quelle a priori pensabili, per una proprietà globale, integrale, nello stesso senso in cui la retta si distingue, tra tutte le curve del piano, come la più breve tra due punti assegnati, o la traiettoria di un raggio luminoso in un mezzo a indice di rifrazione variabile si distingue (principio di Fermat) per il fatto che minimizza il tempo di percorrenza tra due punti assegnati.

Questa nuova via, più geometrica, si può chiamare "via variazionale," in quanto fa riferimento a quel settore dell'analisi chiamato calcolo delle variazioni; punto di arrivo sono i cosiddetti *principi variazionali* della meccanica, che portano il nome di "principi" proprio perchè da essi, assunti come principi primi, si può dedurre l'intera meccanica (in particolare, $m\mathbf{a} = \mathbf{F}$ per i punti materiali). Uno dei motivi di interesse della formulazione variazionale è anche il fatto che essa è particolarmente adatta per passare dalla meccanica classica alle teorie relativistiche, ove gli aspetti globali e geometrici sono determinanti; inoltre a essa è ispirata una formulazione particolarmente significativa della meccanica quantistica (la cosiddetta *"path integral formulation,"* introdotta da Feynman negli anni '50). Non è azzardato affermare che la formulazione variazionale di un qualunque problema è, almeno come linguaggio, forse la più generale possibile, e spesso anche la più profonda.

I primi tre paragrafi sono dedicati a una breve introduzione al calcolo variazionale (che non si suppone noto), utile a inquadrare i principi variazionali della meccanica nel contesto matematico più generale del calcolo delle variazioni. Il procedimento che seguiremo è a volte soltanto intuitivo, e non rigoroso quanto la materia richiederebbe; per una trattazione più estesa e rigorosa si rinvia ai testi di analisi.[1)]

### 4.1.1 Funzionali

Il problema elementare da cui ha origine il calcolo delle variazioni è la ricerca dei massimi e dei minimi, più in generale dei punti di stazionarietà, per funzioni reali di una o più variabili reali; come è ben noto, per una qualsiasi funzione regolare $F : \mathbb{R}^n \to \mathbb{R}$ la condizione perché $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ sia punto di stazionariatà è che si annulli in $\mathbf{x}$ il differenziale di $F$, ovvero che risulti $\frac{\partial F}{\partial x_i}(\mathbf{x}) = 0$, $i = 1, \ldots, n$. Si potrebbero ricordare altri problemi variazionali per funzioni definite in $\mathbb{R}^n$ (o in un suo sottoinsieme aperto $U$), come ad esempio la ricerca dei massimi e minimi condizionati. Il vero e proprio calcolo delle variazioni comincia tuttavia quando l'insieme di definizione di $F$ non è $\mathbb{R}^n$, nè alcun altro spazio finito-dimensionale, ma uno spazio di funzioni. Dato un insieme $U$ di funzioni – ad esempio, le funzioni reali differenziabili definite nell'intervallo $[a, b]$ – si dice *funzionale* definito nel dominio $U$ una legge o applicazione $F$ che ad ogni funzione $u \in U$ associ un numero reale. Una notazione comune per i funzionali è $F[u]$.

Semplici esempi di funzionali definiti sulle funzioni regolari dell'intervallo $[a, b]$ sono l'integrale di $u$, $F[u] = \int_a^b u(x)\,\mathrm{d}x$, una qualsiasi norma, ad esempio $F[u] = \left(\int_a^b u^2(x)\,\mathrm{d}x\right)^{1/2}$, oppure il valore della funzione $u$, o della sua derivata $u'$, in un punto fissato $\bar{x}$: $F[u] = u(\bar{x})$, $F[u] = u'(\bar{x})$. Il primo e i due ultimi sono lineari, perchè risulta $F[c_1 u_1 + c_2 u_2] = c_1 F[u_1] + c_2 F[u_2]$, il secondo evidentemente non lo è (si noti che la linearità del funzionale non ha niente a che vedere con la linearità di $u$). Si osservi che nell'ultimo esempio $F$ dipende da $u$ attraverso la sua derivata $u'$.

Un funzionale di notevole interesse in geometria e in meccanica è quello che dà la lunghezza di una curva. Nel caso più semplice, se $U$ è l'insieme di tutte le curve del piano cartesiano $xy$ della forma $y = u(x)$, $a \le x \le b$, allora la lunghezza della curva è data da $F[u] = \int_a^b [1 + {u'}^2(x)]^{\frac{1}{2}}\,\mathrm{d}x$; se invece usiamo le coordinate polari $r$ e $\vartheta$, e scriviamo la curva nella forma $r = u(\vartheta)$, $\vartheta_0 \le \vartheta \le \vartheta_1$, il funzionale lunghezza si scrive, come è immediato verificare, $F[u] = \int_{\vartheta_0}^{\vartheta_1} [u^2(\vartheta) + {u'}^2(\vartheta)]^{\frac{1}{2}}\,\mathrm{d}\vartheta$. Altre espressioni si trovano con altri sistemi di coordinate, o per superfici non piane.

---

[1)] Si veda ad esempio: I.M. Gelfand e S.V. Fomin, *Calculus of variations,* Prentice Hall, 1963. Per una esposizione matematica moderna si può anche vedere: Palais e Terng, *Critical point theory...,* Lect. Notes in Math. 1353, Springer.

- **Esercizio 4.1:** Per un cono di semi-apertura $\beta$ assegnata si usino come coordinate di un punto $P$ sulla superficie la distanza $r$ dal vertice lungo una direttrice e l'angolo $\varphi$ tra il semipiano uscente dall'asse del cono e passante per $P$, e un semipiano fissato ancora uscente dall'asse del cono (coordinate polari sferiche $r, \vartheta, \varphi$, con $\vartheta = \beta$). Si scriva il funzionale che dà la lunghezza di una curva di equazione $r = u(\varphi)$, tra $\varphi_0$ e $\varphi_1$ assegnati. (*Risposta:* $F[u] = \int_{\varphi_0}^{\varphi_1} [u^2(\varphi) \sin^2\beta + u'^2(\varphi)]^{\frac{1}{2}}\, \mathrm{d}\varphi$). Si scriva il medesimo funzionale, usando le coordinate polari $r, \vartheta$ dello sviluppo piano del cono (*Risposta:* se $r = u(\vartheta)$, allora $F[u] = \int_{\vartheta_0}^{\vartheta_1} [u^2(\vartheta) + u'^2(\vartheta)]^{\frac{1}{2}}\, \mathrm{d}\vartheta$, come per le curve del piano in coordinate polari).

Un classico problema è quello di trovare, per ciascuna superficie, le *geodetiche*, cioè le linee più brevi (più precisamente, di lunghezza stazionaria: si veda oltre) tra due punti qualsiasi della superficie.

Un problema importante in meccanica e in ottica è il calcolo del tempo di percorrenza di una traiettoria $\gamma$ assegnata da parte di un punto materiale, o di un raggio luminoso, la velocità dei quali dipenda in maniera nota dalla posizione: restringendoci per semplicità a moti piani, e a traiettorie rappresentabili come grafico di una funzione $y = u(x)$ tra due ascisse fissate $a$ e $b$, il tempo di percorrenza è dato da

$$T[u] = \int_\gamma \frac{ds}{v} = \int_a^b \frac{\sqrt{1 + u'^2(x)}}{v(x, u(x))}\, \mathrm{d}x \ ,$$

ove $v(x, y)$ è il modulo della velocità in $(x, y)$. In ottica la funzione $v$ è data direttamente dal rapporto $c/n(x, y)$, ove $c$ è la velocità della luce nel vuoto e $n$ l'indice di rifrazione del mezzo:

$$T[u] = c^{-1} \int_a^b n(x, u(x)) \sqrt{1 + u'^2(x)}\, \mathrm{d}x \ ; \tag{4.1}$$

il *principio di Fermat* assicura che le traiettorie effettive sono, tra tutte quelle a priori possibili che passano per due punti assegnati, quelle per cui $T[u]$ è minimo (più in generale stazionario, si veda oltre). Per il caso di un punto materiale che percorra una traiettoria assegnata, in assenza di forze esterne $v$ è costante, e il calcolo di $T$ si riporta a quello della lunghezza di una curva. Più in generale, se il sistema è conservativo, la funzione $v(x, y)$ dipende dall'energia $E$ e dal potenziale $V$: $v(x, y) = \left[\frac{2}{m}\big(E - V(x, y)\big)\right]^{1/2}$. Ad esempio, per un punto materiale che scenda in un piano verticale lungo una curva $y = u(x)$ passante per l'origine, partendo dall'origine con velocità nulla, risulta $v(x, y) = \sqrt{2gy}$ (si è scelto l'asse $y$ verticale discendente), e dunque si ha

$$T[u] = \frac{1}{\sqrt{2g}} \int_0^b \sqrt{\frac{1 + u'^2(x)}{u(x)}}\, \mathrm{d}x \ . \tag{4.2}$$

Si potrebbero considerare funzionali dipendenti esplicitamente da derivate di $u$ di ordine superiore al primo (per esempio il funzionale che dà la concavità massima di una funzione), ma non essendo importanti per i principi variazionali della meccanica cui questa introduzione è rivolta, non ce ne occuperemo, restringendo così la nostra attenzione ai soli funzionali $F[u]$ che dipendono esplicitamente da $u$ stessa e dalla sua derivata prima $u'$, oltre che dalla variabile indipendente $x$, come nell'esempio (4.1).

La nozione di funzionale si estende naturalmente al caso di dipendenza da due o più funzioni: ad esempio, il prodotto scalare

$$F[u,v] = \int_a^b u(x)v(x)\,\mathrm{d}x$$

è un funzionale che dipende da due funzioni $u$ e $v$ (ed è lineare in entrambe, o bilineare); la lunghezza di una curva dello spazio tridimensionale euclideo, definita dalle equazioni parametriche $x = u(t)$, $y = v(t)$, $z = w(t)$, $a \le t \le b$, è il funzionale

$$F[u,v,w] = \int_a^b [u'^2(t) + v'^2(t) + w'^2(t)]^{\frac{1}{2}}\,\mathrm{d}t\ .$$

### 4.1.2 *Variazione di un funzionale*

La nozione di variazione di un funzionale cui faremo riferimento è l'analogo della nozione di derivata direzionale per le funzioni di un numero finito di variabili, che brevemente ricordiamo.[2] Sia $F$ una funzione regolare di $n$ variabili reali, e $\mathbf{u} = (u_1, \ldots, u_n)$ un punto[3] interno al suo dominio di definizione $U \subset \mathbb{R}^n$. Fissata arbitrariamente una $n$–pla[4] $\delta\mathbf{u} = (\delta u_1, \ldots, \delta u_n) \in \mathbb{R}^n$, consideriamo i valori della funzione nei punti variati $\mathbf{u} + \alpha\delta\mathbf{u}$, per $\alpha$ reale in un intorno dell'origine;[5] la derivata direzionale (o variazione) $\delta F$ della funzione $F$, nel punto $\mathbf{u}$ e relativa al vettore (o alla variazione) $\delta\mathbf{u}$, è allora definita da

$$\delta F(\mathbf{u}, \delta\mathbf{u}) = \frac{\mathrm{d}}{\mathrm{d}\alpha} F(\mathbf{u} + \alpha\delta\mathbf{u})\Big|_{\alpha=0}\ .$$

La variazione $\delta F$ è lineare[6] in $\delta\mathbf{u}$, come si vede eseguendo la derivata:

$$\delta F(\mathbf{u}, \delta\mathbf{u}) = \sum_{i=i}^{n} \frac{\partial F}{\partial u_i}(\mathbf{u})\,\delta u_i\ .$$

---

2) Una seconda maniera di definire la variazione di un funzionale, più soddisfacente per certi particolari aspetti del calcolo delle variazioni, ma non necessaria ai nostri scopi, è quella che fa riferimento alla cosiddetta *derivata di Fréchet*; il metodo qui seguito è anche detto metodo della *derivata di Gateaux.*

3) Le variabili indipendenti sono qui indicate con $u_i$ anzichè con $x_i$, perchè risulti più immediata l'analogia con i funzionali: formalmente si passa da un caso all'altro sostituendo l'$n$-pla $u_i$, $i = 1, \ldots, n$, con la funzione $u(x)$, $x \in [a, b]$, ovvero sostituendo l'indice discreto $i$ con la variabile continua $x$.

4) L'$n$–pla $\delta\mathbf{u}$ si può riguardare come vettore tangente a $U$ in $\mathbf{u}$.

5) Si osservi che per $\alpha$ piccolo il punto variato $\mathbf{u} + \alpha\delta\mathbf{u}$ appartiene certamente a $U$.

6) L'applicazione lineare che ad ogni vettore tangente $\delta\mathbf{u}$ associa $\delta F$ (che a sua volta si può riguardare come vettore tangente a $\mathbb{R}$ in $F(\mathbf{u})$), altro non è che il differenziale $\mathrm{d}F$ in $\mathbf{u}$.

Facendo riferimento alla derivata direzionale, possiamo dire che $F$ è stazionaria in $\mathbf{u}$, se e solo se $\delta F$ si annulla in $\mathbf{u}$ per ogni scelta della variazione $\delta\mathbf{u}$.

Consideriamo ora un funzionale $F$; sia $U$ il suo insieme di definizione — pensiamo, per fissare le idee, all'insieme delle funzioni regolari dell'intervallo $[a, b]$ — e sia $u$ un "punto" interno a $U$. Fissata arbitrariamente una *variazione* $\delta u$ — nell'esempio, un'arbitraria funzione regolare su $[a, b]$ — consideriamo la famiglia a un parametro di funzioni variate[7)]

$$u^\alpha(x) = u(x) + \alpha\delta u(x) \ , \tag{4.3}$$

con $\alpha$ reale in un intorno dell'origine. In analogia con il caso finito-dimensionale, la *variazione* $\delta F$ *del funzionale* $F$ in $u$, relativa alla variazione $\delta u$, si definisce considerando $F[u^\alpha] = F[u+\alpha\delta u]$, che per $u$ e $\delta u$ fissati è una funzione ordinaria della variabile reale $\alpha$, e ponendo

$$\delta F[u, \delta u] \ = \ \frac{\mathrm{d}}{\mathrm{d}\alpha} F[u + \alpha\delta u]\Big|_{\alpha=0} \ ; \tag{4.4}$$

a giustificazione della notazione impiegata, si osservi che $\delta F$ è un numero per ogni scelta della funzione $u$ e della variazione $\delta u$, cioè è un funzionale dei due argomenti $u$ e $\delta u$. Il funzionale stesso si dirà differenziabile[8)] in $u$, se la derivata esiste per ogni scelta della variazione $\delta u$.

- **Osservazione.** Procedendo in modo più intuitivo, si potrebbe fare a meno del parametro $\alpha$, e pensare alla variazione $\delta u$ come "piccola". E' allora spontaneo definire la variazione $\delta F$ del funzionale, corrispondente alla variazione $\delta u$ di $u$, come la parte lineare in $\delta u$ dell'incremento $\Delta F = F[u + \delta u] - F[u]$; si pensi a uno sviluppo di Taylor arrestato al termine lineare, nel quale si trascurano termini di ordine superiore al primo in $\delta u$ (o nelle sue derivate, che come si vedrà compaiono naturalmente nel calcolo di $\delta F$).

  La derivata direzionale — si rifletta sul caso finito dimensionale — è proprio una procedura per definire in modo preciso la separazione della parte lineare dai "termini di ordine superiore" (che contengono $\alpha$ a un ordine superiore al primo, e dunque si annullano quando, dopo aver eseguito la derivata, si pone $\alpha = 0$). E' un utile esercizio vedere, in tutti gli esempi che seguono, che il procedimento intuitivo e la definizione precisa conducono al medesimo risultato.

Qualche esempio servirà a chiarire la nozione di variazione di un funzionale. Se $F[u] = \int_a^b u(x)\,\mathrm{d}x$, allora si ha $F[u^\alpha] = \int_a^b (u(x) + \alpha\delta u(x))\,\mathrm{d}x$, e dunque $\delta F = \int_a^b \delta u(x)\,\mathrm{d}x$ (poiché $F$ è lineare, si ha che $\delta F$ dipende solo dalla variazione $\delta u$ e non da $u$). Per

---

7) Si ammette qui implicitamente che l'insieme di definizione del funzionale abbia struttura di spazio vettoriale lineare; per una generalizzazione si veda piu' avanti.

8) Più precisamente, differenziabile secondo Gateaux. Si osservi che si ha $F[u + \alpha\delta u] - F[u] = \alpha\delta F[u, \delta u] + o(\alpha)$, come avviene nel caso finito dimensionale. In analogia al caso finito dimensionale, l'applicazione $\delta u \mapsto \delta F$, lineare, si dice differenziale del funzionale $F$ in $u$.

$F[u] = \int_a^b u^2(x)\,\mathrm{d}x$ risulta invece

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\int_a^b [u(x) + \alpha\delta u(x)]^2\,\mathrm{d}x\,\Big|_{\alpha=0} = \int_a^b 2[u(x) + \alpha\delta u(x)]\delta u(x)\,\mathrm{d}x\,\Big|_{\alpha=0}\ ,$$

e dunque $\delta F = 2\int_a^b u(x)\delta u(x)\,\mathrm{d}x$ (ora $\delta F$ dipende anche dal "punto" $u$ in cui la variazione è calcolata).

Consideriamo poi un funzionale come la lunghezza di una curva, in cui interviene esplicitamente la derivata $u'$ di $u$: $F[u] = \int_a^b [1 + u'^2(x)]^{1/2}\,\mathrm{d}x$. E' questo un caso particolare di un problema più generale alquanto significativo, ovvero quello di un funzionale della forma

$$F[u] = \int_a^b L(u(x), u'(x), x)\,\mathrm{d}x\ , \tag{4.5}$$

ove $L$ è una funzione (regolare) assegnata delle tre variabili reali $u$, $u'$ e $x$ (nell'esempio era $L = (1 + u'^2)^{1/2}$).

Il calcolo della variazione di questo funzionale non è difficile: tenendo presente che si ha $(u^\alpha)' = u' + \alpha(\delta u)'$, e portando la derivata rispetto a $\alpha$ sotto il segno di integrale, si trova

$$\begin{aligned}\delta F[u, \delta u] &= \frac{\mathrm{d}}{\mathrm{d}\alpha}\int_a^b L\big(u(x) + \alpha\delta u(x), u'(x) + \alpha(\delta u)'(x)\,\mathrm{d}x\,\Big|_{\alpha=0} \\ &= \int_a^b \Big[\frac{\partial L}{\partial u}(u(x), u'(x), x)\,\delta u(x) + \frac{\partial L}{\partial u'}(u(x), u'(x), x)\,(\delta u)'(x)\Big]\,\mathrm{d}x\ .\end{aligned} \tag{4.6}$$

Con una integrazione per parti si può poi eliminare $(\delta u)'$, e concludere con la seguente

**Proposizione 4.1:** *La variazione del funzionale $F[u] = \int_a^b L(u(x), u'(x), x)\,\mathrm{d}x$, corrispondente alla variazione $\delta u$, è data da*

$$\delta F = \frac{\partial L}{\partial u'}\delta u\Big]_a^b - \int_a^b \left(\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial u'} - \frac{\partial L}{\partial u}\right)\delta u\,\mathrm{d}x\ . \tag{4.7}$$

In particolare, se ci si restringe a variazioni $\delta u(x)$ nulle agli estremi, si trova

$$\delta F = -\int_a^b \left(\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial u'} - \frac{\partial L}{\partial u}\right)\delta u(x)\,\mathrm{d}x\ . \tag{4.8}$$

- **Esercizio 4.2:** Scrivere la variazione del funzionale esprimente la lunghezza di una curva del piano in coordinate cartesiane, $F[u] = \int_a^b [1 + u'^2(x)]^{1/2}\,\mathrm{d}x$.

### *4.1.3 Stazionarietà di un funzionale ed equazioni di Eulero-Lagrange*

In analogia con il caso delle ordinarie funzioni di più variabili, diciamo che il funzionale $F$ definito in $U$ è *stazionario* in $u \in U$, o che $u \in U$ è punto di stazionarietà per $F$, se risulta $\delta F[u, \delta u] = 0$ comunque si prenda la variazione $\delta u$ (con l'eventuale restrizione che $\delta u$ si annulli agli estremi). Si ha allora la seguente

**Proposizione 4.2:** *Sia dato un funzionale della forma (4.5). Condizione necessaria e sufficiente perché la funzione $u$ sia punto di stazionarietà di $F$, per variazioni che rispettino gli estremi, è che la funzione $u$ soddisfi l'equazione differenziale*

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial u'} - \frac{\partial L}{\partial u} = 0 \ .$$

Tale equazione è detta *equazione di Eulero–Lagrange* associata al funzionale $F$.

**Dimostrazione.** In virtù dell'espressione (4.8), si vede che $\delta F$ si annulla se l'equazione di Eulero–Lagrange è soddisfatta. Viceversa, se l'equazione non fosse soddisfatta in un qualsiasi punto $\bar{x} \in (a, b)$, allora, per continuità, esisterebbe un intorno $I$ di $\bar{x}$ in cui l'espressione $\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial u'} - \frac{\partial L}{\partial u}$ sarebbe di segno costante; scegliendo dunque una variazione $\delta u$ particolare, con $\delta u(x) = 0$ per $x \notin I$, e $\delta u(x) > 0$ in $I$ (non c'è difficoltà a trovare $\delta u$ regolare), risulterebbe in corrispondenza $\delta F \neq 0$, contro l'ipotesi $\delta F = 0$ per ogni $\delta u$.[9] Per continuità l'espressione si annulla anche agli estremi $a$, $b$.
*Q.E.D.*

La stazionarietà di $F = \int_a^b L(u(x), u'(x), x)\,\mathrm{d}x$ per variazioni $\delta u$ arbitrarie, purchè nulle agli estremi, si vede così essere equivalente a un'equazione differenziale per il "punto" di stazionarietà $u$, e precisamente all'equazione di Eulero–Lagrange relativa alla funzione $L(u, u', x)$, con assegnati dati agli estremi $u(a)$, $u(b)$.[10]

Ad esempio, se il funzionale $F$ rappresenta la lunghezza di una curva nel piano, allora si ha $L = (1 + u'^2)^{1/2}$, e l'equazione di Lagrange si scrive

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{u'}{\sqrt{1 + u'^2}} = 0 \ , \tag{4.9}$$

ovvero, con immediati calcoli, $u''(1 + u'^2)^{-3/2} = 0$, che risulta equivalente a $u'' = 0$, ed è risolta dalle rette.[11] Il risultato (certo non sorprendente!) è che le rette sono, tra tutte le curve del piano, quelle per cui il funzionale lunghezza è stazionario (si potrebbe vedere che è non solo stazionario, ma anche minimo): le rette sono le geodetiche del piano.

---

9) In generale, se l'integrale $\int g(x)\delta u(x)\,\mathrm{d}x$, con $g$ continua, si annulla *per ogni scelta* della variazione $\delta u$ (eventualmente nulla agli estremi), allora si ha necessariamente $g(x) = 0$ per ogni $x$. Questa proprietà, elementare ma importante, è a volte chiamata *lemma fondamentale del calcolo delle variazioni.*

10) Ci si riporta in tal modo al cosiddetto *problema di Sturm–Liouville*, cui si è accennato nel primo capitolo; si ricordi che (diversamente dal problema di Cauchy) tale problema non ammette sempre soluzione, né in generale la soluzione è unica.

11) Più semplicemente, dalla (4.9) si ha $u'/\sqrt{1 + u'^2} = c$, per una opportuna costante $c$, ovvero $u'^2 = c^2/(1 - c^2)$. Dunque si ha $u' = \mathrm{cost}$, che è risolta dalle rette.

- **Esercizio 4.3:** Si mostri che le rette sono geodetiche del piano, usando le coordinate polari (*Risposta:* l'equazione di Eulero–Lagrange prende la forma $uu'' - 2u'^2 - u^2 = 0$, ed è risolta dalla funzione che esprime le rette in coordinate polari, $r = u(\vartheta) = a/\sin(\vartheta - \varphi)$, ove $a > 0$ è la distanza della retta dall'origine, e $\varphi$ la sua inclinazione rispetto all'asse polare).

- **Esercizio 4.4:** Si dimostri che le geodetiche del cilindro e del cono appaiono come rette nello sviluppo piano delle superfici. Si trovi la condizione sull'angolo di apertura del cono perché una geodetica possa avere punti doppi (cioè perchè un laccio, fermato in un sol punto, possa restare teso sulla superficie di un cono). *Suggerimento:* si pensi allo sviluppo piano del cono, ottenuto tagliando la superficie lungo la direttrice passante per il punto doppio.

- **Esercizio 4.5:** Si verifichi che gli archi di cerchio massimo sono geodetiche della superficie sferica (è sufficiente verificarlo per l'equatore, o per un meridiano).

- **Esercizio 4.6:** Si consideri l'espressione (4.2) del tempo di percorrenza della curva $y = u(x)$ da parte di un punto materiale che parte dall'origine con velocità nulla e arriva a un punto prefissato del piano $xy$ verticale, con asse $y$ verticale discendente, e si scriva l'equazione differenziale della curva che rende stazionario (in realtà minimo) $T$ (*Risposta:* $uu'' = -\frac{1}{2}(1 + u'^2)$). Si verifichi che tale curva (chiamata *brachistocrona*, cioè "dal tempo più breve") è un arco di cicloide[12)] (equazioni parametriche: $x = R(\varphi - \sin\varphi)$, $y = R(1 - \cos\varphi)$). Si verifichi infine che la quantità $E = u'\frac{\partial L}{\partial u'}$ è costante al variare di $x$.

- **Esercizio 4.7:** Fissati due punti $(x_0, y_0)$ e $(x_1, y_1)$ del piano $xy$, con $y_0, y_1 > 0$, si determini la curva $y = u(x)$ che li collega, e che genera, per rotazione attorno all'asse $x$, la superficie di area minima (*Risposta:* l'equazione della curva è $uu'' = 1 - u'^2$). Si verifichi che tale equazione è risolta dalla catenaria, $y = a\cosh(x - c)$, con $a$ e $c$ fissati dalle condizioni estreme. La catenaria è la curva assunta da una fune pesante sospesa; la forma corrisponde al minimo del funzionale energia potenziale gravitazionale (ma vi è un punto delicato, legato al vincolo di lunghezza costante).

### 4.1.4 Il principio di Hamilton

Proseguiamo da questo momento in poi con le notazioni tipiche della meccanica dei punti materiali, indicando con $t$ (anzichè $x$) la variabile indipendente, e con $q(t)$, $t_0 \le t \le t_1$ (anzichè $u(x)$) una generica funzione regolare di $t$, che pensiamo rappresenti un movimento a priori possibile di un qualsiasi sistema lagrangiano a un grado di libertà.

Sia $L(q, \dot q, t)$ la lagrangiana del sistema; indichiamo con $S[q]$ il funzionale[13)]

---

12) Un'altra interessante proprietà del moto di un punto su di una cicloide, detto *pendolo cicloidale,* è di essere, a differenza del comune pendolo circolare, *esattamente* isocrono: l'equazione del moto per la lunghezza d'arco $s$, misurata a partire dal punto di minima energia, risulta essere esattamente $\ddot s = -\omega^2 s$, $\omega^2 = g/(4R)$.

13) Attenzione: la variabile $q$ che compare ad argomento del funzionale non è la coordinata la-

$\int_{t_0}^{t_1} L(q(t), \dot{q}(t), t)\, \mathrm{d}t$, chiamato *integrale di Hamilton*, o anche *azione hamiltoniana.* Traducendo nel linguaggio della meccanica quanto già imparato, possiamo dire che un particolare movimento $q(t)$, $t_0 \le t \le t_1$, rende stazionario $S[q]$ per variazioni $\delta q(t)$ arbitrarie, purché nulle agli estremi, se e solo se $q(t)$ soddisfa l'equazione di Lagrange

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \ .$$

In altre parole: tra tutti i moti a priori possibili che nell'intervallo di tempo $[t_0, t_1]$ portano da $q_0$ a $q_1$ fissati — cioè tra tutte le funzioni regolari che si possono immaginare, con $q(t_0) = q_0$ e $q(t_1) = q_1$ assegnati — il moto effettivo, o naturale, è quello che rende stazionario il funzionale $S[q]$ per variazioni arbitrarie $\delta q(t)$, $t_0 \le t \le t_1$, purché nulle agli estremi.

La generalizzazione di quanto esposto sopra al caso di un sistema a più gradi di libertà è immediata. Se $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ è la lagrangiana, dove $\mathbf{q} = (q_1, \ldots, q_n)$, si introduce il funzionale $S$, detto funzionale di *azione hamiltoniana*, definito per ogni movimento $\mathbf{q}(t)$, $t_0 \le t \le t_1$, da

$$S[\mathbf{q}] = \int_{t_0}^{t_1} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)\, \mathrm{d}t \ . \tag{4.10}$$

E' allora un semplice esercizio verificare la seguente

**Proposizione 4.3:** *La variazione del funzionale di azione hamiltoniana $S$ è data da*

$$\delta S = \Big[ \sum_{h=1}^{n} \frac{\partial L}{\partial \dot{q}_h} \delta q_h \Big]_{t_0}^{t_1} - \int_{t_0}^{t_1} \Big[ \sum_{h=1}^{n} \Big( \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_h} - \frac{\partial L}{\partial q_h} \Big) \delta q_h \Big]\, \mathrm{d}t \ ; \tag{4.11}$$

*in particolare, per variazioni che rispettino gli estremi, si ha*

$$\delta S = -\int_{t_0}^{t_1} \Big[ \sum_{h=1}^{n} \Big( \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_h} - \frac{\partial L}{\partial q_h} \Big) \delta q_h \Big]\, \mathrm{d}t \ . \tag{4.12}$$

Si ottiene allora immediatamente la

**Proposizione 4.4 (Principio di Hamilton):** *Il moto $\mathbf{q}(t)$, $t_0 \le t \le t_1$, rende stazionario il funzionale di azione $S$, per variazioni $\delta q_1(t), \ldots, \delta q_n(t)$ arbitrarie, nulle agli estremi, se e solo se è un moto naturale, ovvero soddisfa le equazioni di Lagrange*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_h} - \frac{\partial L}{\partial q_h} = 0 \ , \qquad h = 1, \ldots, n \ .$$

Il principio di Hamilton rappresenta dunque la formulazione variazionale delle equazioni di Lagrange.

---

grangiana $q$, ma la funzione (il movimento) $q(t)$, $t_0 \le t \le t_1$; la notazione un po' ambigua non deve ingannare.

- **Osservazione.** Dal principio di Hamilton si deducono immediatamente le due proprietà di invarianza delle equazioni di Lagrange, verificate direttamente nel capitolo precedente. La prima è la proprietà di invarianza per cambiamenti qualsiasi di coordinate: infatti, dato un cambiamento di coordinate (un diffeomorfismo) $\mathbf{q} = \mathbf{q}(\tilde{\mathbf{q}}, t)$, si definisce $\tilde{L}$ per sostituzione di variabili, ovvero
$$\tilde{L}(\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}, t) = L(\mathbf{q}(\tilde{\mathbf{q}}, t), \dot{\mathbf{q}}(\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}, t), t) \ ;$$
allora per movimenti $\mathbf{q}(t)$, $\tilde{\mathbf{q}}(t)$ corrispondentisi risulta evidentemente
$$\int_{t_0}^{t_1} \tilde{L}(\tilde{\mathbf{q}}(t), \dot{\tilde{\mathbf{q}}}(t), t)\, \mathrm{d}t = \int_{t_0}^{t_1} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)\, \mathrm{d}t \ ,$$
e il primo integrale è stazionario se e solo se lo è l'altro. Poichè, come si è visto, la stazionarietà degli integrali è poi equivalente alle rispettive equazioni di Lagrange, si conclude che le equazioni di Lagrange nelle nuove variabili sono soddisfatte se e solo se esse sono soddisfatte nelle vecchie variabili.
Altrettanto immediata è la verifica della seconda proprietà, cioè che due lagrangiane $L$ e $L'$ danno luogo alle medesime equazioni di Lagrange se differiscono per un termine additivo $L_0(\mathbf{q}, \dot{\mathbf{q}}, t)$, che sia la derivata totale rispetto al tempo di una funzione arbitraria $F(\mathbf{q}, t)$: infatti, posto $S = \int_{t_0}^{t_1} L\, \mathrm{d}t$, $S' = \int_{t_0}^{t_1} (L + \frac{\mathrm{d}}{\mathrm{d}t} F)\, \mathrm{d}t$, si ha $S' = S + F\big|_{t_0}^{t_1}$, e dunque, per estremi fissati, $\delta S' = \delta S$; la conclusione è immediata.

### 4.1.5 *Una generalizzazione*

Per definire la variazione di un funzionale $F[u]$, corrispondente alla variazione $\delta u$ della funzione $u$, abbiamo fatto riferimento alla particolare famiglia a un parametro di funzioni variate $u^\alpha = u + \alpha \delta u$, ponendo per definizione
$$\delta F = \frac{\mathrm{d}}{\mathrm{d}\alpha} F[u^\alpha] \Big|_{\alpha=0} \ . \tag{4.13}$$
E' abbastanza spontaneo (e utile) allargare leggermente la nozione di variazione, facendo riferimento a famiglie qualsiasi di funzioni variate, non necessariamente lineari in $\alpha$ come nella (4.3). Precisamente, diremo che una funzione di due variabili reali $u^\alpha(x)$ costituisce una famiglia di variazioni della funzione $u(x)$, se $u^\alpha(x)$ è definita e regolare per $\alpha$ in un intorno dell'origine, e coincide con $u(x)$ per $\alpha = 0$, ovvero si ha $u^0(x) = u(x)$. La definizione data sopra di variazione del funzionale $F$ ha senso evidentemente anche con riferimento a tali più generali famiglie, e non è difficile convincersi che se si denota (coerentemente con la nozione precedente)
$$\delta u(x) = \frac{\mathrm{d}}{\mathrm{d}\alpha} u^\alpha(x) \Big|_{\alpha=0} \ , \tag{4.14}$$
allora la variazione $\delta F$ definita dalla (4.13) dipende solo dalla funzione di riferimento $u$ e da $\delta u$ (ovvero: famiglie diverse di variazioni di $u$, con il medesimo $\delta u$, danno il

medesimo $\delta F$). Possiamo allora dire, coerentemente con la nozione precedente, che il funzionale $F$ è stazionario in $u$, se la variazione $\delta F$ definita dalla (4.13) si annulla in corrispondenza ad ogni famiglia $u^\alpha$ di variazioni di $u$ (con l'eventuale restrizione che le variazioni rispettino gli estremi, ovvero che $u^\alpha(a)$, $u^\alpha(b)$ non dipendano da $\alpha$).

Anche nel linguaggio della meccanica dei punti materiali possiamo evidentemente far riferimento a famiglie qualsiasi $\mathbf{q}^\alpha(t)$ di moti variati rispetto a un moto assegnato $\mathbf{q}(t)$, non necessariamente della forma $\mathbf{q}(t) + \alpha\delta\mathbf{q}(t)$; per ognuna di tali famiglie la variazione del funzionale di azione $S$ definito dalla (4.10), che per definizione è data da $\delta S = \frac{\mathrm{d}}{\mathrm{d}\alpha} S[\mathbf{q}^\alpha]\big|_{\alpha=0}$, risulta ancora espressa dalla (4.11), purchè si ponga, coerentemente con la (4.14),

$$\delta\mathbf{q}(t) = \frac{\mathrm{d}}{\mathrm{d}\alpha}\mathbf{q}^\alpha(t)\bigg|_{\alpha=0} . \tag{4.15}$$

- **Osservazione.** Questa estensione della nozione di variazione a famiglie qualsiasi di funzioni variate, non necessariamente lineari in $\alpha$, ha una ovvia corrispondenza nel caso finito dimensionale: data $F : \mathbb{R}^n \to \mathbb{R}$, e fissato un punto $\mathbf{u} = (u_1, \ldots, u_n)$ nel dominio di definizione (corrispondente alla funzione di riferimento $u$), possiamo considerare una curva qualsiasi $\mathbf{u}^\alpha$ passante per $\mathbf{u}$, rappresentata parametricamente dall' $n$–pla di funzioni regolari $u_i^\alpha$, $i = 1, \ldots, n$ (corrispondente alla famiglia $u^\alpha(x)$, $a \le x \le b$); possiamo poi dire che la funzione $F$ è stazionaria in $\mathbf{u}$, se si ha $\frac{\mathrm{d}}{\mathrm{d}\alpha} F(\mathbf{u}^\alpha)\big|_{\alpha=0} = 0$ per ogni scelta della curva $\mathbf{u}^\alpha$. In questa analogia, la famiglia $u^\alpha$ appare come una curva nello spazio funzionale $U$ considerato. Grazie a questa estensione, per le funzioni ordinarie come per i funzionali cade la necessità che l'insieme di definizione $U$ abbia la sruttura di spazio vettoriale lineare (si osservi che, per funzioni ordinarie definite su di una varietà $U$, la variazione $\delta\mathbf{u} = \frac{\mathrm{d}}{\mathrm{d}\alpha}\mathbf{u}^\alpha\big|_{\alpha=0}$, analoga alla variazione $\delta u$, è un vettore dello spazio tangente a $U$ in $\mathbf{u}$).

### *4.1.6 Il principio di Maupertuis (o di minima azione).*

Il principio di Hamilton studiato sopra fa riferimento al funzionale di azione $S$, definito nello spazio funzionale dei *movimenti* $\mathbf{q}(t)$, $t_0 \le t \le t_1$, che si svolgono tra due configurazioni assegnate $\mathbf{q}_0$ e $\mathbf{q}_1$. Storicamente era già stato considerato fin dal 1744, da Maupertuis, un principio variazionale che fa invece riferimento a un diverso funzionale, che chiameremo *azione ridotta* o *azione di Maupertuis,* e denoteremo con $A$; questo funzionale è definito nello spazio delle *traiettorie* con estremi assegnati, intese come curve geometriche nello spazio delle configurazioni, senza riferimento esplicito alla legge oraria di percorrenza; al posto della legge oraria viene invece assegnato, come vedremo, il valore dell'energia.[14)]

---

[14)] Questo ambito risulta più naturale quando si voglia considerare l'analogia tra la meccanica e l'ottica geometrica, le cui leggi sono condensate nel principio di Fermat, che fa appunto riferimento a traiettorie. Un'analoga visione geometrica, per traiettorie del cosiddetto spazio–tempo, è fondamentale nella teoria della relatività.

Restringiamoci allora al caso di un sistema lagrangiano naturale, con lagrangiana

$$L(\mathbf{q}, \dot{\mathbf{q}}) = T(\mathbf{q}, \dot{\mathbf{q}}) - V(\mathbf{q})$$

indipendente dal tempo, e dunque conservativo (cioè tale che l'energia $E = T + V$ sia una costante del moto). Fissato un valore di $E$, consideriamo l'insieme delle traiettorie orientate $\gamma$ tra due estremi fissati $\mathbf{q}_0$ e $\mathbf{q}_1$, compatibili con la conservazione dell'energia, più precisamente tali che sia sempre $T = E - V(\mathbf{q}) > 0$.[15] La scelta della curva $\gamma$ e del valore di $E$ determinano, punto per punto, $\dot{\mathbf{q}}$ come funsione di $\mathbf{q}$: infatti, $\gamma$ determina la direzione del vettore $\dot{\mathbf{q}}$, mentre a sua volta la scelta di $E$ determina in ciascun punto l'energia cinetica $T = E - V(\mathbf{q})$, e dunque (per fissata direzione) il modulo di $\dot{\mathbf{q}}$. Anche i momenti $p_h = \frac{\partial T}{\partial \dot{q}_h} = \sum_k a_{hk}(\mathbf{q})\dot{q}_k$ risultano allora determinati come funzioni di $\mathbf{q}$, sicché ha senso considerare il funzionale $A[\gamma; E]$, definito per ogni curva $\gamma$, e ogni fissato valore di $E$, dall'integrale curvilineo

$$A[\gamma; E] = \int_\gamma \mathbf{p}(\mathbf{q}) \cdot \mathrm{d}\mathbf{q} = \int_\gamma \sum_{h=1}^{n} p_h(\mathbf{q})\, \mathrm{d}q_h \ ; \tag{4.16}$$

a questo funzionale si dà il nome di *azione ridotta*.

Osserviamo anche che, essendo fissata in ogni punto di $\gamma$ la velocità $\dot{\mathbf{q}}$, per ciascuna traiettoria resta determinata (a meno di una arbitraria scelta dell'istante iniziale) anche la legge oraria di percorrenza;[16] pertanto, fissata $E$, ad ogni traiettoria $\gamma$ è associato un moto $\mathbf{q}(t)$, per $t$ in un intervallo $[t_0, t_1]$, con $t_0$ fissato ad arbitrio, e $t_1$ dipendente esso stesso da $\gamma$.

- **Osservazione.** Questo modo di identificare un movimento $q(t)$ mediante una traiettoria orientata e un valore dell'energia, con la determinazione conseguente della legge oraria, si può considerare complementare a quello utilizzato nel principio di Hamilton, per il quale si assegna direttamente il movimento $q(t)$, ovvero una traiettoria e una legge oraria (in particolare, l'intervallo $[t_0, t_1]$ nel quale si svolge il moto). Scelte indipendenti dell'energia $E$ e del tempo di percorrenza non sono in generale compatibili.

Il riferimento al moto $\mathbf{q}(t)$ associato a $\gamma$ per un assegnato valore di $E$ è utile in particolare per confrontare l'azione di Maupertuis $A$ ora definita, con l'azione di Hamilton $S[\mathbf{q}] = \int_{t_0}^{t_1} L(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, \mathrm{d}t$. Se infatti parametrizziamo la curva $\gamma$ con il tempo del corrispondente moto $\mathbf{q}(t)$, $t_0 \le t \le t_1$, allora il funzionale $A$ prende la forma

$$A[\gamma; E] = \int_{t_0}^{t_1} \mathbf{p}(\mathbf{q}(t)) \cdot \dot{\mathbf{q}}(t)\ \mathrm{d}t \ ;$$

pertanto (ricordando che $E = \mathbf{p} \cdot \dot{\mathbf{q}} - L$) si ottiene $A[\gamma; E] = \int_{t_0}^{t_1} (L + E)\, \mathrm{d}t$, e quindi

$$A[\gamma; E] = S[\mathbf{q}] + E(t_1 - t_0) \ . \tag{4.17}$$

---

15) La disuguaglianza stretta esclude che vi siano sulla curva punti di arresto.

16) E' evidente l'analogia con quanto si è visto nel primo capitolo per il caso a un solo grado di libertà, quando si è osservato che la conservazione dell'energia determina la legge oraria.

Consideriamo ora una curva $\gamma$ tra due estremi $\mathbf{q}_0$ e $\mathbf{q}_1$ assegnati, e una famiglia a un parametro $\gamma^\alpha$ di curve variate che rispettino tali estremi; fissiamo lo stesso valore $E$ per tutte le curve della famiglia, e poniamo per definizione

$$\delta A = \frac{\mathrm{d}}{\mathrm{d}\alpha} A[\gamma^\alpha; E]\bigg|_{\alpha=0} .$$

A ogni curva $\gamma^\alpha$ è associato, come abbiamo visto, un moto $\mathbf{q}^\alpha(t)$, $t_0^\alpha \le t \le t_1^\alpha$ (si osservi che gli estremi dell'intervallo di definizione dipendono da $\alpha$, e che si ha $t_0^0 = t_0$, $t_1^0 = t_1$); alla famiglia di moti è poi associata, secondo la definizione consueta, la variazione $\delta\mathbf{q}$, definita per ogni $t \in [t_0, t_1]$ dalla (4.15). Tale variazione però non si annulla agli istanti $t_0$ e $t_1$: infatti, poiché le curve della famiglia hanno estremi $\mathbf{q}_0$, $\mathbf{q}_1$ fissati, si ha $\mathbf{q}^\alpha(t_i^\alpha) = \mathbf{q}_i$ (indipendente da $\alpha$) per $i = 0, 1$; dunque, per derivazione, si ottiene subito[17] $\frac{\mathrm{d}}{\mathrm{d}\alpha}\mathbf{q}^\alpha(t_i^\alpha)\big|_{\alpha=0} = \delta\mathbf{q}(t_i) + \dot{\mathbf{q}}(t_i)\, \frac{\mathrm{d}}{\mathrm{d}\alpha} t_i^\alpha\big|_{\alpha=0} = 0$, ovvero

$$\delta\mathbf{q}(t_i) = -\dot{\mathbf{q}}(t_i)\delta t_i \ , \qquad i = 0, 1 \ , \tag{4.18}$$

dove si è definito, coerentemente con la notazione introdotta,

$$\delta t_i = \frac{\mathrm{d}}{\mathrm{d}\alpha} t_i^\alpha\bigg|_{\alpha=0} \ , \qquad i = 0, 1 \ .$$

La relazione (4.18) ci sarà utile tra breve.

Dopo queste premesse si dimostra senza difficoltà la seguente

**Proposizione 4.5:** *La variazione del funzionale di azione ridotta $A$, corrispondente a ad un fissato valore $E$ dell'energia, e a una fissata famiglia $\gamma^\alpha$ di curve variate tra le medesime configurazioni estreme, è data da*

$$\delta A = -\int_{t_0}^{t_1} \sum_{h=1}^{n} \Big(\dot p_h - \frac{\partial L}{\partial q_h}\Big)\delta q_h \,\mathrm{d}t \ , \tag{4.19}$$

*dove il fattore $\sum_h \big(\dot p_h - \frac{\partial L}{\partial q_h}\big)$ si intende valutato lungo il movimento $\mathbf{q}_h(t)$, $t_0 \le t \le t_1$, corrispondente alla curva di riferimento $\gamma = \gamma^0$.*

**Dimostrazione.** Dalla (4.17) si ha

$$\delta A = \delta S + E(\delta t_1 - \delta t_0) \ , \tag{4.20}$$

e resta da valutare $\delta S = \frac{\mathrm{d}}{\mathrm{d}\alpha}\int_{t_0^\alpha}^{t_1^\alpha} L(\mathbf{q}^\alpha(t), \dot{\mathbf{q}}^\alpha(t))\,\mathrm{d}t\Big|_{\alpha=0}$. A tal fine si osserva che ora, nel calcolo di $\delta S$, si trova un termine aggiuntivo rispetto alla (4.11), dovuto al fatto che questa volta anche gli estremi di integrazione nell'espressione di $S$ dipendono da $\alpha$; precisamente si ha

$$\delta S = \mathbf{p}\cdot\delta\mathbf{q}\Big|_{t_0}^{t_1} - \int_{t_0}^{t_1}\Big(\dot{\mathbf{p}} - \frac{\partial L}{\partial \mathbf{q}}\Big)\cdot\delta\mathbf{q}\,\mathrm{d}t + L(\mathbf{q}(t_1), \dot{\mathbf{q}}(t_1))\delta t_1 - L(\mathbf{q}(t_0), \dot{\mathbf{q}}(t_0))\delta t_0 \ .$$

---

[17] Per eseguire la derivata può risultare più chiara la notazione $\mathbf{q}(\alpha, t_i(\alpha))$ al posto di $\mathbf{q}^\alpha(t_i^\alpha)$

D'altra parte, dalla (4.18) segue immediatamente

$$\mathbf{p}(t_i)\cdot\delta\mathbf{q}(t_i)+L(\mathbf{q}(t_i),\dot{\mathbf{q}}(t_i))\delta t_i=-E\delta t_i\ ,\qquad i=0,1\ ;$$

per confronto con la (4.20) si ottiene allora la (4.19). *Q.E.D.*

Da questa proposizione segue senza difficoltà il seguente

**Corollario 4.1:** *Principio di Maupertuis* Per ogni fissato valore $E$ dell'energia, le traiettorie corrispondenti ai moti naturali (cioè soddisfacenti le equazioni di Lagrange) hanno la proprietà di rendere stazionario il funzionale di azione ridotta $A=\int_\gamma \mathbf{p}\cdot\mathrm{d}\mathbf{q}$ per variazioni arbitrarie della traiettoria $\gamma$, nulle agli estremi.

- **Osservazione.** Viceversa, si potrebbe vedere che se una curva $\gamma$ rende stazionario il funzionale di azione ridotta $A$, allora il corrispondente moto $\mathbf{q}(t)$ soddisfa le equazioni di Lagrange. Tale conclusione è in realtà un po' delicata, e non la si può trarre direttamente dalla (4.19), perché in essa le variazioni $\delta q_h$, essendo fissata l'energia del moto variato, non sono indipendenti; per lo stesso motivo non sono indipendenti le $n$ quantità da annullare (deve infatti risultare in ciascun punto della traiettoria $\sum_h(\dot p_h-\frac{\partial L}{\partial q_h})\dot q_h=\dot E=0$). Per una trattazione più completa, si veda ad esempio il trattato di Levi-Civita e Amaldi (ove le $\delta q_h$ sono rese indipendenti grazie all'introduzione delle cosiddette "variazioni asincrone").

Si noti come l'espressione (4.19) di $\delta A$ a fissata $E$ sia proprio uguale all'espressione (4.12) di $\delta S$ corrispondente a un fissato intervallo $[t_0,t_1]$ (avendo fissato in entrambi i casi gli estremi $\mathbf{q}_0$, $\mathbf{q}_1$); è proprio per questo motivo che sia il principio di Hamilton sia il principio di Maupertuis corrispondono alle medesime equazioni differenziali, cioè quelle di Lagrange.

### *4.1.7 Applicazione: dinamica e moti geodetici.*

Restringiamoci ora provvisoriamente al caso particolare del moto di un punto materiale su di una superficie liscia, soggetto a un sistema di forze conservative di energia potenziale $V(\mathbf{q})$. In questo caso, avendo fissato una traiettoria $\gamma$ e un valore $E$ dell'energia, possiamo scrivere $\mathbf{p}\cdot\mathrm{d}\mathbf{q}=\mathbf{p}\cdot\dot{\mathbf{q}}\,\mathrm{d}t=2T\,\mathrm{d}t=\sqrt{2T}\sqrt{2T}\,\mathrm{d}t$ $=\sqrt{2(E-V)}\sqrt{m}\,v\,\mathrm{d}t$, ove $v$ è il modulo della velocità; l'azione di Maupertuis prende allora la forma $A=\int_{t_0}^{t_1}\sqrt{2m(E-V)}\,v\,\mathrm{d}t$, o anche, ritornando a una formulazione puramente geometrica,

$$A[\gamma;E]=\int_\gamma\sqrt{2m(E-V(\mathbf{q}))}\ \mathrm{d}l\ ,\tag{4.21}$$

ove $\mathrm{d}l=v\,\mathrm{d}t$ è la lunghezza euclidea dell'arco di curva percorso nel tempo $\mathrm{d}t$. Il funzionale che il moto naturale rende stazionario appare così come una sorta di lunghezza della traiettoria, ove però ogni elemento $\mathrm{d}l$ di linea è pesato con il fattore $\sqrt{2m(E-V(\mathbf{q}))}$; tale fattore è una ben definita funzione del posto, dipendente parametricamente dall'energia $E$. Nel caso speciale del moto libero, cioè con $V$ costante sulla superficie, il funzionale rappresenta proprio (a meno dell'irrilevante costante moltiplicativa $\sqrt{2mE}$) la lunghezza euclidea della traiettoria. Si conclude in particolare che per un punto materiale vincolato a una superficie liscia e fissa, in assenza di forze esterne, le traiettorie coincidono con le geodetiche della superficie. Si osservi che il valore $E$ dell'energia non ha alcun rilievo nel determinare la traiettoria $\gamma$ (mentre è rilevante, evidentemente, per determinare la legge oraria del moto su $\gamma$).

Più in generale, per un qualunque sistema a $n$ gradi di libertà conservativo, è spontaneo definire l'elemento di linea $\mathrm{d}s$ sulla varietà vincolare mediante la relazione

$$\mathrm{d}s^2 = \sum_{hk} a_{hk}(\mathbf{q})\,\mathrm{d}q_h\,\mathrm{d}q_k \tag{4.22}$$

( $\mathrm{d}s = \sqrt{m}\,\mathrm{d}l$ per un singolo punto materiale); per ogni moto si ha allora $\mathrm{d}s = \sqrt{2T}\,\mathrm{d}t$, mentre l'azione di Maupertuis prende la forma

$$A = \int_\gamma \sqrt{2(E - V(\mathbf{q}))}\;\mathrm{d}s\ .$$

Si osservi che la metrica così introdotta tiene conto della struttura cinematica del sistema considerato (eventuali vincoli, scelta delle coordinate libere), riassunta nell'espressione della sua matrice cinetica. Anche in questo caso possiamo concludere che il moto, in assenza di forze attive, segue le geodetiche della varietà vincolare, dotata della metrica (4.22) (e il valore di $E$ è ancora irrilevante ai fini della determinazione della traiettoria).

Una ulteriore interessante variante formale del principio di Maupertuis si ottiene riassorbendo il fattore $\sqrt{(E - V(\mathbf{q}))}$ nella definizione della lunghezza d'arco, ovvero introducendo per la varietà vincolare la nuova metrica

$$\mathrm{d}s^2 = 2\big(E - V(\mathbf{q})\big)\sum_{hk} a_{hk}(\mathbf{q})\,\mathrm{d}q_h\,\mathrm{d}q_k\ ;$$

infatti, in tal modo l'azione di Maupertuis si scrive

$$A = \int_\gamma \mathrm{d}s\ ,$$

ovvero viene a coincidere con la lunghezza della traiettoria. In base al principio di Maupertuis si vede allora che *i moti naturali si possono sempre pensare come moti geodetici, con una opportuna metrica che tiene conto, oltre che della struttura cinematica del sistema, anche del complesso delle forze presenti, e dell'energia assegnata.* Così, nel caso ad esempio di un punto materiale nell'ordinario spazio tridimensionale, la presenza di forze può essere equivalentemente attribuita a un "incurvamento" dello spazio (ovvero alla presenza di una metrica diversa da quella euclidea, come avviene sulle varietà curve); analogamente (in modo forse più facilmente intuibile) il problema di un punto materiale vincolato a un piano, soggetto a un sistema di forze conservative, è equivalente al problema di un punto libero su una opportuna superficie curva. Si giunge in tal modo a una completa geometrizzazione della dinamica, secondo un procedimento che è stato poi portato a compimento nella teoria della relatività generale.

### 4.1.8 Principio di Hamilton ed equazioni canoniche (estensione di Helmholtz)

Non è difficile vedere che anche le equazioni canoniche

$$\dot p_h = -\frac{\partial H}{\partial q_h}\ , \qquad \dot q_h = \frac{\partial H}{\partial p_h}\ , \qquad h = 1,\ldots,n\ ,$$

ammettono una formulazione variazionale.

Sia $p_h(t)$, $q_h(t)$, $h = 1,\ldots,n$, un moto a priori possibile tra gli istanti $t_0$ e $t_1$ (vale la pena di sottolineare che le $n$ funzioni $p_1(t),\ldots,p_n(t)$ sono pensate, in questo contesto hamiltoniano, indipendenti dalle $q_1(t),\ldots,q_n(t)$: un moto a priori possibile è così caratterizzato da $2n$, anziché da $n$, funzioni); costruiamo per esso il funzionale, che come è tradizione continuiamo a chiamare azione hamiltoniana e ad indicare con $S$,

$$S[\mathbf{p},\mathbf{q}] \;=\; \int_{t_0}^{t_1} \Big[\sum_{h=1}^n p_h(t)\dot q_h(t) - H\big(\mathbf{p}(t),\mathbf{q}(t),t\big)\Big]\,\mathrm{d}t \tag{4.23}$$

(si osservi che $S$ coincide con l'azione hamiltoniana precedentemente introdotta, qualora le funzioni $p_h(t)$ si pensino non arbitrarie, ma dipendenti dalle $q_h(t)$ tramite la consueta relazione $p_h = \frac{\partial L}{\partial \dot q_h}$). In corrispondenza alle arbitrarie variazioni $\delta p_h(t)$, $\delta q_h(t)$, $h = 1,\ldots,n$, il funzionale $S$ subisce la variazione

$$\delta S \;=\; \int_{t_0}^{t_1} \Big[\sum_{h=1}^n \delta p_h(t)\dot q_h(t) + p_h(t)\delta\dot q_h(t) - \frac{\partial H}{\partial p_h}\delta p_h(t) - \frac{\partial H}{\partial q_h}\delta q_h(t)\Big]\ \mathrm{d}t\ ,$$

con $\delta\dot q_h(t) = \frac{\mathrm{d}}{\mathrm{d}t}\delta q_h(t)$; eseguendo allora, come di consueto, un'integrazione per parti, si ottiene

$$\delta S \;=\; \int_{t_0}^{t_1} \sum_{h=1}^n \Big[\big(\dot q_h - \frac{\partial H}{\partial p_h}\big)\delta p_h - \big(\dot p_h + \frac{\partial H}{\partial q_h}\big)\delta q_h\Big]\ \mathrm{d}t + \sum_{h=1}^n p_h\delta q_h\Big]_{t_0}^{t_1}\ . \tag{4.24}$$

Si deduce allora la

**Proposizione 4.6:** *Il moto $\mathbf{p}(t),\mathbf{q}(t)$, $t_0 \le t \le t_1$, rende stazionario il funzionale di azione $S$ (4.23) corrispondente a una assegnata hamiltoniana $H$, per variazioni $\delta\mathbf{p}$, $\delta\mathbf{q}$ arbitrarie nulle agli estremi, se e solo se esso soddisfa le equazioni di Hamilton relative all'hamiltoniana $H$.*

- **Osservazione.** Come si vede dalla (4.24), se sono soddisfatte le equazioni di Hamilton, allora $\delta S$ si annulla anche per la classe un po' più vasta di variazioni per le quali $\delta\mathbf{q}$ è nullo agli estremi, mentre $\delta\mathbf{p}$ è arbitrario.

Vale la pena di osservare che $H$ è qui del tutto arbitraria: ad esempio, non è necessariamente di secondo ordine nelle $p_h$, e non è necessariamente ottenuta da una lagrangiana per mezzo di una trasformazione di Legendre. Questa estensione del principio di Hamilton alle equazioni canoniche è dovuta ad Helmholtz.

## 4.2 Le trasformazioni canoniche

### 4.2.1 Nozione ed esempi

Uno degli aspetti più interessanti del formalismo lagrangiano è la sua invarianza per trasformazioni di coordinate: come si è visto, se $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, $\mathbf{q} = (q_1, \ldots, q_n)$, è la lagrangiana di un qualsiasi sistema dinamico, e introduciamo la trasformazione di coordinate $q_h = q_h(\tilde{q}_1, \ldots, \tilde{q}_n, t)$, $h = 1, \ldots, n$ (ponendo corrispondentemente $\dot{q}_h = \sum_k \frac{\partial q_h}{\partial \tilde{q}_k}\dot{\tilde{q}}_k + \frac{\partial q_h}{\partial t}$), allora le equazioni del moto nelle nuove coordinate $\tilde{q}_1, \ldots, \tilde{q}_n$ hanno ancora la forma di equazioni di Lagrange, e la nuova lagrangiana $\tilde{L}$ si ottiene da $L$ per semplice sostituzione di variabili, $\tilde{L}(\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}, t) = L(\mathbf{q}(\tilde{\mathbf{q}}, t), \dot{\mathbf{q}}(\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}, t), t)$. Vogliamo qui studiare il problema analogo per il formalismo hamiltoniano; come avremo modo di vedere, il formalismo hamiltoniano da questo punto di vista è sensibilmente più interessante e più ricco del formalismo lagrangiano, perchè la classe di trasformazioni di coordinate ammissibili in ambito hamiltoniano, chiamate *trasformazioni canoniche*, è sostanzialmente più vasta. Infatti, anzitutto, oltre alle trasformazioni in cui (come avviene in ambito lagrangiano) $q_1, \ldots, q_n$ dipendono dalle sole $\tilde{q}_1, \ldots, \tilde{q}_n$ e da $t$, e i momenti si trasformano di conseguenza (trasformazioni "puntuali", estese naturalmente ai momenti, cui si è già accennato nel capitolo precedente) si prendono in considerazione trasformazioni di coordinate più generali, del tipo

$$p_h = u_h(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t)\ , \qquad q_h = v_h(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t)\ , \qquad h = 1, \ldots, n\ ,$$

in cui cioè coordinate e momenti sono mescolati insieme dalla trasformazione.[18)] Useremo di frequente, per la trasformazione di coordinate e la sua inversa, la notazione più compatta

$$\begin{cases} \mathbf{p} = \mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \\ \mathbf{q} = \mathbf{v}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \end{cases}, \qquad \begin{cases} \tilde{\mathbf{p}} = \tilde{\mathbf{u}}(\mathbf{p}, \mathbf{q}, t) \\ \tilde{\mathbf{q}} = \tilde{\mathbf{v}}(\mathbf{p}, \mathbf{q}, t) \end{cases}, \tag{4.25}$$

o anche

$$(\mathbf{p}, \mathbf{q}) = \mathbf{w}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t)\ , \qquad (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) = \tilde{\mathbf{w}}(\mathbf{p}, \mathbf{q}, t)\ .$$

Queste trasformazioni di coordinate giocano un ruolo essenziale, in particolare, nella cosiddetta "teoria delle perturbazioni", uno dei rami ancor oggi più fecondi della meccanica, nel quale hanno consentito di compiere progressi fondamentali.

Tuttavia, non tutte le trasformazioni generali del tipo (4.25) conservano la forma delle equazioni di Hamilton, ovvero sono, come si suol dire, "canoniche". Quali siano, che proprietà abbiano e come si generino tali trasformazioni, è oggetto della *teoria delle trasformazioni canoniche*, che andiamo a esporre.

Qui e nel seguito ci occuperemo (senza ricordarlo di volta in volta) soltanto di trasformazioni definite localmente, cioè per $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$ in un dominio aperto $\tilde{U} \in \mathbb{R}^{2n}$, e $t$ in

---

[18)] E' inteso che, una volta assegnata la legge di trasformazione delle coordinate, la corrispondente legge di trasformazione per le loro derivate rispetto al tempo resta determinata di conseguenza: ad esempio si ha $\dot{p}_h = \sum_k (\frac{\partial u_h}{\partial \tilde{p}_k}\dot{\tilde{p}} + \frac{\partial u_h}{\partial \tilde{q}_k}\dot{\tilde{q}}) + \frac{\partial u_h}{\partial t}$, e similmente per $\dot{q}_h$.

un intervallo contenente l'origine; tali trasformazioni saranno sempre supposte invertibili in $\tilde{U}$. Evitando qualunque considerazione sui requisiti minimi di differenziabilità delle funzioni che definiscono la trasformazione, assumeremo che esse, ed ogni altra funzione di cui faremo uso, siano di volta in volta "regolari quanto basta" a rendere corrette le affermazioni che faremo.

Cominciamo col dare una definizione precisa, del tutto generale, di trasformazione canonica.

**Definizione 4.1:** *La trasformazione di coordinate (4.25), regolare e invertibile, si dice* canonica, *se muta qualunque sistema canonico in un sistema canonico, cioè se, comunque si prenda una hamiltoniana $H(\mathbf{p},\mathbf{q},t)$, esiste sempre una hamiltoniana $K(\tilde{\mathbf{p}},\tilde{\mathbf{q}},t)$, tale che le equazioni canoniche relative ad $H$:*

$$\dot{p}_h = -\frac{\partial H}{\partial q_h} \ , \qquad \dot{q}_h = \frac{\partial H}{\partial p_h} \ , \qquad h = 1,\ldots,n \ ,$$

*sono mutate dalla trasformazione di coordinate nelle equazioni canoniche relative a $K$:*

$$\dot{\tilde{p}}_h = -\frac{\partial K}{\partial \tilde{q}_h} \ , \qquad \dot{\tilde{q}}_h = \frac{\partial K}{\partial \tilde{p}_h} \ , \qquad h = 1,\ldots,n \ .$$

*Attenzione:* non è necessario che $K$ si ottenga da $H$ per sostituzione di variabili, ovvero che sia $K = H \circ \mathbf{w}$;[19] basta che le nuove equazioni del moto siano ancora canoniche, non importa con quale hamiltoniana (questa maggior libertà è essenziale, vedremo, per le trasformazioni canoniche dipendenti dal tempo). Le hamiltoniane $H$ e $K$ si dicono *canonicamente coniugate* dalla (4.25).

Diamo qui alcuni esempi di trasformazioni di coordinate, la cui canonicità si riconosce senza difficoltà sulla base della definizione:

i) La traslazione nello spazio delle fasi:

$$\mathbf{p} = \tilde{\mathbf{p}} + \mathbf{a} \ , \qquad \mathbf{q} = \tilde{\mathbf{q}} + \mathbf{b} \ ,$$

è canonica, e muta $H(\mathbf{p},\mathbf{q},t)$ in $K(\tilde{\mathbf{p}},\tilde{\mathbf{q}},t) = H(\tilde{\mathbf{p}}+\mathbf{a},\tilde{\mathbf{q}}+\mathbf{b},t)$.

ii) La dilatazione nello spazio delle fasi

$$\mathbf{p} = \alpha\tilde{\mathbf{p}} \ , \qquad \mathbf{q} = \beta\tilde{\mathbf{q}} \ , \qquad \alpha,\beta \neq 0 \ ,$$

è canonica, e muta una generica hamiltoniana $H(\mathbf{p},\mathbf{q},t)$ nell'hamiltoniana $K(\tilde{\mathbf{p}},\tilde{\mathbf{q}},t) = (\alpha\beta)^{-1}H(\alpha\tilde{\mathbf{p}},\beta\tilde{\mathbf{q}},t)$. In particolare, la trasformazione $p = m^{1/2}\tilde{p}$, $q = k^{-1/2}\tilde{q}$, muta l'hamiltoniana dell'oscillatore armonico, $H = \frac{1}{2m}p^2 + \frac{k}{2}q^2$,

---

[19] Il simbolo $\circ$ indica, in generale, la composizione di funzioni: se $f,g$ sono due funzioni reali di variabile reale, allora $(f\circ g)(x) \equiv f(g(x))$; la scrittura sopra utilizzata corrisponde allora a $K(\tilde{\mathbf{p}},\tilde{\mathbf{q}},t) = H(\mathbf{w}(\tilde{\mathbf{p}},\tilde{\mathbf{q}},t),t) = H(\mathbf{u}(\tilde{\mathbf{p}},\tilde{\mathbf{q}},t),\mathbf{v}(\tilde{\mathbf{p}},\tilde{\mathbf{q}},t),t)$.

nell'hamiltoniana $K = \frac{1}{2}\omega(\tilde{p}^2 + \tilde{q}^2)$, $\omega^2 = k/m$; si osservi che $\tilde{p}$ e $\tilde{q}$ hanno la stessa dimensione fisica, e che le traiettorie divengono cerchi percorsi con velocità angolare $\omega$.

iii) La trasformazione lineare

$$\mathbf{p} = (J^T)^{-1}\tilde{\mathbf{p}} \ , \qquad \mathbf{q} = J\tilde{\mathbf{q}},$$

ove $J$ è una qualunque matrice invertibile, è canonica, e la nuova hamiltoniana si ottiene per sostituzione di variabili in $H$, cioè si ha $K(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = H((J^T)^{-1}\tilde{\mathbf{p}}, J\tilde{\mathbf{q}}, t)$ (la verifica è un po' laboriosa, ma non difficile).

iv) La trasformazione che scambia le coordinate con i momenti, cioè

$$\mathbf{p} = \tilde{\mathbf{q}} \ , \qquad \mathbf{q} = \tilde{\mathbf{p}} \ ,$$

è canonica, e $K(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = -H(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}, t)$. Questa trasformazione, di nessuna utilità concreta, mostra tuttavia come, in ambito hamiltoniano, non vi sia alcuna sostanziale differenza tra coordinate e momenti coniugati.

v) La trasformazione di Galileo, cioè

$$\mathbf{p} = \tilde{\mathbf{p}} \ , \qquad \mathbf{q} = \tilde{\mathbf{q}} + \mathbf{c}t \ ,$$

ove $\mathbf{c}$ è un vettore costante, è canonica, e si ha $K(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = H(\tilde{\mathbf{p}}, \tilde{\mathbf{q}} + \mathbf{c}t, t) - \mathbf{c} \cdot \tilde{\mathbf{p}}$.

vi) La trasformazione

$$\mathbf{p} = \tilde{\mathbf{p}} \ , \qquad \mathbf{q} = \tilde{\mathbf{q}} + \alpha\tilde{\mathbf{p}}t \ ,$$

ove $\alpha$ è una costante reale, è canonica, e muta $H(\mathbf{p}, \mathbf{q}, t)$ in $K(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = H(\tilde{\mathbf{p}}, \tilde{\mathbf{q}} + \alpha\tilde{\mathbf{p}}t, t) - \frac{1}{2}\alpha\tilde{\mathbf{p}} \cdot \tilde{\mathbf{p}}$. Si osservi un fatto curioso:[20] se $\alpha = m^{-1}$, l'hamiltoniana della particella libera, $H = \frac{1}{2m}\mathbf{p} \cdot \mathbf{p}$, è mutata da questa trasformazione nella hamiltoniana costante $K = 0$. La dinamica nelle nuove coordinate è allora del tutto banale ($\tilde{\mathbf{p}}$ e $\tilde{\mathbf{q}}$ conservano il valore iniziale!); la parte interessante della dinamica è finita nelle equazioni della trasformazione, che per il problema della particella libera forniscono l'integrale generale ($\tilde{\mathbf{p}}$ e $\tilde{\mathbf{q}}$ sono le costanti arbitrarie).

Non è difficile vedere che non tutte le trasformazioni di coordinate della forma generica (4.25) sono canoniche. Consideriamo ad esempio la trasformazione che introduce, nel piano cartesiano $p\,q$, le coordinate polari $(\tilde{p}, \tilde{q}) \equiv (r, \vartheta)$, cioè

$$p = \tilde{p}\cos\tilde{q} \ , \qquad q = \tilde{p}\sin\tilde{q} \ , \tag{4.26}$$

e prendiamo le equazioni canoniche della particella libera di massa unitaria in coordinate cartesiane, $\dot{p} = 0$ e $\dot{q} = p$; è immediato verificare, per sostituzione nelle equazioni del moto, che le nuove equazioni sono $\dot{\tilde{p}} = \tilde{p}\cos\tilde{q}\sin\tilde{q}$, $\dot{\tilde{q}} = \cos^2\tilde{q}$. E' chiaro allora che non esiste nessuna funzione $K$ tale che sia $\dot{\tilde{p}} = -\frac{\partial K}{\partial \tilde{q}}$, $\dot{\tilde{q}} = \frac{\partial K}{\partial \tilde{p}}$: dovrebbe infatti risultare $-\frac{\partial \dot{\tilde{p}}}{\partial \tilde{p}} = \frac{\partial \dot{\tilde{q}}}{\partial \tilde{q}}$, mentre invece $-\frac{\partial \dot{\tilde{p}}}{\partial \tilde{p}} = -\cos\tilde{q}\sin\tilde{q}$, e $\frac{\partial \dot{\tilde{q}}}{\partial \tilde{q}} = -2\cos\tilde{q}\sin\tilde{q}$.

---

[20] Ma ispiratore di sviluppi interessanti: si veda più avanti l'*equazione di Hamilton–Jacobi*.

Una trasformazione simile alla (4.26), ma canonica, si può però costruire: si dimostra infatti che

vii) la trasformazione

$$p = \sqrt{2\tilde{p}}\cos\tilde{q} \ , \qquad q = \sqrt{2\tilde{p}}\sin\tilde{q} \ , \tag{4.27}$$

è canonica, e la nuova hamiltoniana si ottiene per sostituzione, cioè risulta $K = H(\sqrt{2\tilde{p}}\cos\tilde{q}, \sqrt{2\tilde{p}}\sin\tilde{q}, t)$; la verifica della canonicità è rinviata a più avanti. Vale la pena di osservare che l'hamiltoniana dell'oscillatore armonico, $H = \frac{1}{2}\omega(p^2+q^2)$, è mutata da questa trasformazione nell'hamiltoniana, particolarmente semplice, $K = \omega\tilde{p}$. Questa hamiltoniana serve come base per lo studio del comportamento (ancora oggi non compreso a sufficienza) di un sistema di oscillatori armonici debolmente accoppiati.

- **Osservazione.** Alcuni degli esempi dati sopra, assieme anche alla trasformazione (4.27), mostrano in particolare come l'insieme delle trasformazioni canoniche sia più vasto delle trasformazioni ammissibili in ambito lagrangiano. Viceversa, data una qualunque trasformazione di coordinate puntuale $\mathbf{q} = \mathbf{v}(\tilde{\mathbf{q}}, t)$, la cui inversa indichiamo con $\tilde{\mathbf{q}} = \tilde{\mathbf{v}}(\mathbf{q}, t)$, la si può sempre estendere canonicamente ai momenti, ponendo $p_h = u_h(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t)$, con

$$u_h(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = \sum_k \tilde{p}_k \frac{\partial \tilde{v}_k}{\partial q_h}(\mathbf{v}(\tilde{\mathbf{q}}, t), t) \ , \qquad h = 1, \ldots, n \ , \tag{4.28}$$

e la nuova hamiltoniana è data da

$$K(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = H(\mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t), \mathbf{v}(\tilde{\mathbf{q}}, t), t) + \tilde{\mathbf{p}} \cdot \frac{\partial \tilde{\mathbf{v}}}{\partial t}(\mathbf{v}(\tilde{\mathbf{q}}, t), t) \ .$$

La verifica diretta della canonicità di questa trasformazione, per hamiltoniane del tutto generiche, è possibile, ma molto laboriosa (non è però difficile verificare che la (4.28) corrisponde alla definizione dei momenti data in ambito lagrangiano). Una dimostrazione semplicissima si vedrà più avanti, quando avremo a disposizione opportuni criteri di canonicità.

- **Osservazione.** Dagli esempi si vede che, come già si è osservato, in generale la nuova hamiltoniana non si ottiene dalla vecchia per semplice sostituzione di variabili. Si osservi però che in tutti gli esempi sopra riportati la nuova hamiltoniana è legata alla vecchia da una relazione lineare del tipo

$$K \ = \ c\,H \circ \mathbf{w} \ + \ K_0 \ ,$$

ove $c$ è una costante, mentre $K_0$ dipende dalla sola trasformazione (cioè non dalla hamiltoniana $H$ che si vuole trasformare); inoltre, $K_0$ è presente per le sole trasformazioni esplicitamente dipendenti dal tempo. Vedremo che questo è il caso generale. Come si intuisce dall'esempio ii), la costante $c$ si può sempre modificare a piacere, e in particolare riportare a uno, con un una opportuna dilatazione (un cambio di unità di misura); tale costante gioca dunque un ruolo del tutto inessenziale.

- **Osservazione.** Le trasformazioni canoniche costituiscono un gruppo. La verifica è però più delicata di quello che sembra: mentre infatti è chiaro, sulla base della definizione, che la trasformazione identica $\mathbf{p} = \tilde{\mathbf{p}}$, $\mathbf{q} = \tilde{\mathbf{q}}$ è canonica, e che la composizione di due trasformazioni canoniche è canonica, non è altrettanto evidente (avendo in mano la sola definizione) che l'inversa di una trasformazione canonica sia necessariamente canonica: non è chiaro infatti che *comunque si prenda l'hamiltoniana* $K$ esista $H$ ad essa coniugata (mentre è vero, per definizione, che *comunque si prenda* $H$ esiste $K$ ad essa coniugata). Questa piccola difficoltà sarà superata più avanti.

Naturalmente, perché le trasformazioni canoniche possano essere uno strumento utile, sono indispensabili dei "criteri di canonicità", per quanto possibile nella forma di condizioni necessarie e sufficienti, che permettano di garantire a priori la canonicità di una trasformazione, e consentano anche di scrivere direttamente la nuova hamiltoniana, senza passare attraverso la sostituzione di variabili nelle equazioni del moto, come si è fatto sopra. A questo argomento sono dedicati i prossimi paragrafi.

### *4.2.2 Una condizione necessaria e sufficiente per la canonicità*

In questo paragrafo studieremo una condizione necessaria e sufficiente per la canonicità di una trasformazione di coordinate, che chiameremo *condizione di Lie*; tale condizione è particolarmente importante, sia perché costituisce una caratterizzazione profonda della canonicità, sia perché è una buona base per poi introdurre i criteri di canonicità cui si è sopra accennato.

Per una generica funzione $f(\mathbf{p}, \mathbf{q}, t)$, definiamo il differenziale "a tempo bloccato", o *differenziale virtuale* $\mathrm{d}^* f$ nel modo seguente:

$$\mathrm{d}^* f \;=\; \sum_{h=1}^{n} \Big( \frac{\partial f}{\partial p_h}\, \mathrm{d}p_h + \frac{\partial f}{\partial q_h}\, \mathrm{d}q_h \Big) \;=\; \mathrm{d}f - \frac{\partial f}{\partial t}\, \mathrm{d}t \ .$$

Per i differenziali delle variabili indipendenti risulta, evidentemente, $\mathrm{d}^* p_h = \mathrm{d}p_h$, $\mathrm{d}^* q_h = \mathrm{d}q_h$, $h = 1, \ldots, n$. Assegnata una trasformazione di coordinate

$$\mathbf{p} = \mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \ , \qquad \mathbf{q} = \mathbf{v}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \ , \tag{4.29}$$

nella quale compaiono $\tilde{p}_h$, $\tilde{q}_h$, $h = 1, \ldots, n$, come variabili indipendenti, restano in particolare definiti i differenziali virtuali delle funzioni $u_h$, $v_h$ che esprimono le vecchie variabili in funzione delle nuove, ad esempio $\mathrm{d}^* v_h = \sum_k (\frac{\partial v_h}{\partial \tilde{p}_k}\, \mathrm{d}\tilde{p}_k + \frac{\partial v_h}{\partial \tilde{q}_k}\, \mathrm{d}\tilde{q}_k) = \mathrm{d}v_h - \frac{\partial v_h}{\partial t}\, \mathrm{d}t$. Ciò premesso, diamo la seguente

**Definizione 4.2:** *Diremo che la trasformazione di coordinate (4.29) soddisfa la "condizione di Lie", se esistono una costante $c \neq 0$ e una funzione $F(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t)$, tali che sia* *identicamente soddisfatta l'uguaglianza differenziale*[21)]

$$c\,\mathbf{u}\cdot \overset{*}{\mathrm{d}}\mathbf{v} \,=\, \tilde{\mathbf{p}}\cdot \overset{*}{\mathrm{d}}\tilde{\mathbf{q}} + \overset{*}{\mathrm{d}}F \tag{4.30}$$

*(cioè* $c\sum_h u_h\,\overset{*}{\mathrm{d}}v_h = \sum_h \tilde{p}_h\,\overset{*}{\mathrm{d}}\tilde{q}_h + \overset{*}{\mathrm{d}}F$*).*

- **Osservazione.** Con un piccolo abuso di notazione si può scrivere la condizione di Lie nella forma più agevole
$$c\mathbf{p}\cdot \overset{*}{\mathrm{d}}\mathbf{q} \,=\, \tilde{\mathbf{p}}\cdot \overset{*}{\mathrm{d}}\tilde{\mathbf{q}} + \overset{*}{\mathrm{d}}F\ ;$$
da questa scrittura non è immediatamente evidente quali siano le variabili indipendenti (le $\tilde{\mathbf{p}},\tilde{\mathbf{q}}$ o le $\mathbf{p},\mathbf{q}$) cui si sta facendo riferimento, ma la cosa non è grave perché, come è ben noto, la scelta delle variabili indipendenti nelle quali si verifica l'uguaglianza è irrilevante. E' poi immediato verificare che la condizione di Lie si può anche scrivere nella forma equivalente (che sarà impiegata tra breve)
$$c\,\mathbf{u}\cdot \mathrm{d}\mathbf{v} \,=\, \tilde{\mathbf{p}}\cdot \mathrm{d}\tilde{\mathbf{q}} + \mathrm{d}F - K_0\,\mathrm{d}t\ , \qquad K_0 = \frac{\partial F}{\partial t} - c\,\mathbf{u}\cdot\frac{\partial \mathbf{v}}{\partial t}\ , \tag{4.31}$$
nella quale non compaiono differenziali virtuali.

- **Esempio.** Tutte le trasformazioni i) – vii) introdotte nel paragrafo 4.2.1 soddisfano la condizione di Lie, precisamente

| | |
|---|---|
| i) | con $c=1$ e $F=\mathbf{a}\cdot\tilde{\mathbf{q}}$; |
| ii) | con $c=(\alpha\beta)^{-1}$ e $F=0$; |
| iii, v) | con $c=1$ e $F=0$; |
| iv) | con $c=-1$ e $F=-\tilde{\mathbf{p}}\cdot\tilde{\mathbf{q}}$; |
| vi) | con $c=1$ e $F=\frac{1}{2}\alpha t\tilde{\mathbf{p}}\cdot\tilde{\mathbf{p}}$; |
| vii) | con $c=1$ e $F=\tilde{p}\sin\tilde{q}\cos\tilde{q}$. |

Vale ora il seguente fondamentale teorema:

**Proposizione 4.7:** *Se la trasformazione di coordinate (4.29) soddisfa la condizione di Lie, con riferimento a una costante* $c$ *e a una funzione* $F$*, allora essa è canonica, e muta* $H$ *in*

$$K \,=\, c\,H\circ\mathbf{w} + K_0\ , \tag{4.32}$$

*con* $K_0$ *dato dalla (4.31). Viceversa, se la (4.29) è canonica, allora esistono una costante* $c\neq 0$ *e una funzione* $F$*, in corrispondenza alle quali la condizione di Lie è soddisfatta (e di conseguenza la nuova hamiltoniana* $K$ *ha l'espressione (4.32)).*

La costante $c$ prende il nome di *valenza* della trasformazione; le trasformazioni canoniche con $c = 1$ sono chiamate *univalenti.* E' chiaro, dall'esempio ii), che ci si può sempre riportare a trasformazioni univalenti con una opportuna dilatazione (con

---

21) Nel linguaggio delle forme differenziali, ciò significa che, pur non essendo esatte le 1-forme $c\mathbf{u}\cdot \overset{*}{\mathrm{d}}\mathbf{v}$ e $\tilde{\mathbf{p}}\cdot \overset{*}{\mathrm{d}}\tilde{\mathbf{q}}$, lo è però la loro differenza (a tempo, s'intende, fissato).

$\alpha\beta = c$). La valenza non gioca dunque nessun ruolo importante nella teoria delle trasformazioni canoniche; anzi, in taluni testi, il nome di trasformazioni canoniche è riservato alle sole trasformazioni univalenti, e corrispondentemente la condizione di Lie si scrive senza la costante $c$:

$$\mathbf{p} \cdot \overset{*}{\mathrm{d}}\mathbf{q} = \tilde{\mathbf{p}} \cdot \overset{*}{\mathrm{d}}\tilde{\mathbf{q}} + \overset{*}{\mathrm{d}}F \ .$$

Le trasformazioni univalenti e indipendenti dal tempo, per le quali risulta evidentemente $K = H \circ \mathbf{w}$, sono a volte chiamate trasformazioni *completamente canoniche.*

- **Osservazione.** Resta così dimostrata l'affermazione, anticipata sopra, che la nuova hamiltoniana ha sempre la forma (4.32), con $K_0$ dipendente dalla sola trasformazione.

Diamo qui una dimostrazione della sola prima parte della proposizione, facendo uso dei principi variazionali, precisamente della generalizzazione di Helmholtz del principio di Hamilton (paragrafo 4.1.8). Una dimostrazione indiretta della seconda parte si vedrà più avanti.

**Dimostrazione. (Prima parte)** Consideriamo una trasformazione di variabili soddisfacente la condizione di Lie (4.31); allora, per ogni hamiltoniana $H$, resta definita dalla (4.32) la funzione $K$, con $K_0$ data dalla (4.31). Siano ora $\big(\mathbf{p}(t), \mathbf{q}(t)\big)$, e $\big(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t)\big)$ dei moti a priori possibili tra gli istanti arbitrari $t_0$ e $t_1$, coniugati dalla trasformazione (4.29), cioè tali che sia, istante per istante,

$$\mathbf{p}(t) = \mathbf{u}\big(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t\big) \ , \qquad \mathbf{q}(t) = \mathbf{v}\big(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t\big) \ .$$

Per per tali moti scriviamo l'azione hamiltoniana, precisamente

$$S = \int_{t_0}^{t_1} (\mathbf{p} \cdot \dot{\mathbf{q}} - H) \,\mathrm{d}t \ , \qquad \tilde{S} = \int_{t_0}^{t_1} (\tilde{\mathbf{p}} \cdot \dot{\tilde{\mathbf{q}}} - K) \,\mathrm{d}t \ ,$$

e mostriamo innanzitutto che si ha

$$cS = \tilde{S} + F(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t)\Big|_{t_0}^{t_1} \ . \tag{4.33}$$

Infatti, dalla condizione di Lie (4.31) segue subito, per integrazione,

$$\begin{aligned} c\int_{t_0}^{t_1} \mathbf{p}(t) \cdot \dot{\mathbf{q}}(t) \,\mathrm{d}t &= c\int_{t_0}^{t_1} \mathbf{u}(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t) \cdot \dot{\mathbf{v}}(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t) \,\mathrm{d}t \\ &= \int_{t_0}^{t_1} [\tilde{\mathbf{p}}(t) \cdot \dot{\tilde{\mathbf{q}}}(t) - K_0(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t)] \,\mathrm{d}t + \int_{t_0}^{t_1} \dot{F}(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t) \,\mathrm{d}t \ , \end{aligned}$$

e di conseguenza, per la definizione di $K$, si ha

$$\begin{aligned} &c\int_{t_0}^{t_1} [\mathbf{p}(t) \cdot \dot{\mathbf{q}}(t) - H(\mathbf{p}(t), \mathbf{q}(t), t)] = \\ &\qquad \int_{t_0}^{t_1} [\tilde{\mathbf{p}}(t) \cdot \dot{\tilde{\mathbf{q}}}(t) - K(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t)] \,\mathrm{d}t + F(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t), t)\Big|_{t_0}^{t_1} \ ; \end{aligned}$$

la (4.33) è allora immediata.

La conclusione della dimostrazione è ora una facile conseguenza del principio variazionale: poichè infatti la (4.33) vale per il moto di riferimento come per tutti i moti variati, ed in essa $F$ compare solo attraverso i valori agli estremi, allora per variazioni arbitrarie, purché nulle agli estremi, si ha $c\delta S = \delta\tilde{S}$; pertanto, $\delta S$ e $\delta\tilde{S}$ si annullano simultaneamente, e dunque il moto $(\mathbf{p}(t), \mathbf{q}(t))$ soddisfa le equazioni canoniche relative all'hamiltoniana $H$, se e solo se il moto $(\tilde{\mathbf{p}}(t), \tilde{\mathbf{q}}(t))$ soddisfa le equazioni canoniche relativa all'hamiltoniana $K$. *Q.E.D.*

- **Osservazione.** Non è difficile verificare che le trasformazioni che soddisfano la condizione di Lie formano un gruppo (per acquistare familiarità con la condizione è bene eseguire esplicitamente la verifica). Resta allora dimostrato che le trasformazioni canoniche costituiscono un gruppo (senza la piccola difficoltà legata alla canonicità della trasformazione inversa, discussa nel paragrafo 4.2.1).
- **Osservazione.** Data la trasformazione di coordinate dipendente dal tempo (4.29), ad essa è naturalmente associata una famiglia ad un parametro di trasformazioni indipendenti dal tempo (o "a tempo bloccato")
$$\mathbf{p} = \mathbf{u}^\alpha(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) , \qquad \mathbf{q} = \mathbf{v}^\alpha(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) , \tag{4.34}$$
ottenuta ponendo $\mathbf{u}^\alpha(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) = \mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, \alpha)$, $\mathbf{v}^\alpha(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) = \mathbf{v}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, \alpha)$. Viceversa, data una famiglia ad un parametro di trasformazioni indipendenti dal tempo, con essa si può sempre costruire una singola trasformazione dipendente dal tempo. Poichè nella condizione di Lie (4.30) il tempo non gioca alcun ruolo (grazie al fatto che vi compaiono soltanto differenziali virtuali), si deduce immediatamente il seguente corollario:

  *La trasformazione (4.29) è canonica, se e solo se ciascuna delle trasformazioni a tempo bloccato (4.34) è canonica, con valenza $c$ indipendente da $\alpha$.*

### 4.2.3 Funzioni generatrici di trasformazioni canoniche

Ci restringiamo qui, per semplicità di linguaggio, alle sole trasformazioni univalenti; l'estensione di quanto diremo al caso $c \neq 1$ è del tutto banale.

La presenza nella condizione di Lie (4.31) dei differenziali $\mathrm{d}\mathbf{v} \equiv \mathrm{d}\mathbf{q}$ e $\mathrm{d}\tilde{\mathbf{q}}$ "invita", per così dire, a usare come variabili indipendenti le variabili "miste" $q_h$, $\tilde{q}_h$, al posto delle originali $\tilde{p}_h$, $\tilde{q}_h$, $h = 1, \ldots, n$. Ciò è possibile, evidentemente, se tali variabili sono effettivamente indipendenti, cioè se le equazioni della trasformazione si possono equivalentemente scrivere nella forma implicita
$$\tilde{p}_h = \tilde{p}_h(\mathbf{q}, \tilde{\mathbf{q}}, t) , \qquad p_h = p_h(\mathbf{q}, \tilde{\mathbf{q}}, t) , \qquad h = 1, \ldots, n . \tag{4.35}$$
A questo scopo è sufficiente che risulti
$$\det\left(\frac{\partial v_h}{\partial \tilde{p}_k}\right) \neq 0 ; \tag{4.36}$$

in tal caso infatti il sistema

$$q_h = v_h(\tilde p_1, \ldots, \tilde p_n, \tilde q_1, \ldots, \tilde q_n, t) \ , \qquad h = 1, \ldots, n \tag{4.37}$$

si può invertire rispetto a $\tilde p_1, \ldots, \tilde p_n$, fornendo così la prima delle (4.35); per sostituzione in $p_h = u_h(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t)$ si ottiene poi la seconda.

Se ciò è possibile, allora, posto

$$F_1(\mathbf{q}, \tilde{\mathbf{q}}, t) = F\big(\tilde{\mathbf{p}}(\mathbf{q}, \tilde{\mathbf{q}}, t), \tilde{\mathbf{q}}, t\big) \ , \tag{4.38}$$

possiamo scrivere la condizione di Lie nella forma

$$\mathbf{p} \cdot \mathrm{d}\mathbf{q} - \tilde{\mathbf{p}} \cdot \mathrm{d}\tilde{\mathbf{q}} + K_0 \, \mathrm{d}t \ = \ \mathrm{d}F_1$$

(anche $K_0$ è qui pensato funzione di $\mathbf{q}$ e $\tilde{\mathbf{q}}$). Tale forma è sostanzialmente più semplice della (4.30), perché a sinistra compaiono i soli differenziali delle variabili indipendenti di cui è funzione $F_1$: l'immediata conseguenza è che si ha, necessariamente,

$$\tilde p_h(\mathbf{q}, \tilde{\mathbf{q}}, t) \ = \ -\frac{\partial F_1}{\partial \tilde q_h}(\mathbf{q}, \tilde{\mathbf{q}}, t) \ , \qquad p_h(\mathbf{q}, \tilde{\mathbf{q}}, t) \ = \ \frac{\partial F_1}{\partial q_h}(\mathbf{q}, \tilde{\mathbf{q}}, t) \qquad K_0 = \frac{\partial F_1}{\partial t} \ ; \tag{4.39}$$

pertanto $F_1$ fornisce direttamente, con le sue derivate, le equazioni della trasformazione in variabili miste (cioè le (4.35)), e insieme la relazione tra la vecchia e la nuova hamiltoniana (ancora in variabili miste), precisamente

$$K \ = \ H + \frac{\partial F_1}{\partial t} \ . \tag{4.40}$$

Osserviamo ora che (per il teorema della funzione implicita, ricordando che le $\tilde p_h = \tilde p_h(q_1, \ldots, q_n, \tilde q_1, \ldots, \tilde q_n, t)$ sono state ottenuta per inversione) la matrice $\big(\frac{\partial \tilde p_h}{\partial q_k}\big)$ ha certamente determinante diverso da zero; pertanto si ha necessariamente

$$\det\Big(\frac{\partial^2 F_1}{\partial q_h \partial \tilde q_k}\Big) \ \neq \ 0 \ . \tag{4.41}$$

Si vede dunque che, se la condizione di invertibilità (4.36) è soddisfatta, allora esiste una funzione $F_1(\mathbf{q}, \tilde{\mathbf{q}}, t)$, soddisfacente la (4.41), che determina implicitamente la trasformazione canonica tramite le (4.39).

Viceversa (ed è questo il vero motivo di interesse per quello che stiamo facendo) vale la seguente

**Proposizione 4.8:** *Per ogni funzione regolare $F_1(\mathbf{q}, \tilde{\mathbf{q}}, t)$ che soddisfi la condizione di invertibilità (4.41), le (4.39) definiscono implicitamente (cioè a meno di una inversione) una trasformazione canonica, e il legame tra la vecchia e la nuova hamiltoniana è dato dalla (4.40).*

**Dimostrazione.** La (4.41) garantisce l'invertibilità delle (4.39) rispetto alle $\tilde{\mathbf{q}}$, e dunque la possibilità di esprimere $\mathbf{q}$ e $\mathbf{p}$ come funzioni di $\tilde{\mathbf{p}}$, $\tilde{\mathbf{q}}$ e $t$ (si effettua l'inversione nella prima delle (4.39), e si sostituisce nella seconda); pertanto le (4.39) definiscono effettivamente una trasformazione di variabili. Per quanto visto sopra, la condizione di Lie è certamente soddisfatta, con $K_0 = \frac{\partial F_1}{\partial t}$, per cui la trasformazione è canonica, e vale la (4.40). *Q.E.D.*

Si dice che la funzione $F_1$ *genera*, tramite le (4.39), una trasformazione canonica. E' chiaro che si ottengono, per questa via, tutte e sole le trasformazioni canoniche che consentono di prendere $q_h$ e $\tilde{q}_h$ come variabili indipendenti (assieme a $t$). Queste trasformazioni si chiamano *libere*.[22] E' evidente che tali trasformazioni (pur molto interessanti per taluni aspetti, ad esempio per la facilità di comporle tra loro) non solo non esauriscono le trasformazioni canoniche, ma anzi ne escludono una classe importantissima, precisamente tutte le trasformazioni puntuali, compresa la trasformazione più elementare, cioè la trasformazione identica $p_h = \tilde{p}_h$, $q_h = \tilde{q}_h$.

E' tuttavia chiaro che la scelta delle $q_h$ e $\tilde{q}_h$ come variabili miste indipendenti è solo una delle scelte possibili. Ad esempio, se si può invertire il sistema (4.37) rispetto alle $\tilde{q}_h$, allora si possono usare le variabili indipendenti $q_h$ e $\tilde{p}_h$, e scrivere le equazioni della trasformazione nella diversa forma mista

$$\tilde{q}_h \;=\; \tilde{q}_h(\tilde{\mathbf{p}}, \mathbf{q}, t) \;, \qquad p_h \;=\; p_h(\tilde{\mathbf{p}}, \mathbf{q}, t) \;;$$

ciò è possibile se risulta $\det\left(\frac{\partial v_h}{\partial \tilde{q}_k}\right) \neq 0$. Vale la pena di osservare che questa condizione è certamente verificata da tutte trasformazioni "prossime all'identità", ad esempio trasformazioni del tipo

$$p_h = \tilde{p}_h + \varepsilon f_h(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \;, \qquad q_h = \tilde{q}_h + \varepsilon g_h(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \;, \tag{4.42}$$

per $\varepsilon$ abbastanza piccolo: per convincersene, basta osservare che si ha $\det\left(\frac{\partial v_h}{\partial \tilde{q}_k}\right) = \det\left(\delta_{hk} + \varepsilon \frac{\partial g_h}{\partial \tilde{q}_k}\right) = 1 + \mathcal{O}(\varepsilon)$.

Posto allora $F_2 = F + \tilde{\mathbf{p}} \cdot \tilde{\mathbf{q}}$, più precisamente

$$F_2(\tilde{\mathbf{p}}, \mathbf{q}, t) = F\big(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}(\tilde{\mathbf{p}}, \mathbf{q}, t), t\big) + \tilde{\mathbf{p}} \cdot \tilde{\mathbf{q}}(\tilde{\mathbf{p}}, \mathbf{q}, t) \;,$$

la condizione di Lie prende la forma

$$\mathbf{p} \cdot \mathrm{d}\mathbf{q} + \tilde{\mathbf{q}} \cdot \mathrm{d}\tilde{\mathbf{p}} + K_0 \,\mathrm{d}t \;=\; \mathrm{d}F_2 \;;$$

si ottiene allora immediatamente la seguente

**Proposizione 4.9:** *Per ogni funzione regolare $F_2(\tilde{\mathbf{p}}, \mathbf{q}, t)$, purché sia soddisfatta la condizione di invertibilità*

$$\det\left(\frac{\partial^2 F_2}{\partial \tilde{p}_h \partial q_k}\right) \neq 0 \;, \tag{4.43}$$

*le equazioni*

$$p_h \;=\; \frac{\partial F_2}{\partial q_h} \;, \qquad \tilde{q}_h \;=\; \frac{\partial F_2}{\partial \tilde{p}_h} \;, \tag{4.44}$$

*definiscono implicitamente una trasformazione canonica, mentre la relazione tra la vecchia e la nuova hamiltoniana è data da*

$$K \;=\; H + \frac{\partial F_2}{\partial t} \;. \tag{4.45}$$

---

[22] Il termine ricorda appunto il fatto che le coordinate configurazionali $\mathbf{q}$ e $\tilde{\mathbf{q}}$ sono "libere" le une rispetto alle altre.

**Dimostrazione.** Come per la proposizione precedente, la (4.43) garantisce l'invertibilità delle (4.44), e dunque l'effettiva esistenza di una trasformazione; la condizione di Lie è automaticamente soddisfatta, con $K_0 = \frac{\partial F_2}{\partial t}$, pertanto la trasformazione è canonica, e vale anche la (4.45). *Q.E.D.*

Con una funzione di tipo $F_2(\tilde{\mathbf{p}}, \mathbf{q})$ si possono generare sostanzialmente tutte le trasformazioni, tranne quelle, ben poco utili, in cui almeno alcune delle vecchie coordinate sono funzioni dei soli nuovi momenti (come nell'esempio iv. del paragrafo 4.2.1). L'identità è generata da $F_2 = \tilde{\mathbf{p}} \cdot \mathbf{q}$, come si riconosce immediatamente; funzioni del tipo

$$F_2 = \tilde{\mathbf{p}} \cdot \mathbf{q} + \varepsilon G(\tilde{\mathbf{p}}, \mathbf{q}, t, \varepsilon)$$

(con $G$ regolare in tutte le variabili) generano invece trasformazioni prossime all'identità, come le (4.42), di uso prezioso nella teoria classica delle perturbazioni (si osservi che per trasformazioni di questa forma la condizione di invertibilità (4.42) è sempre soddisfatta, se $\varepsilon$ è abbastanza piccolo). Infine, non è difficile vedere che le trasformazioni puntuali sono generate da funzioni $F_2$ del tipo $F_2 = \tilde{\mathbf{p}} \cdot \tilde{\mathbf{v}}(\mathbf{q}, t)$, e che l'estensione ai momenti è data dalla (4.28).

- **Osservazione.** Si dimostra così, come già anticipato, che la (4.28) è canonica.

- **Esercizio 4.8:** Si determini la funzione generatrice $F_2$ per le trasformazioni canoniche i), iii), v), vi) e vii) introdotte nel paragrafo 4.2.1.
  Risposte:
  i) $F_2 = \sum_h (\tilde{p}_h + a_h)(q_h - b_h)$ ;
  iii) $F_2 = \sum_{hk} \tilde{p}_h (J^{-1})_{hk} q_k$ ;
  v) $F_2 = \sum_h \tilde{p}_h (q_h - c_h t)$ ;
  vi) $F_2 = \sum_h \tilde{p}_h (q_h - \frac{1}{2}\alpha \tilde{p}_h t)$ ;
  vii) $F_2 = \int_0^q \sqrt{2\tilde{p} - x^2}\, \mathrm{d}x$ ;

Non vi è difficoltà ad usare le altre coppie di variabili miste indipendenti, precisamente $\tilde{p}_h$, $p_h$ e $\tilde{q}_h$, $p_h$, ottenendo, come si verifica facilmente,

$$q_h = -\frac{\partial F_3}{\partial p_h}\,, \qquad \tilde{p}_h = -\frac{\partial F_3}{\partial \tilde{q}_h}\,, \qquad K = H + \frac{\partial F_3}{\partial t}$$
$$q_h = -\frac{\partial F_4}{\partial p_h}\,, \qquad \tilde{q}_h = \frac{\partial F_4}{\partial \tilde{p}_h}\,, \qquad K = H + \frac{\partial F_4}{\partial t}\,,$$

con $F_3 = F - \mathbf{p} \cdot \mathbf{q}$ e $F_4 = F + \tilde{\mathbf{p}} \cdot \tilde{\mathbf{q}} - \mathbf{p} \cdot \mathbf{q}$; questo modo di generare le trasformazioni canoniche è tuttavia di uso meno frequente.

L'utilità principale delle funzioni generatrici è quella di fornire in modo semplice delle classi ampie di trasformazioni di coordinate, la cui canonicità sia garantita a priori; ad esempio, invece di affrontare direttamente le equazioni del moto di un sistema, in alcuni casi conviene prima cercare una trasformazione canonica che sostituisca all'hamiltoniana $H$ assegnata una hamiltoniana $K$ in qualche senso più semplice. Una via possibile è allora quella di scrivere le equazioni della trasformazione per mezzo di una funzione generatrice, cercando di determinare quest'ultima in modo di ottenere la semplificazione desiderata. Le difficoltà, si può immaginare, non spariscono (come vedremo tra breve, per questa via ci si imbatte immediatamente in equazioni alle derivate parziali, per la funzione generatrice, di non facile soluzione); nondimeno, in alcuni casi questo modo di procedere conduce a dei progressi veramente sostanziali.

- **Osservazione.** Non sarebbe difficile vedere che si possono anche introdurre funzioni generatrici "miste", facendo riferimento, ad esempio, alle coordinate indipendenti $q_h, \tilde q_h$, come con la funzione generatrice $F_1$, per alcuni gradi di libertà, e alle coordinate $\tilde p_h, q_h$ (o $p_h, \tilde q_h$, o $\tilde p_h, p_h$), come con la funzione generatrice $F_2$ (o $F_3$, o $F_4$), per gli altri gradi di libertà. Tali funzioni generatrici miste sono di scarsa utilità; si potrebbe però dimostrare una cosa abbastanza interessante, e cioè che con tali funzioni generatrici miste (anzi, è sufficiente introdurre coppie di variabili di due soli tipi) si possono generare *tutte* le trasformazioni canoniche.

### *4.2.4 Il "flusso hamiltoniano" come trasformazione canonica.*

Data una equazione differenziale in $\mathbb{R}^n$, $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, con $\mathbf{f}$ regolare, il suo integrale generale $\boldsymbol{\Phi}(\tilde{\mathbf{x}}, t)$, espresso in termini del dato iniziale $\tilde{\mathbf{x}}$, fornisce per ogni $t$ una trasformazione regolare e invertibile: $\mathbb{R}^n \to \mathbb{R}^n$; la famiglia di queste trasformazioni, che possiamo anche considerare come una singola trasformazione dipendente dal tempo, prende il nome di *flusso* generato dall'equazione differenziale.[23] Il flusso generato dalle equazioni di Hamilton prende il nome di *flusso hamiltoniano*, e costituisce una trasformazione canonica, come è stabilito dalla seguente importante

**Proposizione 4.10:** *Si consideri un qualsiasi sistema hamiltoniano, di hamiltoniana $H$, e si indichi con*

$$\mathbf{p} = \mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \ , \qquad \mathbf{q} = \mathbf{v}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \ , \tag{4.46}$$

*il suo flusso, dove $\tilde{\mathbf{p}}$, $\tilde{\mathbf{q}}$ sono i dati iniziali. La trasformazione (4.46) è canonica, precisamente soddisfa la condizione di Lie (4.30), con $c = 1$ e*

$$F(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = \int_0^t \left[\mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t') \cdot \frac{\partial \mathbf{v}}{\partial t}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t') - H\big(\mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t'), \mathbf{v}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t'), t'\big)\right] \mathrm{d}t' \ . \tag{4.47}$$

**Dimostrazione.** Scrivendo esplicitamente la condizione di Lie (4.30), si vede facilmente che bisogna e basta verificare che la trasformazione soddisfa, per ogni $h = 1, \dots, n$, le condizioni

$$\begin{aligned} \frac{\partial F}{\partial \tilde p_h}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) &= \sum_{k=1}^n u_k(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \frac{\partial v_k}{\partial \tilde p_h}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \\ \frac{\partial F}{\partial \tilde q_h}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) &= \sum_{k=1}^n u_k(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \frac{\partial v_k}{\partial \tilde q_h}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) - \tilde p_h \ . \end{aligned} \tag{4.48}$$

[23] Si riveda, per la nozione di flusso, il primo capitolo.

Vediamo la prima di esse. Dalla definizione di $F$ segue

$$\frac{\partial F}{\partial \tilde{p}_h} = \int_0^t \sum_{k=1}^n \Big[\frac{\partial u_k}{\partial \tilde{p}_h}\frac{\partial v_k}{\partial t} + u_k \frac{\partial^2 v_k}{\partial t \partial \tilde{p}_h} - \frac{\partial H}{\partial p_k}\frac{\partial u_k}{\partial \tilde{p}_h} - \frac{\partial H}{\partial q_k}\frac{\partial v_k}{\partial \tilde{p}_h}\Big]\,\mathrm{d}t' \ ;$$

ma per ipotesi le (4.46) sono soluzioni delle equazioni di Hamilton relative all'hamiltoniana $H$, e dunque si ha $\frac{\partial H}{\partial p_k} = \frac{\partial v_k}{\partial t}$, $-\frac{\partial H}{\partial q_k} = \frac{\partial u_k}{\partial t}$. Segue allora

$$\frac{\partial F}{\partial \tilde{p}_h} = \int_0^t \sum_{k=1}^n \Big[u_k \frac{\partial^2 v_k}{\partial t \partial \tilde{p}_h} + \frac{\partial u_k}{\partial t}\frac{\partial v_k}{\partial \tilde{p}_h}\Big]\,\mathrm{d}t' = \int_0^t \frac{\partial}{\partial t}\sum_{k=1}^n u_k \frac{\partial v_k}{\partial \tilde{p}_h}\,\mathrm{d}t' = \sum_{k=1}^n u_k \frac{\partial v_k}{\partial \tilde{p}_h}\Big|_0^t$$

(nell'ultimo passaggio si è tenuto conto che, essendo costanti $\tilde{\mathbf{p}}$ e $\tilde{\mathbf{q}}$, derivata parziale e totale rispetto a $t$ coincidono). Infine, osserviamo che si ha $v_k(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, 0) = \tilde{q}_k$ (si ricordi che $\tilde{\mathbf{p}}$, $\tilde{\mathbf{q}}$ sono i dati iniziali), e dunque $\frac{\partial v_k}{\partial \tilde{p}_h}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, 0) = 0$; la prima delle (4.48) è allora immediata. Procedendo in modo analogo si ottiene facilmente anche la seconda. *Q.E.D.*

- **Osservazione.** Dall'espressione di $F$, tenendo conto del significato di $\mathbf{u}$ e $\mathbf{v}$, si vede che $F$ coincide con l'azione hamiltoniana $S$, calcolata per il moto vero, ed espressa in funzione del dato iniziale $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$ e del tempo.

Questo risultato fornisce un metodo particolarmente efficace per generare trasformazioni canoniche prossime all'identità, chiamato *metodo di Lie*, il cui uso va diffondendosi nell'ambito della teoria delle perturbazioni per sistemi hamiltoniani. Il caso più semplice è quello in cui si genera una trasformazione canonica "$\varepsilon$–vicina all'identità" come flusso al tempo $t = \varepsilon$ generato da una opportuna hamiltoniana $H$ indipendente dal tempo;[24] in questo caso, benchè la trasformazione stessa sia definita solo implicitamente, si ottiene immediatamente, in corrispondenza a una qualsiasi funzione $f$, l'espressione della funzione trasformata $\tilde{f} = f \circ \mathbf{w}$, in forma di serie di Taylor in $\varepsilon$: si ha infatti $\dot{f} = \{f, H\}$, e dunque $\frac{\mathrm{d}^r f}{\mathrm{d}t^r} = \mathcal{L}^r f$, $r = 1, 2 \ldots$, ove $\mathcal{L}$ indica l'operatore differenziale $\{\,.\,, H\}$ (ovvero la derivata di Lie associata al campo vettoriale Hamiltoniano; si riveda, al Capitolo 3, il paragrafo sulle parentesi di Poisson). Segue immediatamente l'espressione cercata per $\tilde{f}$:

$$\tilde{f} = \sum_{r=0}^{\infty} \frac{\varepsilon^r}{r!}\mathcal{L}^r f = e^{\varepsilon \mathcal{L}} f$$

(non ci addentriamo qui nei problemi, un po' delicati anche se non difficili, della convergenza di questa serie). Si osservi che, in questo modo, la funzione trasformata è nota fino all'ordine voluto in $\varepsilon$, senza bisogno di inversioni. Una buona introduzione al metodo di Lie per la generazione di trasformazioni canoniche (o anche non canoniche, con opportune generalizzazioni) si trova in W. Gröbner, *Serie di Lie e loro applicazioni* (Cremonese, Roma 1973).

---

[24] Si tratta di una hamiltoniana "fittizia", che non ha niente a che vedere con il problema hamiltoniano che si sta studiando.

- **Osservazione.** Indichiamo con $\hat{H}$ la particolare hamiltoniana che genera con il suo flusso la trasformazione (4.46); poichè tale trasformazione, come abbiamo visto, è canonica, allora essa associa ad ogni hamiltoniana $H$ una opportuna hamiltoniana $K$. E' facile vedere che all'hamiltoniana $\hat{H}$ resta associata $\hat{K} = 0$: ciò segue immediatamente dall'espressione (4.32) di $K$, con $c = 1$ e $K_0$ dato dalla (4.31), nel nostro caso $K_0 = -H \circ \mathbf{w}$; ma ciò era evidente a priori, perché in accordo con la (4.46) l'immagine dei moti di $H$ è data dai particolarissimi moti $\tilde{p} = \text{cost}$, $\tilde{q} = \text{cost}$, che appunto corrispondono all'hamiltoniana $\hat{K} = 0$. Si riveda, alla luce di questa osservazione, l'esempio vi) del paragrafo 4.2.1. Il fatto poi che ogni flusso hamiltoniano si possa interpretare come una trasformazione canonica, porta a pensare che alle equazioni di Hamilton si possa equivalentemente sostituire una equazione per la funzione generatrice di questa trasformazione canonica (equazione di Hamilton–Jacobi). Questo punto di vista sarà ripreso più avanti.

### 4.2.5 Trasformazioni canoniche e parentesi di Poisson

#### a) Trasformazioni che preservano le parentesi di Poisson

Come si è fatto notare nel capitolo precedente, la definizione di parentesi di Poisson di due variabili dinamiche $\{f, g\}$ fa esplicito riferimento a un particolare sistema di coordinate. Si pone pertanto il problema di caratterizzare la classe di cambiamenti di coordinate nello spazio delle fasi, rispetto alle quali la definizione di parentesi di Poisson è invariante.

Cominciamo allora col dare la seguente precisa

**Definizione 4.3:** *Diciamo che la trasformazione di coordinate*

$$(\mathbf{p}, \mathbf{q}) = \mathbf{w}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) \tag{4.49}$$

*preserva le parentesi di Poisson, se, comunque si prendano due funzioni $f$ e $g$, indicando con $F = f \circ \mathbf{w}$, $G = g \circ \mathbf{w}$ le funzioni trasformate, cioè $F(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = f(\mathbf{w}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t), t)$, $G(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = g(\mathbf{w}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t), t)$, risulta $\{F, G\} = \{f, g\} \circ \mathbf{w}$, ovvero*

$$\{F, G\}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t) = \{f, g\}(\mathbf{w}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, t), t) \ .$$

Sottintendendo il cambiamento di variabili, si usa a volte la scrittura breve $\{F, G\} = \{f, g\}$. E' evidente, sulla base della sola definizione, che le trasformazioni che preservano le parentesi di Poisson costituiscono un gruppo.

Il gruppo si può allargare ammettendo che le parentesi di Poisson siano preservate a meno di una costante moltiplicativa $c$, cioè che risulti

$$c\,\{F, G\} = \{f, g\} \circ \mathbf{w} \ ; \tag{4.50}$$

in questo modo si comprendono, in particolare, le dilatazioni, definite da $p_h = \alpha_h \tilde{p}_h$, $q_h = \beta_h \tilde{q}_h$, con $\alpha_h \beta_h = c$ ($c$ indipendente da $h$), $h = 1, \ldots, n$.

Se una trasformazione preservi o meno le parentesi di Poisson è cosa che si verifica con facilità, nonostante la definizione faccia riferimento a infinite funzioni: si vede infatti immediatamente che risulta

$$\begin{aligned}\{F,G\} = &\sum_{h,k=1}^{n} \frac{\partial f}{\partial p_h}\frac{\partial g}{\partial p_k}\{u_h,u_k\} + \sum_{h,k=1}^{n} \frac{\partial f}{\partial p_h}\frac{\partial g}{\partial q_k}\{u_h,v_k\} \\ &+ \sum_{h,k=1}^{n} \frac{\partial f}{\partial q_h}\frac{\partial g}{\partial p_k}\{v_h,u_k\} + \sum_{h,k=1}^{n} \frac{\partial f}{\partial q_h}\frac{\partial g}{\partial q_k}\{v_h,v_k\} \ ;\end{aligned}$$

è evidente allora che tutte le parentesi di Poisson sono preservate, se sono preservate le parentesi di Poisson fondamentali, cioè se risulta, per $h,k = 1,\ldots,n$,

$$\{u_h,u_k\} = 0 \ , \quad \{v_h,v_k\} = 0 \ , \quad \{u_h,v_k\} = -\delta_{hk} \ .$$

- **Esercizio 4.9:** Si verifichi che tutte le trasformazioni i) – vii) del paragrafo 4.2.1 preservano le parentesi di Poisson (eventualmente a meno della costante $c$).

Ci proponiamo di dimostrare che le trasformazioni che preservano le parentesi di Poisson (eventualmente a meno di una costante moltiplicativa) coincidono con le trasformazioni canoniche.[25) ] Premettiamo tuttavia un breve paragrafo, ove introduciamo per il formalismo canonico una notazione più compatta e un linguaggio appropriato.

*b) Notazione compatta per il formalismo canonico.*

E' utile introdurre per il formalismo canonico la seguente notazione compatta: poniamo $\mathbf{x} = (x_1,\ldots,x_{2n}) = (p_1,\ldots,p_n,q_1,\ldots,q_n)$, e introduciamo la matrice antisimmetrica

$$E = \begin{pmatrix} O & -I \\ I & O \end{pmatrix} \ ,$$

ove $O$ e $I$ sono rispettivamente la matrice nulla e la matrice identità $n \times n$; si verifica immediatamente che risulta $E^T = E^{-1} = -E$. Infine, per una generica funzione $f$ indichiamo con $\partial_x f$ il suo gradiente, $\partial_x f = (\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_{2n}})$. Con queste notazioni, è immediato verificare che le $2n$ equazioni di Hamilton prendono la forma compatta

$$\dot{\mathbf{x}} = E\,\partial_x H \ ,$$

mentre le parentesi di Poisson si scrivono

$$\{f,g\} = \partial_x f \cdot E\,\partial_x g \ .$$

Presi due vettori $\mathbf{a}$, $\mathbf{b} \in \mathbf{R}^{2n}$, l'espressione $\mathbf{a} \cdot E\,\mathbf{b}$ è una operazione bilineare antisimmetrica, chiamata *prodotto scalare simplettico;* la parentesi di Poisson $\{f,g\}$ si presenta

[25)] Poiché, come già si è osservato, queste trasformazioni costituiscono un gruppo, si verifica ancora una volta che le trasformazioni canoniche formano un gruppo.

dunque come il prodotto scalare simplettico tra i gradienti $\partial_x f$ e $\partial_x g$. Analogamente, data una qualsiasi funzione $f(\mathbf{x})$, il campo vettoriale $\mathbf{f}(\mathbf{x}) = E\partial_x f = \{\mathbf{x}, f\}$ prende il nome di *gradiente simplettico* di $f$[26)]

Consideriamo allora una trasformazione di coordinate (4.49), e sia $J$ la corrispondente matrice jacobiana, di elementi $J_{ij} = \frac{\partial w_i}{\partial \tilde{x}_j}$, o in altra notazione

$$J = \begin{pmatrix} \frac{\partial \mathbf{u}}{\partial \tilde{\mathbf{p}}} & \frac{\partial \mathbf{u}}{\partial \tilde{\mathbf{q}}} \\ \frac{\partial \mathbf{v}}{\partial \tilde{\mathbf{p}}} & \frac{\partial \mathbf{v}}{\partial \tilde{\mathbf{q}}} \end{pmatrix} .$$

Si dimostra molto facilmente la seguente

**Proposizione 4.11:** *La trasformazione di coordinate (4.49) preserva le parentesi di Poisson, se e solo se la corrispondente matrice jacobiana $J$ soddisfa la relazione*

$$J\,E\,J^T = E \ . \tag{4.51}$$

**Dimostrazione.** La conservazione delle parentesi di Poisson si scrive, con la notazione sopra introdotta,

$$\partial_{\tilde{x}} F \cdot E\,\partial_{\tilde{x}} G = (\partial_x f \cdot E\,\partial_x g) \circ \mathbf{w} \ ,$$

o anche, eseguendo a sinistra il cambio di variabili da $\tilde{\mathbf{x}}$ a $\mathbf{x}$, e osservando che $\partial_{\tilde{x}} F = J^T \partial_x f \circ \mathbf{w}$, $\partial_{\tilde{x}} G = J^T \partial_x g \circ \mathbf{w}$,

$$(\partial_x f \cdot J\,E\,J^T \partial_x g) \circ \mathbf{w} = (\partial_x f \cdot E\,\partial_x g) \circ \mathbf{w} \ ;$$

eliminando l'ormai inessenziale composizione con $\mathbf{w}$, si vede allora che questa uguaglianza è soddisfatta per ogni di $f$ e $g$, se e solo se si ha $JEJ^T = E$. *Q.E.D.*

Matrici che soddisfano la proprietà (4.51) si dicono *simplettiche;* è immediato verificare che esse formano un gruppo,[27)] che $E$ stessa è simplettica, e che se $J$ è simplettica lo è anche $J^T$.[28)] Dalla (4.51) si deduce immediatamente $(\det J)^2 = 1$, e dunque $\det J = \pm 1$; in realtà si può dimostrare che si ha sempre $\det J = 1$ (anzi, si verifica immediatamente che per il caso particolare delle matrici $2 \times 2$ la proprietà di simpletticità coincide con la proprietà di avere determinante uguale a uno). E' infine immediato verificare che se le parentesi di Poisson sono preservate a meno di una costante $c$, cioè si ha $\{f, g\} = c\{F, G\}$, allora la matrice $J$ soddisfa la relazione $JEJ^T = c^{-1}E$, e viceversa; in questo caso si ha $\det J = c^{-n}$.

---

26) Come è noto, un campo vettoriale $\mathbf{f}$ (definito in una dominio semplicemente connesso) è il gradiente di una funzione, se e solo se è verificata la condizione di chiusura $\partial f_i/\partial x_j = \partial f_j/\partial x_i$ $\forall i, j$. Per il gradiente simplettico sussiste una condizione analoga, nella quale la derivata rispetto alla coordinata $x_i$ è sostituita dalla parentesi di Poisson con $x_i$; si veda il Lemma 1 in appendice.

27) Infatti, la matrice identità è evidentemente simplettica; $J^{-1}$ è simplettica, come segue da $J^{-1}E(J^{-1})^T = J^{-1}(JEJ^T)(J^{-1})^T = E$; infine, se $J$ e $K$ sono simplettiche, allora $JK$ è simplettica, come segue da $(JK)E(JK)^T = JKEK^TJ^T = JEJ^T = E$.

28) Infatti, da $J^{-1}E(J^{-1})^T = E$ segue, per inversione, $J^TE^{-1}J = E^{-1}$, e dunque (ricordando $E^{-1} = -E$) $J^TEJ = E$.

*c) Trasformazioni canoniche e parentesi di Poisson.*

Ci proponiamo finalmente di dimostrare che le trasformazioni che preservano le parentesi di Poisson coincidono con le trasformazioni canoniche. A questo scopo basta dimostrare la seguente

**Proposizione 4.12:** *La trasformazione di coordinate (4.25) preserva le parentesi di Poisson, a meno della costante $c$, se e solo se la condizione di Lie (4.30) è soddisfatta con la medesima costante $c$, e una opportuna funzione $F(\tilde{\mathbf{p}},\tilde{\mathbf{q}},t)$.*

**Dimostrazione.** Limitiamoci, per semplicità, al caso $c=1$. La condizione di Lie (4.30) si può anche scrivere nella forma completamente equivalente

$$\mathbf{u}\cdot \overset{*}{\mathrm{d}}\mathbf{v} - \mathbf{v}\cdot \overset{*}{\mathrm{d}}\mathbf{u} = \tilde{\mathbf{p}}\cdot \overset{*}{\mathrm{d}}\tilde{\mathbf{q}} - \tilde{\mathbf{q}}\cdot \overset{*}{\mathrm{d}}\tilde{\mathbf{p}} + \overset{*}{\mathrm{d}}G \ , \qquad G = 2F - \mathbf{u}\cdot\mathbf{v} + \tilde{\mathbf{p}}\cdot\tilde{\mathbf{q}} \ ,$$

che si ottiene sommando alla (4.30) la relazione equivalente $\overset{*}{\mathrm{d}}(\mathbf{u}\cdot\mathbf{v}) - \mathbf{v}\cdot\overset{*}{\mathrm{d}}\mathbf{u} = \overset{*}{\mathrm{d}}(\tilde{\mathbf{p}}\cdot\tilde{\mathbf{q}}) - \tilde{\mathbf{q}}\cdot\overset{*}{\mathrm{d}}\tilde{\mathbf{p}} + \overset{*}{\mathrm{d}}F$; tale forma, più simmetrica, si presta alla notazione compatta

$$E\mathbf{w}\cdot\overset{*}{\mathrm{d}}\mathbf{w} - E\tilde{\mathbf{x}}\cdot\mathrm{d}\tilde{\mathbf{x}} = \overset{*}{\mathrm{d}}G \ .$$

Ricordando che si ha $\overset{*}{\mathrm{d}}\mathbf{w} = J\,\overset{*}{\mathrm{d}}\tilde{\mathbf{x}}$, si ottiene allora la relazione, ancora equivalente alla condizione di Lie, $(J^T E\mathbf{w} - E\tilde{\mathbf{x}})\cdot\overset{*}{\mathrm{d}}\tilde{\mathbf{x}} = \overset{*}{\mathrm{d}}G$; di conseguenza, $G$ (e dunque $F$) esiste, se e solo se il campo vettoriale

$$\mathbf{f} = J^T E\,\mathbf{w}(\tilde{\mathbf{x}}) - E\tilde{\mathbf{x}}$$

è il gradiente di una funzione. A sua volta, $\mathbf{f}$ esiste (in un dominio semplicemente connesso, come abbiamo sempre assunto) se, e solo se, è verificata la condizione di chiusura $\frac{\partial f_i}{\partial\tilde{x}_j} - \frac{\partial f_j}{\partial\tilde{x}_i} = 0\ \forall i,j$. Ora, non è difficile, anche se un po' laborioso, verificare che si ha

$$\frac{\partial f_i}{\partial\tilde{x}_j} - \frac{\partial f_j}{\partial\tilde{x}_i} = 2\big(J^T E J - E\big)_{ij}$$

(non si dimentichi che anche $J$ dipende da $\tilde{\mathbf{x}}$); pertanto, $G$ esiste se, e solo se, la matrice $J$ è simplettica, cioè se la (4.25) preserva le parentesi di Poisson. *Q.E.D.*

La proposizione ora dimostrata, assieme alla 4.7, porta a concludere che una trasformazione è canonica, se e solo se preserva le parentesi di Poisson; più precisamente, vale la seguente

**Proposizione 4.13:** *Se la trasformazione di coordinate*

$$\mathbf{x} = \mathbf{w}(\tilde{\mathbf{x}},t) \tag{4.52}$$

*preserva le parentesi di Poisson, a meno di una costante $c$, allora la trasformazione è canonica di valenza $c$, e muta la generica hamiltoniana $H(\mathbf{x},t)$ in una nuova hamiltoniana $K$ della forma*

$$K(\tilde{\mathbf{x}},t) = c\,H(\mathbf{w}(\tilde{\mathbf{x}},t),t) + K_0(\tilde{\mathbf{x}},t) \ , \tag{4.53}$$

*con opportuna $K_0$ indipendente da $H$. Viceversa, se la trasformazione (4.52) è canonica di valenza $c$, allora essa preserva le parentesi di Poisson a meno della costante $c$, ed esiste $K_0$ tale che per ogni hamiltoniana $H$ la hamiltoniana coniugata $K$ è data dalla (4.53).*

Questa proposizione è una evidente conseguenza della Proposizione 4.7, attraverso la Proposizione 4.12. Poichè però, come si ricorderà, della 4.7 abbiamo dimostrato solo la prima parte, una dimostrazione diretta della Proposizione 4.13 è comunque interessante; anzi, poichè a sua volta, come si vede immediatamente, la Proposizione 4.13 implica la 4.7, tale dimostrazione diretta è utile a colmare la lacuna allora lasciata. La dimostrazione diretta della Proposizione 4.13, elementare ma un po' laboriosa, è riportata in Appendice.

Dalla Proposizione 4.12, ricordando quanto si è detto a proposito del determinante delle matrici simplettiche, segue immediatamente il seguente

**Corollario 4.2:** *Le trasformazioni canoniche univalenti, e tra esse il flusso hamiltoniano, preservano il volume euclideo $\mathrm{d}p_1 \cdots \mathrm{d}p_n \, \mathrm{d}q_1 \cdots \mathrm{d}q_n$.*

- **Osservazione.** Il fatto che il flusso hamiltoniano preservi il volume dello spazio delle fasi è un risultato importante, noto in letteratura col nome di *teorema di Liouville.* Qui abbiamo dedotto questo teorema nel corso del nostro studio sulle trasformazioni canoniche; più avanti se ne vedrà una dimostrazione molto più semplice e diretta.

## 4.3 L'equazione di Hamilton –Jacobi

### 4.3.1 Sistemi hamiltoniani "integrabili"

Consideriamo un sistema hamiltoniano autonomo a $n$ gradi di libertà, di hamiltoniana $H(\mathbf{p},\mathbf{q})$; come abbiamo già avuto modo di osservare, se $H$ non dipende da una o più coordinate, ad esempio $q_{m+1},\ldots,q_n$, $0 \le m < n$, allora i corrispondenti momenti coniugati $p_{m+1},\ldots,p_n$ sono integrali del moto del sistema, cosicché in pratica le $2n$ equazioni canoniche si riducono a $2m$ equazioni per le sole coordinate $p_1,\ldots,p_m$, $q_1,\ldots,q_m$, nelle quali i momenti $p_{m+1},\ldots,p_n$ giocano il ruolo di semplici parametri. Se queste si sanno integrare, allora si trovano facilmente anche $q_{m+1}(t),\ldots,q_n(t)$, con una quadratura (cioè con una integrazione ordinaria): si ha infatti, per $m < h \le n$,

$$q_h(t) = \int_0^t \dot q_h(t)\,\mathrm{d}t\,, \quad \dot q_h(t) \;=\; \frac{\partial H}{\partial p_h}\big(p_1(t),\ldots,p_m(t),p_{m+1},\ldots,p_n,q_1(t),\ldots,q_m(t)\big)\,.$$

Se poi $H$ è indipendente da tutte le coordinate $q_1,\ldots,q_n$, cioè si ha $H(\mathbf{p},\mathbf{q}) = E(\mathbf{p})$, allora l'integrazione delle equazioni del moto è del tutto banale: posto infatti $\omega_h(\mathbf{p}) = \frac{\partial E}{\partial p_h}$, $h = 1,\ldots,n$, le equazioni del moto divengono $\dot p_h = 0$, $\dot q_h = \omega_h(\mathbf{p})$, e quindi l'integrale generale si scrive

$$p_h(t) = p_h^\circ \ , \qquad q_h(t) = q_h^\circ + \omega_h(\mathbf{p}^\circ)t \ , \qquad h = 1, \ldots, n \ , \tag{4.54}$$

dove $p_h^\circ$, $q_h^\circ$, $h = 1, \ldots, n$, indicano i dati iniziali. Le coordinate assenti dall'hamiltoniana sono chiamate *variabili cicliche.*[29]

Data ora una hamiltoniana $H(\mathbf{p}, \mathbf{q})$, dipendente in generale da tutte le variabili, è chiaramente interessante studiare la possibilità che tutte le $n$ coordinate diventino cicliche a seguito di una opportuna trasformazione canonica. E' questo il caso, ad esempio, di un sistema di $n$ oscillatori armonici indipendenti, di pulsazioni $\omega_1, \ldots, \omega_n$. L'hamiltoniana del sistema si scrive $H(\mathbf{p}, \mathbf{q}) = \frac{1}{2} \sum_h (p_h^2 + \omega_h^2 q_h^2)$, e dunque dipende da tutte le coordinate canoniche. Tuttavia, con una ovvia generalizzazione dell'esempio vii) del paragrafo 4.2.1, è facile introdurre una trasformazione canonica, tale che la nuova hamiltoniana $K$ dipenda dai soli (nuovi) momenti: usando per le nuove coordinate canoniche la notazione, tipica in questo genere di problemi, $I_1, \ldots, I_n$, $\varphi_1, \ldots, \varphi_n$, si vede che la trasformazione

$$p_h = \sqrt{2\omega_h I_h} \, \cos\varphi_h \ , \qquad q_h = \sqrt{2I_h/\omega_h} \, \sin\varphi_h \tag{4.55}$$

è canonica, e muta $H$ in

$$K(I_1, \ldots, I_n, \varphi_1, \ldots, \varphi_n) \ = \ \sum_{h=1}^n \omega_h I_h \ .$$

La verifica della canonicità, lasciata per esercizio, è una facile generalizzazione della verifica fatta nel paragrafo 4.2.1 a proposito dell'esempio vii); è molto semplice anche vedere che la (4.55) preserva le parentesi di Poisson. Nelle nuove variabili l'integrazione delle equazioni del moto dà $I_h(t) = I_h^\circ$, $\varphi_h(t) = \varphi_h^\circ + \omega_h t$; le equazioni (4.55) della trasformazione forniscono poi il moto nelle variabili originarie, precisamente

$$p_h(t) \ = \ \omega_h A_h \cos(\varphi_h^\circ + \omega_h t) \ , \quad q_h(t) \ = \ A_h \sin(\varphi_h^\circ + \omega_h t) \ , \qquad A_h \equiv \sqrt{2I_h^\circ/\omega_h} \ .$$

Sistemi hamiltoniani per i quali esista una trasformazione canonica (che possiamo supporre univalente) $\mathbf{p} = \mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$, $\mathbf{q} = \mathbf{v}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$, tale che la nuova hamiltoniana $K$ sia indipendente da $\tilde{\mathbf{q}}$, si dicono *integrabili.* E' importante però, a questo proposito, introdurre una distinzione alquanto significativa: precisamente, se il cambio di coordinate è definito solo localmente, in un generico aperto $U$, diremo che il sistema è *localmente integrabile* in $U$; vedremo più avanti che tutti i sistemi hamiltoniani sono localmente integrabili, intorno a un qualsiasi punto che non sia di equilibrio, e che corrispondentemente l'integrabilità locale è una proprietà abbastanza povera. Diremo invece che il sistema è *globalmente integrabile* (o semplicemente integrabile) quando l'aperto $U$ in cui il cambio di coordinate è definito è *invariante* (ovvero "contiene intere orbite"; si veda il Capitolo 1). Ciò avviene, innanzitutto, se l'insieme $U$ dove è definito il cambiamento di variabili è determinato da una condizione del tipo $E_0 < K(\tilde{\mathbf{p}}) < E_1$,

---

[29] La ragione di questa terminologia sta nel fatto che, tipicamente, queste variabili sono angoli, che come mostrano le (4.54) avanzano (ruotano) con velocità angolare costante.

o più in generale da una condizione sulle sole $\tilde{\mathbf{p}}$, della forma $\tilde{\mathbf{p}} \in U_0 \subset \mathbb{R}^n$ (mentre $\tilde{\mathbf{q}}$ resta arbitrario); ad esempio il sistema di oscillatori armonici sopra considerato è integrabile globalmente in un qualunque insieme di questo tipo (purché in $U_0$ nessuna delle $I_h$ si annulli, altrimenti il cambio di coordinate è mal definito, non essendo differenziabile). L'integrabilità globale è una proprietà estremamente interessante, molto forte (si pensi al fatto che vi è coinvolto il comportamento asintotico delle orbite per $t \to \pm\infty$), e in un certo senso – difficile però da precisare in queste note – eccezionale; ciò si è cominciato a capire soprattutto dagli studi di Poincaré in poi ($\simeq 1900$), ed è diventato più chiaro di recente, a partire dagli anni '60. L'argomento è abbastanza difficile e delicato, e molti problemi, non privi di profondo significato fisico, sono ancora sostanzialmente aperti, nonostante la ricerca in questo campo sia oggi particolarmente attiva. Qualche breve osservazione su questo punto sarà fatta più avanti.

### 4.3.2 L'equazione di Hamilton–Jacobi

Da quanto abbiamo detto nel paragrafo precedente, si pone in modo naturale il problema della ricerca di una funzione generatrice, ad esempio di tipo $F_2(\tilde{\mathbf{p}}, \mathbf{q})$, che generi la trasformazione di coordinate $\mathbf{p} = \mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$, $\mathbf{q} = \mathbf{v}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$, capace di eliminare dalla nuova hamiltoniana le coordinate $\tilde{q}_1, \ldots, \tilde{q}_n$.

Consideriamo a questo scopo, in corrispondenza a una assegnata hamiltoniana $H$, la seguente equazione alle derivate parziali, detta *equazione ridotta di Hamilton–Jacobi:*

$$H\Big(\frac{\partial W}{\partial q_1}, \ldots, \frac{\partial W}{\partial q_n}, q_1, \ldots, q_n\Big) = E \ , \tag{4.56}$$

nella quale sia la funzione $W(q_1, \ldots, q_n)$, sia la costante $E$ sono incognite.[30] Ad esempio, in corrispondenza all'hamiltoniana $H$ dell'oscillatore armonico, $H = \frac{1}{2}(p^2 + \omega^2 q^2)$, l'equazione si scrive

$$\frac{1}{2}\Big(\frac{\partial W}{\partial q}\Big)^2 + \frac{1}{2}\omega^2 q^2 = E \ . \tag{4.57}$$

Si dice che l'equazione ammette un *integrale completo*[31] se esiste una famiglia di soluzioni $W_{\alpha_1,\ldots,\alpha_n}(q_1, \ldots, q_n)$, $E_{\alpha_1,\ldots,\alpha_n}$, dipendente da $n$ parametri reali arbitrari $\alpha_1, \ldots, \alpha_n$, tale che sia[32]

$$\det\Big(\frac{\partial^2 W}{\partial \alpha_h \partial q_k}\Big) \neq 0 \ . \tag{4.58}$$

La connessione tra l'equazione ridotta di Hamilton–Jacobi ora introdotta e il problema che ci siamo posti è espressa dalla seguente facile

[30] La presenza di una costante libera nell'equazione ricorda, seppur vagamente (l'equazione non è lineare) il problema agli autovalori.

[31] L'integrale completo non va confuso con l'integrale generale (che per le equazioni alle derivate parziali contiene sempre una funzione arbitraria, anziché un numero finito di costanti arbitrarie).

[32] Questa proprietà di non degenerazione assicura che le derivate di $W_{\alpha_1,\ldots,\alpha_n}$ rispetto alle $q_h$, che compaiono nell'equazione, effettivamente dipendono dalle $\alpha_h$. I parametri $\alpha_1, \ldots, \alpha_n$ non possono essere, ad esempio, costanti additive.

**Proposizione 4.14:** *Se $W_{\alpha_1,\ldots,\alpha_n}$, $E_{\alpha_1,\ldots,\alpha_n}$ costituiscono un integrale completo per l'equazione ridotta di Hamilton–Jacobi (4.56), relativa a una assegnata hamiltoniana $H$, allora la funzione $F_2$ definita da*

$$F_2(\tilde{p}_1,\ldots,\tilde{p}_n,q_1,\ldots,q_n) \;=\; W_{\tilde{p}_1,\ldots,\tilde{p}_n}(q_1,\ldots,q_n)$$

*genera una trasformazione canonica che muta $H$ in*

$$K(\tilde{p}_1,\ldots,\tilde{p}_n,\tilde{q}_1,\ldots,\tilde{q}_n) = E_{\tilde{p}_1,\ldots,\tilde{p}_n} \; . \tag{4.59}$$

**Dimostrazione.** In base alle (4.44) e (4.45), poiché $F_2$ non dipende dal tempo, si ha (in variabili miste)

$$K(\tilde{p}_1,\ldots,\tilde{p}_n,q_1,\ldots,q_n) = H\Big(\frac{\partial F_2}{\partial q_1},\ldots,\frac{\partial F_2}{\partial q_n},q_1,\ldots,q_n\Big) \; ;$$

dall'equazione di Hamilton–Jacobi (4.56) segue allora la (4.59). La condizione di invertibilità (4.43) è garantita dalla (4.58). *Q.E.D.*

Vediamo così che, se esiste un integrale completo della (4.56), allora esiste un cambiamento di variabili $\mathbf{p} = \mathbf{u}(\tilde{\mathbf{p}},\tilde{\mathbf{q}})$, $\mathbf{q} = \mathbf{v}(\tilde{\mathbf{p}},\tilde{\mathbf{q}})$, implicitamente definito dalle (4.44), a seguito del quale la soluzione delle equazioni del moto assume la forma speciale

$$\tilde{\mathbf{p}}(t) = \tilde{\mathbf{p}}^\circ \; , \qquad \tilde{\mathbf{q}}(t) = \tilde{\mathbf{q}}^\circ + \boldsymbol{\omega}(\tilde{\mathbf{p}}^\circ)t \; ; \tag{4.60}$$

corrispondentemente si ha, nelle vecchie variabili, l'andamento (anch'esso assai particolare)

$$\mathbf{p}(t) = \mathbf{u}\big(\tilde{\mathbf{p}}^\circ,\tilde{\mathbf{q}}^\circ + \boldsymbol{\omega}(\tilde{\mathbf{p}}^\circ)\,t\big) \; , \qquad \mathbf{q}(t) = \mathbf{v}\big(\tilde{\mathbf{p}}^\circ,\tilde{\mathbf{q}}^\circ + \boldsymbol{\omega}(\tilde{\mathbf{p}}^\circ)\,t\big) \; . \tag{4.61}$$

Le (4.60) rappresentano, sostanzialmente, un moto libero con velocità costante; le (4.61) mostrano lo stesso moto attraverso la trasformazione canonica. Si osservi che il dominio ove variano i parametri $\alpha_1,\ldots,\alpha_n$ dell'integrale completo determina il dominio di definizione della trasformazione canonica, per la parte riguardante le variabili $\tilde{p}_1,\ldots,\tilde{p}_n$.

Una formulazione equivalente della Proposizione 4.14 è la seguente

**Proposizione 4.14':** *Nelle stesse ipotesi della Proposizione 4.14, la funzione*

$$F_2(\tilde{p}_1,\ldots,\tilde{p}_n,q_1,\ldots,q_n,t) \;=\; W_{\tilde{p}_1,\ldots,\tilde{p}_n}(q_1,\ldots,q_n) - t\,E_{\tilde{p}_1,\ldots,\tilde{p}_n}$$

*genera una trasformazione canonica dipendente dal tempo, che coniuga l'hamiltoniana $H$ all'hamiltoniana $K=0$.*

**Dimostrazione.** Come per la Proposizione 4.14, è sufficiente usare le (4.44) e (4.45), ricordando che questa volta $F_2$ dipende anche da $t$. *Q.E.D.*

- **Esempio.** Consideriamo il caso particolare dell'oscillatore armonico. L'equazione ridotta di Hamilton–Jacobi ha la forma (4.57), cosicché si ha $\frac{\partial W}{\partial q} = \pm\sqrt{2E-\omega^2q^2}$.

L'equazione si risolve allora con una quadratura,[33] precisamente si trova, ad esempio in corrispondenza al segno $+$ (ed a meno di una inessenziale costante additiva), $W = \int_0^q \sqrt{2E - \omega^2 x^2}\,\mathrm{d}x$. Un integrale completo dell'equazione si trova allora scegliendo a piacere la dipendenza di $E$ da $\alpha$. Ad esempio, si può porre $E_\alpha = \alpha$, e corrispondentemente $W_\alpha = \int_0^q \sqrt{2\alpha - \omega^2 x^2}\,\mathrm{d}x$; la scelta più comune, che conduce alle nuove variabili $(I, \varphi)$ sopra introdotte, è però[34] $E_\alpha = \omega\alpha$: con questa posizione si ottiene infatti l'integrale completo nella forma

$$W_\alpha(q) = \int_0^q \sqrt{2\omega\alpha - \omega^2 x^2}\,\mathrm{d}x \ , \qquad E_\alpha = \omega\alpha \ ,$$

e dunque la funzione generatrice si scrive $F_2(I, q) = \int_0^q \sqrt{2\omega I - \omega^2 x^2}\,\mathrm{d}x$. La trasformazione canonica è allora definita implicitamente da

$$\begin{aligned} p &= \frac{\partial F_2}{\partial q} = \sqrt{2\omega I - \omega^2 q^2} \\ \varphi &= \frac{\partial F_2}{\partial I} = \int_0^q \frac{\mathrm{d}x}{\sqrt{2I/\omega - x^2}} = \arcsin \sqrt{\omega/2I}\, q \ , \qquad -\frac{\pi}{2} < \varphi < \frac{\pi}{2} \ , \end{aligned}$$

e per inversione si ottengono facilmente le equazioni esplicite

$$q = \sqrt{2I/\omega}\, \sin\varphi \ , \qquad p = \sqrt{2\omega I}\, \cos\varphi \ , \tag{4.62}$$

limitate, per il momento, a $-\frac{\pi}{2} < \varphi < \frac{\pi}{2}$. Non sarebbe difficile vedere che, in corrispondenza al segno $-$, si ottengono ancora queste equazioni, con determinazione dell'angolo $\varphi$ tra $\frac{1}{2}\pi$ e $\frac{3}{2}\pi$; in definitiva, le (4.62) forniscono, per ogni valore di $\varphi$, la trasformazione canonica cercata. L'hamiltoniana dell'oscillatore armonico è mutata da questa trasformazione nell'hamiltoniana $K(I, \varphi) = E_I = \omega I$; le corrispondenti equazioni del moto $\dot I = 0$ e $\dot\varphi = \omega$ sono risolte da $I = I_0$, $\varphi = \varphi_0 + \omega t$, e infine le equazioni della trasformazione forniscono l'integrale generale nelle vecchie variabili:

$$p = \sqrt{2\omega I_0}\, \cos(\omega t + \varphi_0) \ , \qquad q = \sqrt{2I_0/\omega}\, \sin(\omega t + \varphi_0) \ , \tag{4.63}$$

con $I_0$, $\varphi_0$ costanti arbitrarie.

In maniera del tutto equivalente possiamo utilizzare la funzione generatrice dipendente dal tempo $F_2(I, q) = W_I(q) - E_I t = \int_0^q \sqrt{2\omega I - \omega^2 x^2}\,\mathrm{d}x - \omega I\, t$; in questo caso la trasformazione canonica resta definita da

$$\begin{aligned} p &= \frac{\partial F_2}{\partial q} = \sqrt{2\omega I - \omega^2 q^2} \\ \varphi &= \frac{\partial F_2}{\partial I} = \int_0^q \frac{\mathrm{d}x}{\sqrt{2I/\omega - x^2}} - \omega t = \arcsin \sqrt{\omega/2I}\, q - \omega\, t \ , \end{aligned}$$

e per inversione si ottiene senza difficoltà $p = \sqrt{2\omega I}\, \cos(\omega t + \varphi)$, $q = \sqrt{2I/\omega}\, \sin(\omega t + \varphi)$. La nuova hamiltoniana è ora $K(I, \varphi) = 0$, pertanto le equazioni del moto $\dot I = 0$, $\dot\varphi = 0$ sono risolte da $I = I_0$, $\varphi = \varphi_0$; per sostituzione si riottiene l'integrale generale nell'identica forma (4.63).

---

33) E' utile rivedere, nel primo capitolo, il paragrafo sull'analisi quantitativa delle soluzioni per sistemi conservativi a un solo grado di libertà.

34) Si osservi che in questo modo il parametro $\alpha$ ha le dimensioni di una azione.

- **Osservazione.** E' del tutto evidente che questo metodo di integrazione, qui illustrato in dettaglio per l'oscillatore armonico, si applica in realtà a qualunque sistema a un grado di libertà; gli integrali che compaiono nelle equazioni della trasformazione non saranno in generale elementari, tuttavia è chiaro che in questo modo la soluzione di qualunque problema a un grado di libertà si riporta (come peraltro già sapevamo) a una quadratura.

Alla Proposizione 4.14', del tutto equivalente alla 4.14 quando si consideri una hamiltoniana indipendente dal tempo, si può dare una formulazione più generale, adatta anche al caso di dipendenza esplicita di $H$ dal tempo. A questo scopo consideriamo la seguente equazione, detta *equazione di Hamilton–Jacobi:*

$$H\Big(\frac{\partial S}{\partial q_1},\ldots,\frac{\partial S}{\partial q_n},q_1,\ldots,q_n,t\Big)+\frac{\partial S}{\partial t} \;=\; 0\;; \tag{4.64}$$

in analogia a quanto visto sopra, chiameremo *integrale completo* della (4.64) ogni famiglia di soluzioni $S_{\alpha_1,\ldots,\alpha_n}(q_1,\ldots,q_n,t)$, dipendente da $n$ parametri arbitrari $\alpha_1,\ldots,\alpha_n$, tale che sia

$$\det\Big(\frac{\partial^2 S}{\partial\alpha_h\partial q_k}\Big)\;\neq\;0\,. \tag{4.65}$$

E' di immediata verifica la seguente

**Proposizione 4.15:** *Se $S_{\alpha_1,\ldots,\alpha_n}(q_1,\ldots,q_n,t)$ è un integrale completo dell'equazione di Hamilton–Jacobi (4.64), relativa a una assegnata hamiltoniana $H$, allora la funzione*

$$F_2(\tilde p_1,\ldots,\tilde p_n,q_1,\ldots,q_n,t)\;=\;S_{\tilde p_1,\ldots,\tilde p_n}(q_1,\ldots,q_n,t)$$

*genera una trasformazione canonica che coniuga $H$ a $K$ identicamente nulla.*

**Dimostrazione.** Ancora una volta, basta usare le (4.44) (4.45); la (4.65) garantisce ancora la condizione di invertibilità (4.43). *Q.E.D.*

Le equazioni canoniche corrispondenti a $K=0$ sono evidentemente risolte da $\tilde p_h(t)=\tilde p_h^\circ$, $\tilde q_h(t)=\tilde q_h^\circ$. Le equazioni della trasformazione forniscono allora l'integrale generale delle equazioni canoniche relative all'hamiltoniana $H$: infatti, per la canonicità della trasformazione, le

$$\begin{aligned} p_h(t) &= u_h(\tilde p_1^\circ,\ldots,\tilde p_n^\circ,\tilde q_1^\circ,\ldots,\tilde q_n^\circ,t)\\ q_h(t) &= v_h(\tilde p_1^\circ,\ldots,\tilde p_n^\circ,\tilde q_1^\circ,\ldots,\tilde q_n^\circ,t)\end{aligned}\qquad h=1,\ldots,n\,,$$

sono soluzioni per ogni scelta delle $2n$ costanti arbitrarie $\tilde p_1^\circ,\ldots,\tilde p_n^\circ,\tilde q_1^\circ,\ldots,\tilde q_n^\circ$. Per questo, quanto abbiamo visto è anche chiamato *metodo di integrazione di Hamilton–Jacobi.*

- **Osservazione.** Nel caso in cui $H$ non dipenda esplicitamente dal tempo, è chiaro che ogni integrale completo della (4.56) fornisce un integrale completo della (4.64): basta infatti porre $S_{\alpha_1,\ldots,\alpha_n}(q_1,\ldots,q_n,t)\;=\;W_{\alpha_1,\ldots,\alpha_n}(q_1,\ldots,q_n)\,-$

$t\, E_{\alpha_1,\ldots,\alpha_n}$. Procedendo in direzione opposta, si vede bene che dalla (4.64) si giunge alla (4.56) cercando soluzioni a variabili separate, del tipo $S(q_1,\ldots,q_n,t) = W(q_1,\ldots,q_n) - E\,t$.

### 4.3.3 *Soluzioni locali e globali dell'equazione di Hamilton–Jacobi (cenno)*

E' naturale a questo punto chiedersi che cosa si può dire sulla possibilità di risolvere l'equazione di Hamilton–Jacobi; ragioneremo, per fissare le idee, sull'equazione ridotta (4.56). Il problema è interessante sia perchè la ricerca di un integrale completo dell'equazione può essere, in qualche caso, più semplice o più naturale della soluzione diretta delle equazioni differenziali del moto, sia soprattutto perchè, al di là delle difficoltà tecniche di scrivere esplicitamente l'integrale completo, semplicemente sapere che tale integrale esiste permette di ottenere importanti informazioni qualitative sulle soluzioni, precisamente di sapere che esse hanno, nelle nuove e vecchie variabili, la forma specialissima (4.60) e (4.61), ovvero di sapere che le soluzioni delle equazioni differenziali del moto sono, nella sostanza, quelle del moto libero.

E' facile vedere che questo non è in generale possibile, nemmeno localmente, nell'intorno di un punto di equilibrio. Ciò è evidente nel caso di un punto di equilibrio instabile (si pensi, per fissare le idee, al pendolo rovesciato), per la chiara impossibilità delle (4.60), (4.61) di descrivere i moti sulle separatrici: se infatti $(\mathbf{p}^*,\mathbf{q}^*)$ è il punto di equilibrio nelle vecchie variabili, allora in corrispondenza ai moti sulle separatrici si ha $(\mathbf{p}(t),\mathbf{q}(t)) \to (\mathbf{p}^*,\mathbf{q}^*)$ per $t\to+\infty$ oppure $t\to-\infty$; corrispondentemente si dovrebbe avere, nelle nuove variabili, $(\tilde{\mathbf{p}}(t),\tilde{\mathbf{q}}(t)) \to (\tilde{\mathbf{p}}^*,\tilde{\mathbf{q}}^*)$, e ciò è chiaramente incompatibile con la forma (4.60) delle soluzioni. Difficoltà più sottili, che non illustreremo in queste note, sono connesse alla presenza di punti di equilibrio stabili.[35] Viceversa, è possibile dimostrare che *l'equazione di Hamilton–Jacobi ammette sempre un integrale completo locale, nell'intorno di un punto che non sia di equilibrio.* Pertanto, comunque si prenda un dato iniziale, non di equilibrio, $(\mathbf{p}^\circ,\mathbf{q}^\circ)$, in un intervallo di tempo $|t|<T$ sufficientemente breve la soluzione delle equazioni del moto ammette, in opportune nuove variabili, la forma (4.60), e corrispondentemente la forma (4.61) nelle variabili originarie. Dato il carattere strettamente locale, questo risultato non è in realtà molto forte; inoltre, il contesto meccanico e hamiltoniano, nel quale stiamo ragionando, non gioca in realtà alcun ruolo essenziale: si dimostra infatti del tutto in generale che per un sistema di equazioni differenziali in $\mathbb{R}^m$, $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, nell'intorno di un punto $\mathbf{x}^\circ$ non singolare esiste sempre un cambiamento di coordinate $\mathbf{x} = \mathbf{w}(\tilde{\mathbf{x}})$, tale che le soluzioni assumono localmente la forma, ancor più particolare della (4.60), $\tilde{x}_1(t) = \tilde{x}_1^\circ + t$, $\tilde{x}_i(t) = \tilde{x}_i^\circ$, $i>1$.

Per una esposizione più dettagliata di questo risultato, che porta il nome di *teorema di rettificazione*, si rinvia ai testi di analisi. L'idea comunque è la seguente: se $\mathbf{x}$ è tale che $\mathbf{f}(\mathbf{x}) \neq 0$, allora possiamo introdurre una *sezione di Poincaré locale*, cioè una superficie $(m-1)$-dimensionale $\Sigma$, trasversale al vettore $\mathbf{f}(\mathbf{x})$, munita di $m-1$ coordi-

---

[35] Nella sostanza, attorno ai punti di equilibrio stabili le orbite restano per tempi infiniti, e ciò consente ricorrenze e comportamenti complessi incompatibili con la forma troppo semplice (4.60).

nate locali arbitrarie $\sigma_1, \ldots, \sigma_{m-1}$; in un intorno $U$ di $\mathbf{x}$ possiamo allora introdurre un sistema di coordinate adattato, prendendo, in aggiunta a $\sigma_1, \ldots, \sigma_{m-1}$, un'ulteriore coordinata $\tau$, le cui linee coordinate coincidano con le traiettorie del sistema, e il cui valore rappresenti il tempo di percorrenza sulle traiettorie a partire da $\Sigma$. Le coordinate $\sigma_1, \ldots, \sigma_{m-1}$ sono allora costanti, mentre $\tau$ segue, evidentemente, l'equazione $\dot{\tau} = 1$. Le coordinate $\tilde{\mathbf{x}}$ cui si è sopra accennato sono date allora da $\tilde{x}_1 = \tau$, $\tilde{x}_i = \sigma_{i-1}$, $i = 2, \ldots, m$.

- **Osservazione.** Per un'equazione differenziale $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ in $\mathbb{R}^m$ è dunque vero che nell'intorno di un punto non singolare si possono introdurre nuove coordinate $\tilde{\mathbf{x}}$ tali che le equazioni del moto abbiano la forma $\dot{\tilde{x}}_1 = 1$, $\dot{\tilde{x}}_i = 0$, $i = 2, \ldots, m$. Vale allora la pena di osservare che, se $m$ è pari, $m = 2n$,[36] queste equazioni si possono sempre interpretare come equazioni di Hamilton: basta porre $q_i = \tilde{x}_i$, $p_i = \tilde{x}_{i+n}$, $i = 1, \ldots, n$, e prendere $H = p_1$. Si vede così che *qualunque sistema dinamico, osservato localmente nell'intorno di un punto non singolare, è hamiltoniano.* Sono dunque localmente hamiltoniani anche i più tipici sistemi dissipativi, come l'oscillatore armonico smorzato, o il sistema di Van der Pol, pur di non osservarli nell'intorno di un punto singolare. In altre parole: non solo la proprietà di essere integrabile, ma anche la più fondamentale proprietà di essere hamiltoniano, è in realtà molto povera, se intesa solo localmente (nell'intorno di un punto che non sia di equilibrio). La distinzione tra sistemi hamiltoniani, in particolare integrabili, e sistemi che non lo sono, è realmente significativa solo se si osservano proprietà globali, che coinvolgono il comportamento delle orbite per tempi infiniti (si ricordi che in ogni intorno di un punto singolare, stabile o instabile, vi sono orbite che restano nell'intorno per tempi infiniti, ed è questo a fare la differenza).

Veniamo allora al problema, ben più complesso e difficile, della ricerca di soluzioni globali all'equazione di Hamilton–Jacobi, ovvero il problema della ricerca di trasformazioni canoniche, definite in insiemi invarianti, che, come nel caso visto sopra di un sistema di oscillatori armonici disaccoppiati, coniughino una assegnata hamiltoniana a una hamiltoniana dipendente dai soli momenti. Contributi essenziali allo studio di questo problema sono stati dati, nel secolo scorso, da Liouville, e più tardi, al volger del secolo, da Poincaré, che vi dedicò molta attenzione;[37] più di recente vi si è dedicata, con splendidi risultati, soprattutto la scuola matematica russa.[38] Non è difficile comprendere il motivo di tanto interesse: ad esempio, la possibilità di avere risultati

---

[36] Se anche $m$ è dispari, ci si riconduce al caso di $m$ pari aggiungendo una ulteriore coordinata fittizia $x_{m+1}$, ponendo $\dot{\tilde{x}}_{m+1} = 0$.

[37] Nel suo “Méthodes Nouvelles de la Mechanique Céleste” Poincaré dà molto rilievo a questo problema, cui si riferisce col nome significativo di *problème genéral de la dynamique.*

[38] Principalmente A.N. Kolmogorov (un risultato veramente fondamentale nel 1954), V.I. Arnold (diversi importanti risultati soprattutto nel corso degli anni '60), e un allievo di Arnold, N.N. Nekhoroshev (un teorema assai rilevante nel 1976); ma bisognerebbe citare altri autori, come Littlewood, Siegel e soprattutto Moser, che pure hanno compiuto in questo campo studi fondamentali.

globali consente di trarre conclusioni sul comportamento asintotico delle soluzioni per tempi grandi, e dunque, in particolare, di studiare la stabilità dei sistemi, la loro risposta a perturbazioni, l'esistenza di integrali del moto, la rapidità con cui i diversi gradi di libertà scambiano tra loro energia...; dall'astronomia alla meccanica statistica si può dire che non vi sia campo della fisica ove non si pongano domande di questo genere.

Concludiamo questo paragrafo accennando (in modo purtroppo incompleto e molto affrettato) a un risultato particolarmente significativo, che consente di dare una caratterizzazione più profonda dei sistemi (globalmente) integrabili; unico scopo è la speranza di suscitare nel lettore curiosità e interesse ad approfondire personalmente l'argomento nei libri adatti (ad esempio, i libri di Arnold più volte citati).

E' chiaro che se, risolvendo l'equazione di Hamilton–Jacobi, si costruisce una trasformazione canonica $\mathbf{p} = \mathbf{u}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$, $\mathbf{q} = \mathbf{v}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$, che muta l'hamiltoniana $H(\mathbf{p}, \mathbf{q})$ in $K(\tilde{\mathbf{p}})$, allora il sistema ammette $n$ integrali del moto: se infatti $\tilde{\mathbf{p}} = \tilde{\mathbf{u}}(\mathbf{p}, \mathbf{q})$, $\tilde{\mathbf{q}} = \tilde{\mathbf{v}}(\mathbf{p}, \mathbf{q})$, sono le equazioni della trasformazione inversa, allora $\tilde{u}_1(\mathbf{p}, \mathbf{q}), \ldots, \tilde{u}_n(\mathbf{p}, \mathbf{q})$ sono integrali del moto. Poichè tali funzioni si possono prendere come nuovi momenti coniugati, allora (ricordando che le trasformazioni canoniche preservano le parentesi di Poisson) si vede che, necessariamente, questi $n$ integrali del moto sono in involuzione (cioè hanno mutue parentesi di Poisson nulle). Infine, trattandosi di un cambiamento di variabili, la matrice jacobiana $\frac{\partial(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})}{\partial(\mathbf{p}, \mathbf{q})}$ ha rango $2n$, e dunque, in particolare, le $n$ funzioni $\tilde{u}_1(\mathbf{p}, \mathbf{q}), \ldots, \tilde{u}_n(\mathbf{p}, \mathbf{q})$ sono indipendenti (più precisamente, gli $n$ gradienti $\frac{\partial \tilde{\mathbf{u}}}{\partial(\mathbf{p}, \mathbf{q})}$, corrispondenti alle prime $n$ righe della matrice jacobiana, sono linearmente indipendenti). Dunque, *condizione necessaria* perchè un sistema sia globalmente integrabile in un aperto invariante $U$, è che esista in $U$ un sistema completo di integrali del moto, indipendenti e in involuzione.[39] Un importante teorema, detto *teorema di Liouville–Arnold*, afferma che *la condizione è anche sufficiente.*[40] Precisamente, se esiste un sistema completo di integrali del moto indipendenti in involuzione, $F_1(\mathbf{p}, \mathbf{q}), \ldots, F_n(\mathbf{p}, \mathbf{q})$, allora in aperti invarianti definiti da condizioni del tipo $a_h < F_h(\mathbf{p}, \mathbf{q}) < b_h$, $h = 1, \ldots, n$, si ha che: i) l'equazione di Hamilton–Jacobi è risolubile per quadrature; ii) nel caso, particolarmente rilevante, in cui tale regione sia limitata, si possono scegliere le nuove coordinate $\tilde{q}_1, \ldots, \tilde{q}_n$ in modo che le equazioni della trasformazione canonica siano periodiche di periodo $2\pi$ in ciascuna $\tilde{q}_h$. Le nuove coordinate $\tilde{q}_1, \ldots, \tilde{q}_n$ sono dunque *angoli*; i momenti coniugati sono dimensionalmente *azioni.* Una notazione frequente per le nuove coordinate canoniche, dette *variabili di azione-angolo*, è $(\mathbf{I}, \boldsymbol{\varphi})$. Come mostrano le (4.61), che nella nuova notazione si scrivono

$$\mathbf{p} = \mathbf{u}(\mathbf{I}^\circ, \boldsymbol{\omega}(\mathbf{I}^\circ)t + \boldsymbol{\varphi}^\circ) \ , \qquad \mathbf{q} = \mathbf{v}(\mathbf{I}^\circ, \boldsymbol{\omega}(\mathbf{I}^\circ)t + \boldsymbol{\varphi}^\circ) \ ,$$

---

39) Coerentemente con quanto si è detto sopra, si potrebbe vedere che nei punti di equilibrio l'indipendenza lineare si perde. Si pensi al caso del pendolo: l'unico integrale del moto esistente, cioè l'hamiltoniana stessa (o una sua funzione), ha gradiente $\partial H/\partial(p, q)$ nullo nei punti di equilibrio.

40) Avendo in vista questo risultato, in molti testi la presenza di $n$ integrali del moto indipendenti in involuzione è presa come definizione stessa di integrabilità.

se esiste una trasformazione canonica che conduce alle variabili di azione-angolo, allora *il moto complessivo è quasi periodico,* con $n$ pulsazioni dipendenti in generale dal dato iniziale.

**Esempi.** Qualunque sistema a un solo grado di libertà è ovviamente integrabile (la sola conservazione di $H$ è sufficiente), in ogni regione compresa tra due superfici di energia costante che non contenga punti singolari;[41] è poi evidentemente integrabile[42] qualunque sistema a $n$ gradi di libertà "disaccoppiato", cioè con $H(\mathbf{p},\mathbf{q}) = \sum_h H_h(p_h, q_h)$ (sono infatti integrali del moto in involuzione le singole hamiltoniane $H_h$). In particolare sono dunque integrabili i sistemi incontrati nello studio delle piccole oscillazioni attorno alle configurazioni di equilibrio, a seguito del troncamento della lagrangiana ai soli termini quadratici: come si è visto, in coordinate opportune (coordinate normali) il sistema appare costituito da $n$ oscillatori armonici indipendenti, di hamiltoniana $H = \frac{1}{2}\sum_h(p_h^2 + \omega_h^2 q_h^2)$. Un altro esempio di sistema integrabile è offerto dal caso di forza centrale. Come sappiamo, in tal caso si conserva, oltre all'energia, anche il momento della quantità di moto $\mathbf{M}$; le sue tre componenti $M_x$, $M_y$, $M_z$ sono dunque integrali del moto, ma, come si è visto nel Capitolo 3, non sono in involuzione (si ha, ad esempio, $\{M_x, M_y\} = M_z$). Sono tuttavia in involuzione $H$, $M^2$ e (ad esempio) $M_z$, e dunque si conclude che un sistema con forze centrali di qualunque tipo (cioè con qualsiasi dipendenza dal raggio) è sempre integrabile. In modo analogo si vede che è integrabile il problema del corpo rigido con punto fisso, in assenza di forze attive esterne (problema di Eulero): anche in questo caso infatti si conservano l'energia totale, $M^2$ e $M_z$.

## 4.4 Introduzione alla teoria delle perturbazioni per sistemi hamiltoniani

### 4.4.1 *Sistemi prossimi a sistemi integrabili*

Supponiamo di avere un sistema hamiltoniano prossimo a un sistema integrabile, precisamente un sistema la cui hamiltoniana dipenda (in modo regolare) da un parametro $\varepsilon$, e sia integrabile per $\varepsilon = 0$:

$$H^\varepsilon(\mathbf{p},\mathbf{q}) = H_0(\mathbf{p},\mathbf{q}) + \varepsilon H_1(\mathbf{p},\mathbf{q},\varepsilon) \ , \tag{4.66}$$

con $H_1$ regolare in tutte le variabili. Il termine $H_0$ costituisce l'*hamiltoniana imperturbata*; $\varepsilon H_1$ rappresenta invece la *perturbazione.* Il problema che ci si pone naturalmente, di fronte all'hamiltoniana (4.66), è quale sia l'effetto della perturbazione, ovvero di quanto e in che cosa i moti del sistema perturbato differiscano da quelli del sistema imperturbato. Una prima informazione, significativa per tempi brevi, ma del tutto priva di utilità per tempi lunghi, viene dal teorema di esistenza e unicità delle

---

41) Nei punti singolari si ha $\frac{\partial H}{\partial(p,q)} = 0$, e viene meno l'indipendenza lineare.

42) Ancora in regioni invarianti che escludano punti di equilibrio delle singole hamiltoniane $H_h$.

soluzioni per le equazioni differenziali ordinarie, nella parte riguardante le continuità delle soluzioni al variare dei parametri: se indichiamo con $d_\varepsilon(t)$ la distanza tra la traiettoria perturbata e quella imperturbata al tempo $t$ (per un fissato dato iniziale), è noto che il teorema (applicato alle equazioni canoniche associate a $H^\varepsilon$) fornisce una stima del tipo $d_\varepsilon(t) < C\,\varepsilon\, e^{\lambda|t|}$, con $C$ e $\lambda$ costanti positive; a causa della dipendenza esponenziale da $t$, si vede bene che si perde in pratica ogni informazione su una scala di tempo $|t| \sim \lambda^{-1}\log\varepsilon^{-1}$, teoricamente infinita per $\varepsilon \to 0$, ma sostanzialmente limitata, a meno di valori di $\varepsilon$ assurdamente piccoli.

Informazioni molto più interessanti si ottengono tenendo conto che $H_0$ è integrabile, ovvero che esiste una trasformazione canonica $(\mathbf{p},\mathbf{q}) = \mathbf{w}(\mathbf{I},\boldsymbol{\varphi})$ che muta $H_0(\mathbf{p},\mathbf{q})$ in una hamiltoniana $K_0(\mathbf{I})$ dipendente dalle sole azioni. Usando per tutto $H^\varepsilon$ questa trasformazione canonica adattata a $H_0$ — ovvero introducendo nel nostro sistema le variabili di azione–angolo del problema imperturbato — si ottiene per la nuova hamiltoniana $K^\varepsilon$ la forma

$$K^\varepsilon(\mathbf{I},\boldsymbol{\varphi}) = K_0(\mathbf{I}) + \varepsilon K_1(\mathbf{I},\boldsymbol{\varphi},\varepsilon)\ ;$$

la funzione $K_1 = H_1 \circ \mathbf{w}$ esprime la perturbazione nelle nuove variabili. Posto allora $\omega_h(\mathbf{I}) = \frac{\partial K_0}{\partial I_h}$, le nuove equazioni del moto si scrivono

$$\begin{aligned} \dot I_h &= \qquad\quad - \varepsilon\frac{\partial K_1}{\partial \varphi_h} \\ \dot\varphi_h &= \omega_h(\mathbf{I}) + \varepsilon\frac{\partial K_1}{\partial I_h}\ . \end{aligned} \tag{4.67}$$

Queste equazioni mostrano che nel sistema perturbato le azioni (che nel sistema imperturbato sarebbero costanti) evolvono lentamente, con velocità proporzionale a $\varepsilon$: più precisamente, scrivendo $I_h(t) = I(0) - \varepsilon\int_0^t \frac{\partial K_1}{\partial\varphi_h}(\mathbf{I}(t),\boldsymbol{\varphi}(t))\,dt$, si ottiene immediatamente una stima del tipo

$$|I_h(t) - I_h(0)| < C\varepsilon t\ , \tag{4.68}$$

ove $C$ è il massimo di $|\frac{\partial K_1}{\partial\varphi_h}|$ nel dominio ove si svolge il moto. La scala di tempo sulla quale le azioni evolvono in maniera significativa è ora $|t| \sim C^{-1}\varepsilon^{-1}$, ben più lunga della precedente per $\varepsilon$ piccolo; su scale di tempo più brevi (ma sempre lunghe per $\varepsilon$ piccolo), ad esempio $|t| \sim \varepsilon^{-\frac12}$, le azioni variano di quantità piccole con $\varepsilon$, $|\mathbf{I}(t) - \mathbf{I}(0)| \sim \varepsilon^{\frac12}$. Le variabili di azione sono dette *variabili lente* del sistema. Al contrario, come mostrano le (4.67), gli angoli (salvo il caso in cui una o più pulsazioni si annullino) evolvono su una scala di tempo breve; anzi, fintantoché le azioni si discostano di poco dal valore iniziale, anche le velocità angolari si mantengono pressoché costanti, e dunque gli angoli ruotano, in prima approssimazione, in modo uniforme. Le variabili di angolo sono dette *variabili veloci* del sistema.[43)]

[43)] La separazione delle variabili in lente e veloci è però più delicata di quello che sembra. Ad esempio, se per un particolare valore di $\mathbf{I}$ si ha $\omega_1 = \omega_2$, allora $\varphi_1$ e $\varphi_2$ sono veloci, ma la loro differenza $\varphi_1 - \varphi_2$ è lenta; più in generale, se le pulsazioni soddisfano una *relazione di risonanza* $\sum_i k_i\omega_i = 0$, con i $k_i$ interi, allora la corrispondente combinazione di angoli $\sum_i k_i\varphi_i$

Come si è visto, la possibilità di scrivere le equazioni del moto nella forma (4.67) consente di trarre conclusioni sull'effetto della perturbazione che vanno ben al di là delle informazioni fornite dal teorema di esistenza e unicità. Un progresso ulteriore, assai significativo, si ottiene guardando in maniera più attenta alle (4.67), dalle quali si è dedotta in modo ancora troppo rozzo la stima "di ordine zero" (4.68). Questa stima si ottiene infatti supponendo, pessimisticamente, che gli effetti della perturbazione $\varepsilon\frac{\partial K_1}{\partial\varphi_h}$ vadano sommandosi; viceversa, è chiaro che sono da aspettarsi grosse compensazioni: infatti il termine $\frac{\partial K_1}{\partial\varphi_h}$, essendo la derivata rispetto a un angolo di una funzione periodica, ha necessariamente *media sugli angoli* nulla (la sua serie di Fourier manca del termine costante), e dunque la perturbazione ha segno alterno (l'effetto sarebbe anzi esattamente nullo, se fosse nulla la *media temporale lungo l'orbita* del termine $\frac{\partial K_1}{\partial\varphi_h}$).

Lo studio di queste compensazioni, o più in generale, lo studio della dinamica dei sistemi prossimi a sistemi integrabili su scale di tempo maggiori di $\varepsilon^{-1}$, è l'oggetto tipico della *teoria classica delle perturbazioni*, che si applica, sostanzialmente, ovunque vi siano, in un sistema dinamico (anche non hamiltoniano), variabili che evolvono su scale di tempo ben separate. L'interesse di questa teoria, nata con Lagrange e Laplace (in connessione ai moti planetari), e giunta a un primo livello di maturazione un secolo dopo, con Poincaré, deriva dal fatto che, mentre i sistemi esattamente integrabili sono eccezionali, invece la fisica offre una grande varietà di problemi ove intervengono sistemi dinamici prossimi a sistemi integrabili (più in generale, problemi caratterizzati dalla presenza di variabili lente e veloci). Un ulteriore livello di maturazione, assai rilevante per le sue implicazioni fisiche (ancora non abbastanza esplorate), si è avuto soprattutto a seguito dell'attività della scuola russa, cui si è sopra accennato. Nel prossimo paragrafo faremo un rapido cenno a uno dei metodi più significativi della teoria classica delle perturbazioni, basato sulla manipolazione dell'hamiltoniana per mezzo di costruzione di trasformazioni canoniche prossime all'identità; rinunciando a una trattazione generale, fisseremo l'attenzione su di un esempio particolare (ma significativo), precisamente un sistema di oscillatori armonici debolmente accoppiati. Per una trattazione più generale si rinvia, ancora una volta, ai libri di Arnold.

### 4.4.2 Equazione di Hamilton–Jacobi e teoria classica delle perturbazioni

Consideriamo un sistema di oscillatori armonici debolmente accoppiati, già scritto utilizzando le variabili di azione–angolo del sistema imperturbato,

$$H^\varepsilon(\mathbf{I},\boldsymbol{\varphi}) = H_0(\mathbf{I}) + \varepsilon H_1(\mathbf{I},\boldsymbol{\varphi},\varepsilon) \ , \qquad H_0 = \boldsymbol{\omega}\cdot\mathbf{I} \ , \tag{4.69}$$

con $\boldsymbol{\omega} = \text{cost}$. Pur essendo un caso abbastanza particolare, in quanto perturbazione di oscillatori esattamente armonici, una hamiltoniana di questa forma interviene in modo naturale in un gran numero di problemi fisici, in particolare nella trattazione hamil-

---

è lenta. Come si vedrà nel prossimo paragrafo, la presenza di tali variabili lente "nascoste" è una potenziale sorgente di difficoltà nella teoria delle perturbazioni.

toniana delle piccole oscillazioni attorno alle configurazioni di equilibrio stabile.[44] Ci proponiamo di trovare una trasformazione canonica $(\mathbf{I}, \boldsymbol{\varphi}) = \mathbf{w}(\mathbf{J}, \boldsymbol{\psi})$ prossima all'identità, tale che la nuova hamiltoniana $K^\varepsilon(\mathbf{J}, \boldsymbol{\psi}) = H^\varepsilon\big(\mathbf{w}(\mathbf{J}, \boldsymbol{\psi})\big)$ abbia la forma "integrabile a meno di termini di ordine $\varepsilon^2$"

$$K^\varepsilon = \boldsymbol{\omega} \cdot \mathbf{J} + \varepsilon K_1(\mathbf{J}) + \varepsilon^2 K_2(\mathbf{J}, \boldsymbol{\psi}, \varepsilon) \tag{4.70}$$

(i termini contenenti gli angoli, capaci di far evolvere le azioni, compaiono solo dall'ordine $\varepsilon^2$ in poi). Ricordando che l'identità è generata[45] da $F(\mathbf{J}, \boldsymbol{\varphi}) = \mathbf{J} \cdot \boldsymbol{\varphi}$, è spontaneo cercare una funzione generatrice della forma

$$F = \mathbf{J} \cdot \boldsymbol{\varphi} + \varepsilon \mathcal{F}(\mathbf{J}, \boldsymbol{\varphi}) \ . \tag{4.71}$$

Le equazioni della trasformazione implicitamente generata dalla (4.71) si scrivono

$$I_h = J_h + \varepsilon \frac{\partial \mathcal{F}}{\partial \varphi_h}(\mathbf{J}, \boldsymbol{\varphi}) \ , \qquad \psi_h = \varphi_h + \varepsilon \frac{\partial \mathcal{F}}{\partial J_h}(\mathbf{J}, \boldsymbol{\varphi}) \ , \tag{4.72}$$

ed è evidente che per inversione nella seconda, e sostituzione nella prima, si ottiene una trasformazione della forma[46]

$$I_h = J_h + \mathcal{O}(\varepsilon) \ , \qquad \varphi_h = \psi_h + \mathcal{O}(\varepsilon) \ . \tag{4.73}$$

Ora imponiamo che la nuova hamiltoniana $K^\varepsilon$ abbia la forma (4.70). Poichè risulta, in variabili miste,

$$\begin{aligned} K^\varepsilon &= \boldsymbol{\omega} \cdot \big(\mathbf{J} + \varepsilon \frac{\partial \mathcal{F}}{\partial \boldsymbol{\varphi}}(\mathbf{J}, \boldsymbol{\varphi})) + \varepsilon H_1\big(\mathbf{J} + \mathcal{O}(\varepsilon), \boldsymbol{\varphi}, \varepsilon\big) \\ &= \boldsymbol{\omega} \cdot \mathbf{J} + \varepsilon\big(\boldsymbol{\omega} \cdot \frac{\partial \mathcal{F}}{\partial \boldsymbol{\varphi}}(\mathbf{J}, \boldsymbol{\varphi}) + H_1(\mathbf{J}, \boldsymbol{\varphi}, 0)\big) + \mathcal{O}(\varepsilon^2) \ , \end{aligned}$$

è chiaro che la forma (4.70) si ottiene se e solo se si trovano $\mathcal{F}$ e $K_1$ tali che sia

$$\boldsymbol{\omega} \cdot \frac{\partial \mathcal{F}}{\partial \boldsymbol{\varphi}}(\mathbf{J}, \boldsymbol{\varphi}) + \mathcal{H}(\mathbf{J}, \boldsymbol{\varphi}) = K_1(\mathbf{J}) \ , \tag{4.74}$$

---

[44] Si giunge in particolare a tale forma studiando le piccole oscillazioni in sistemi lagrangiani. L'idea è questa: la lagrangiana troncata, scritta nelle coordinate normali, ha la forma $L_0 = \frac{1}{2}\sum(\dot{q}_h^2 + \omega_h^2 q_h^2)$; la lagrangiana vera è allora $L = L_0 + L_1$, con $L_1$ almeno cubico nelle $q_h, \dot{q}_h$. L'hamiltoniana corrispondente si mette allora subito nella forma $\frac{1}{2}\sum \omega_h(p_h^2 + q_h^2) + H_1(\mathbf{q})$, con $H_1$ almeno cubico nelle $p_h, q_h$. In un intorno di raggio $\varepsilon$ del punto di equilibrio $\mathbf{p} = 0$, $\mathbf{q} = 0$, è poi naturale introdurre la dilatazione $\mathbf{p} = \varepsilon\tilde{\mathbf{p}}$, $\mathbf{q} = \varepsilon\tilde{\mathbf{q}}$; la nuova hamiltoniana $K(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) = \varepsilon^{-2} H(\varepsilon\tilde{\mathbf{p}}, \varepsilon\tilde{\mathbf{q}})$ ha allora la forma $K = \frac{1}{2}\sum \omega_h(\tilde{p}_h^2 + \tilde{q}_h^2) + \mathcal{O}(\varepsilon)$, che a sua volta, con l'introduzione delle variabili di azione–angolo, conduce alla (4.69).

[45] Si tratta di una funzione generatrice "di tipo $F_2$"; qui e nel seguito l'indice 2 è omesso per non complicare la notazione.

[46] L'inversione è certamente possibile se $\varepsilon$ è abbastanza piccolo, ed è facile scriverla esplicitamente all'ordine $\varepsilon$: poichè si ha $\psi_h = \varphi_h + \mathcal{O}(\varepsilon)$, si ottiene subito

$$I_h = J_h + \varepsilon \frac{\partial \mathcal{F}}{\partial \varphi_h}(\mathbf{J}, \psi) + \mathcal{O}(\varepsilon^2) \ , \qquad \varphi_h = \psi_h + \varepsilon \frac{\partial \mathcal{F}}{\partial J_h}(\mathbf{J}, \psi) + \mathcal{O}(\varepsilon^2) \ .$$

ove si è posto $\mathcal{H}(\mathbf{J},\boldsymbol{\varphi}) = H_1(\mathbf{J},\boldsymbol{\varphi},0)$. Infatti, se tale equazione è soddisfatta, allora $K$, all'ordine $\varepsilon$, non dipende da $\boldsymbol{\varphi}$ e dunque, dopo la sostituzione con la seconda delle (4.73), neanche da $\boldsymbol{\psi}$; viceversa, se $K$ all'ordine $\varepsilon$ dipendesse da $\boldsymbol{\varphi}$, dopo la sostituzione dipenderebbe da $\boldsymbol{\psi}$.

- **Osservazione.** L'equazione(4.74) non è altro che l'equazione (ridotta) di Hamilton–Jacobi "all'ordine $\varepsilon$". Infatti, l'equazione di Hamilton–Jacobi per l'hamiltoniana $H^\varepsilon$ si scrive
$$\boldsymbol{\omega}\cdot\frac{\partial W}{\partial\boldsymbol{\varphi}} + \varepsilon H_1\Big(\frac{\partial W}{\partial\boldsymbol{\varphi}},\boldsymbol{\varphi},\varepsilon\Big) = E \ ; \tag{4.75}$$
per $\varepsilon = 0$ un integrale completo è dato da $W_{\alpha_1,\ldots,\alpha_n} = \boldsymbol{\alpha}\cdot\boldsymbol{\varphi}$, $E_{\alpha_1,\ldots,\alpha_n} = \boldsymbol{\omega}\cdot\boldsymbol{\alpha}$ (la corrispondente $F$ genera l'identità), e dunque, per trovare un integrale completo approssimato, che risolva l'equazione a meno di termini di ordine $\varepsilon^2$, è spontaneo porre
$$W_\alpha = \boldsymbol{\alpha}\cdot\boldsymbol{\varphi} + \varepsilon\,\mathcal{W}_\alpha(\boldsymbol{\varphi}) \ , \qquad E_\alpha = \boldsymbol{\omega}\cdot\boldsymbol{\alpha} + \varepsilon\,\mathcal{E}_\alpha \ . \tag{4.76}$$
Sostituendo le (4.76) nella (4.75), e raccogliendo i termini di ordine $\varepsilon$, si ottiene l'equazione
$$\boldsymbol{\omega}\cdot\frac{\partial}{\partial\boldsymbol{\varphi}}\mathcal{W}_\alpha + \mathcal{H}_\alpha = \mathcal{E}_\alpha \ ,$$
che coincide con la (4.74) (ove si ponga $\mathcal{F}(\mathbf{J},\boldsymbol{\varphi}) = \mathcal{W}_{\mathbf{J}}(\boldsymbol{\varphi})$, $\mathcal{H}_{\mathbf{J}}(\boldsymbol{\varphi}) = \mathcal{H}(\mathbf{J},\boldsymbol{\varphi})$, $K_1(\mathbf{J}) = \mathcal{E}_{\mathbf{J}}$).

Veniamo allora alla soluzione della (4.74); per semplicità ci restringiamo al caso di due soli oscillatori armonici, e per evitare problemi di convergenza supponiamo anche che la serie di Fourier della perturbazione sia finita, ovvero che si abbia, con opportuno $N$,
$$\mathcal{H}(J_1,J_2,\varphi_1,\varphi_2) = \sum_{-N\le k_1,k_2\le N} \hat{\mathcal{H}}_{k_1k_2}(J_1,J_2)e^{i(k_1\varphi_1+k_2\varphi_2)}$$
(ciò si verifica quando la perturbazione, espressa in funzione delle variabili $\mathbf{p}$, $\mathbf{q}$, prima dell'introduzione delle variabili di azione–angolo, è polinomiale). Guardando la (4.74), si vede che serve $\mathcal{F}$ capace, per così dire, di compensare tutti i termini della serie di Fourier di $\mathcal{H}$, tranne il valor medio $\hat{\mathcal{H}}_{00}$ (che non dipende dagli angoli). E' pertanto spontaneo cercare $\mathcal{F}$ della forma
$$\mathcal{F}(J_1,J_2,\varphi_1,\varphi_2) = \sum_{-N\le k_1,k_2\le N} \hat{\mathcal{F}}_{k_1k_2}(J_1,J_2)e^{i(k_1\varphi_1+k_2\varphi_2)} \ ,$$
al fine di soddisfare poi la (4.74) componente per componente. Con calcoli elementari si trova allora, per $(k_1,k_2) = (0,0)$,
$$\hat{\mathcal{H}}_{00}(J_1,J_2) = K_1(J_1,J_2) \ , \tag{4.77}$$
mentre per gli altri valori di $(k_1,k_2)$ si ha
$$i(k_1\omega_1 + k_2\omega_2)\hat{\mathcal{F}}_{k_1,k_2}(J_1,J_2) + \hat{\mathcal{H}}_{k_1,k_2}(J_1,J_2) = 0 \ . \tag{4.78}$$

La (4.77) mostra che *la prima correzione all'hamiltoniana integrabile, dovuta alla perturbazione, coincide con il valor medio della perturbazione sugli angoli.* Veniamo ora alla (4.78). Perchè questa si possa risolvere, è necessario che sia

$$k_1\omega_1 + k_2\omega_2 \neq 0 \;, \qquad \text{per} \quad -N \leq k_1, k_2 \leq N \;, \quad (k_1, k_2) \neq (0,0) \;; \tag{4.79}$$

questa condizione, detta *condizione di non risonanza*, esclude la presenza di rapporti razionali tra le frequenze $\omega_1$ e $\omega_2$, $\frac{\omega_1}{\omega_2} = -\frac{k_2}{k_1}$, con $|k_1|, |k_2| \leq N$.

Se questa condizione è soddisfatta, allora si può scrivere $\hat{\mathcal{F}}_{k_1,k_2} = i\frac{\hat{\mathcal{H}}_{k_1,k_2}}{k_1\omega_1 + k_2\omega_2}$, e pertanto la (4.74) è risolta da

$$\mathcal{F}(J_1, J_2, \varphi_1, \varphi_2) \;=\; i \sum_{\substack{-N \leq k_1, k_2 \leq N \\ (k_1,k_2)\neq(0,0)}} \frac{\hat{\mathcal{H}}_{k_1,k_2}(J_1, J_2)}{k_1\omega_1 + k_2\omega_2} \; e^{i(k_1\varphi_1 + k_2\varphi_2)} \tag{4.80}$$

(assieme a $K_1(\mathbf{J}) = \hat{\mathcal{H}}_{00}(\mathbf{J})$). Si vede così che, se la condizione di non risonanza (4.79) è soddisfatta, allora esiste[47)] una trasformazione canonica che muta l'hamiltoniana $H^\varepsilon$ in una hamiltoniana $K^\varepsilon$ integrabile, come richiesto, a meno di termini di ordine $\varepsilon^2$.

L'interesse per tale trasformazione è evidente: in sostanza, per $K^\varepsilon$ si possono ripetere tutte le considerazioni svolte nel paragrafo precedente a proposito dei sistemi prossimi a sistemi integrabili, con $\varepsilon^2$ in luogo di $\varepsilon$; in particolare, le nuove azioni evolvono solo su una scala di tempo $\sim \varepsilon^{-2}$, assai più lunga di $\varepsilon^{-1}$ per $\varepsilon$ piccolo. Come poi mostrano le (4.72), le vecchie e nuove azioni coincidono a meno di termini di ordine $\varepsilon$, cioè sono $\varepsilon$–vicine (uniformemente in $t$): si conclude che *anche le vecchie azioni evolvono solo su una scala di tempo* $\sim \varepsilon^{-2}$. Questo fatto non traspariva dall'analisi "di ordine zero" fatta nel paragrafo precedente, ed è chiaramente dovuto alle compensazioni cui si è sopra accennato.

Abbiamo qui ragionato, per comodità, su due soli oscillatori, ma il risultato si trasporta immediatamente al caso di $n$ oscillatori, semplicemente sostituendo alla condizione (4.79) che le frequenze non abbiano rapporto razionale (con interi piccoli) la condizione che non vi siano *relazioni di risonanza* della forma

$$k_1\omega_1 + \cdots k_n\omega_n = 0 \;,$$

per $k_1, \ldots, k_n$ non tutti nulli (e piccoli). Un risultato di questo tipo ha implicazioni fisiche abbastanza notevoli: ad esempio, riferendoci al problema sopra introdotto delle piccole oscillazioni attorno a una configurazione di equilibrio stabile, quanto abbiamo visto vuol dire che, in assenza di relazioni di risonanza con interi piccoli, *gli oscillatori corrispondenti ai modi normali di oscillazione sono sostanzialmente disaccoppiati* (mantengono ampiezza di oscillazione costante, e non scambiano tra loro energia), *fino a tempi* $\sim \varepsilon^{-2}$.

---

47) Come si vede dalla (4.77), la nuova hamiltoniana è data dalla vecchia, più il valor medio della perturbazione sugli angoli, e dunque la si conosce esplicitamente fino all'ordine $\varepsilon$. Con poca fatica in più si potrebbero anche stimare i termini di ordine superiore dipendenti dagli angoli, in modo da poter usare la nuova hamiltoniana in modo quantitativo.

La domanda spontanea sarebbe, a questo punto, se per caso non esista una successione di trasformazioni canoniche, piccole con $\varepsilon^2$, $\varepsilon^3$,... capaci di spostare la dipendenza della hamiltoniana dagli angoli a ordini perturbativi sempre più elevati. La questione è difficile, e del tutto impossibile da affrontare in queste note; diciamo soltanto che sia per il caso qui considerato degli oscillatori armonici debolmente interagenti, sia per altre iteressanti classi di sistemi dinamici hamiltoniani, questo argomento è stato ed è attualmente oggetto di studio intenso nell'ambito della teoria classica delle perturbazioni. Ad esempio, sotto opportune ipotesi, si riesce a dimostrare che la scala di tempo sulla quale le azioni evolvono in modo significativo (e dunque i diversi gradi di libertà del sistema interagiscono in modo efficace) è "estremamente lunga" per $\varepsilon$ piccolo, più precisamente cresce al calare di $\varepsilon$ con andamento *esponenziale* $\sim e^{1/\varepsilon}$;[48)] oppure, che l'interazione non avviene in modo significativo neanche su un tempo infinito, per dati iniziali scelti in un conveniente insieme dello spazio delle fasi.

Dalle considerazioni sopra svolte si intuisce anche il ruolo delicato delle risonanze nel determinare il comportamento di un sistema.[49)] Con ciò non si vuol dire che in presenza di risonanze la teoria delle perturbazioni non si possa applicare; tuttavia sono necessarie modifiche significative allo schema sopra esposto, ed è facile vedere, ad esempio, che una stessa perturbazione produce in generale effetti qualitativamente diversi, a seconda delle relazioni di risonanza presenti nell'hamiltoniana imperturbata.

---

48) Più in generale $e^{a/\varepsilon^b}$, con $a, b > 0$; la scala di tempo cresce dunque più rapidamente di qualunque potenza negativa di $\varepsilon$

49) Si ricordi a questo proposito quanto si è visto nel Capitolo 1, a proposito delle oscillazioni forzate.

# APPENDICE

Riportiamo qui la dimostrazione della Proposizione 4.13, cominciando col dimostrare due semplici lemmi.

**Lemma 4.1:** *Condizione necessaria e sufficiente perché un assegnato campo vettoriale* $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_n(\mathbf{x}))$ *sia il gradiente simplettico di una funzione* $\Phi$*, cioè risulti*

$$\mathbf{f} = \{\mathbf{x}, \Phi\} \ , \tag{A1}$$

*è che sia verificata la condizione di compatibilità*

$$\{x_i, f_j\} \ = \ \{x_j, f_i\} \ , \qquad i, j = 1, \dots, n \ . \tag{A2}$$

**Dimostrazione.** Per dimostrare la sufficienza, supponiamo vera la (A2), a sua volta equivalente a $\sum_k E_{ik} \frac{\partial f_j}{\partial x_k} = \sum_k E_{jk} \frac{\partial f_i}{\partial x_k}$. Moltiplicando entrambi i membri per $E_{ri} E_{sj}$, sommando su $i$ e $j$ e ricordando che $E^2 = -1$, si trova, con un po' di pazienza,

$$\frac{\partial g_r}{\partial x_s} = \frac{\partial g_s}{\partial x_r} \ ,$$

ove si è posto $\mathbf{g} = E\mathbf{f}$. Ma se questa condizione di compatibilità è soddisfatta, allora esiste $\Phi$ tale che sia $\mathbf{g} = -\partial_x \Phi$, e dunque si ha $\mathbf{f} = E \partial_x \Phi = \{\mathbf{x}, \Phi\}$.

Ripercorrendo al contrario la dimostrazione della sufficienza, si prova facilmente la necessità. *Q.E.D.*

**Lemma 4.2:** *Se la matrice* $A(x)$ *ha la proprietà che, comunque si prenda la funzione* $H(\mathbf{x})$*, il campo vettoriale* $A\partial_x H$ *è irrotazionale, allora* $A = cI$*, con* $c$ *costante.*

**Dimostrazione.** Se $A\partial_x H$ è irrotazionale, allora

$$\frac{\partial}{\partial x_r}(A\partial_x H)_s \ = \ \frac{\partial}{\partial x_s}(A\partial_x H)_r \qquad \forall r, s \ . \tag{A3}$$

Per $H(\mathbf{x}) = x_r$ segue allora

$$\frac{\partial}{\partial x_r} A_{sr} \ = \ \frac{\partial}{\partial x_s} A_{rr} \ , \tag{A4}$$

mentre per $H(\mathbf{x}) = x_r^2$ si trova

$$\frac{\partial}{\partial x_r}(A_{sr} x_r) \ = \ \frac{\partial}{\partial x_s}(A_{rr} x_r) \ .$$

Da queste relazioni si deduce immediatamente $A_{sr} = A_{rr}\delta_{rs}$, e dunque la matrice $A$ è diagonale. La (A4) dà allora

$$\frac{\partial}{\partial x_s} A_{rr} = 0 \ , \qquad r \neq s \ ; \tag{A5}$$

infine, usando tutte queste informazioni, la (A3) si riduce a

$$A_{ss}\frac{\partial^2 H}{\partial x_r \partial x_s} = A_{rr}\frac{\partial^2 H}{\partial x_s \partial x_r} \ .$$

Pretanto si ha $A_{ss} = A_{rr}$, e tenendo conto della (A5) si vede che la matrice, oltre a essere diagonale, ha elementi costanti. *Q.E.D.*

Veniamo ora alla dimostrazione della Proposizione 4.13.

**Prima parte.** Mostriamo innanzitutto che, se la trasformazione preserva le parentesi di Poisson, ovvero si ha

$$c\{F,G\} = \{f,g\}\circ \mathbf{w} \ , \tag{4.81}$$

allora esiste $K_0(\tilde{\mathbf{x}},t)$ per cui, indicando con $\tilde{\mathbf{x}} = \mathbf{w}(\mathbf{x},t)$ la trasformazione inversa alla (4.52), si ha

$$\frac{\partial \tilde{\mathbf{w}}}{\partial t} = \{\tilde{\mathbf{x}}, K_0\}\circ \tilde{\mathbf{w}} \ . \tag{A6}$$

Infatti, dalla (4.81) segue in particolare $\{\tilde{w}_i, \tilde{w}_j\} = -cE_{ij}\ \forall i,j$, da cui per derivazione si ottiene

$$\{\frac{\partial \tilde{w}_i}{\partial t}, \tilde{w}_j\} + \{\tilde{w}_i, \frac{\partial \tilde{w}_j}{\partial t}\} = 0 \qquad \forall i,j \ ;$$

posto allora $F_i = \frac{\partial \tilde{w}_i}{\partial t}\circ \mathbf{w}$, utilizzando ancora la (4.81) si ha $\{x_i, F_j\} + \{F_i, x_j\} = 0$, e per il Lemma 4.1, ciò basta a concludere che esiste $K_0$, tale che sia $\mathbf{F} = \{\tilde{\mathbf{x}}, K_0\}$, cioè che vale la (A6).

E' facile ora concludere che il moto nelle nuove coordinate $\tilde{\mathbf{x}}$ è soluzione delle equazioni di Hamilton relative a $K = \hat{H} + K_0$, con $\hat{H} = cH\circ \mathbf{w}$: infatti, preso $H$ ad arbitrio, lungo il moto si ha

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{w}} = \frac{\partial \tilde{\mathbf{w}}}{\partial t} + \{\tilde{\mathbf{w}}, H\}\circ \tilde{\mathbf{w}} \ ,$$

e dunque, usando la (A6) e la (4.81),

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{w}} = \{\tilde{\mathbf{x}}, K_0\}\circ \tilde{\mathbf{w}} + \{\tilde{\mathbf{x}}, \hat{H}\}\circ \tilde{\mathbf{w}}\{\tilde{\mathbf{x}}, K\}\circ \tilde{\mathbf{w}} \ .$$

*Q.E.D.*

**Seconda parte.** Presa $f(\mathbf{x},t)$ ad arbitrio, poniamo $F(\tilde{\mathbf{x}},t) = f(\mathbf{w}(\tilde{\mathbf{x}},t),t)$; per moti che siano coniugati dalla (4.52) si ha, evidentemente, $\dot{F} = \dot{f}\circ \mathbf{w}$ (istante per istante, su traiettorie coniugate $F$ e $f$ assumono per definizione il medesimo valore). La funzione $f$ soddisfa, sappiamo, la relazione $\dot{f} = \frac{\partial f}{\partial t} + \{f,H\}$; d'altra parte, per l'assunzione fatta che la trasformazione (4.52) sia canonica, allora anche $F$ soddisfa la relazione $\dot{F} = \frac{\partial F}{\partial t} + \{F,K\}$, ove $K$ è la nuova hamiltoniana. Pertanto si ha

$$\{F,K\} + \frac{\partial F}{\partial t} = \{f,H\}\circ \mathbf{w} + \frac{\partial f}{\partial t}\circ \mathbf{w} \ . \tag{A7}$$

Sia ora $K_0(\tilde{\mathbf{x}}, t)$ l'hamiltoniana associata dalla trasformazione a $H = 0$; dalla (A7) si ricava, in particolare,

$$\frac{\partial F}{\partial t} = \frac{\partial f}{\partial t} \circ \mathbf{w} - \{F, K_0\} \ ,$$

cosicché, per sostituzione nella (A7), si trova

$$\{F, \hat{K}\} = \{f, H\} \circ \mathbf{w} \ , \tag{A8}$$

con $\hat{K} = K - K_0$.

Da questa relazione vogliamo dedurre (sfruttando l'arbitrarietà di $H$ e di $f$) che risulta necessariamente $\hat{K} = cH \circ \mathbf{w}$, cosicché $K$ ha la forma voluta (4.53). A questo scopo scriviamo le parentesi di Poisson in forma di prodotto scalare simplettico, precisamente

$$\partial_{\tilde{x}} \hat{K} \cdot E \partial_{\tilde{x}} F = (\partial_x H \cdot E \partial_x f) \circ \mathbf{w} \ . \tag{A9}$$

Poniamo ora $\hat{H} = \hat{K} \circ \tilde{\mathbf{w}}$, ove $\tilde{\mathbf{w}}$ è l'inversa di $\mathbf{w}$, cosicché si ha $\hat{K} = \hat{H} \circ \mathbf{w}$. Omettendo, per brevità di notazione, la composizione delle funzioni con $\mathbf{w}$ (sottintesa ovunque necessaria per dar senso alle espressioni), il membro di sinistra si può scrivere

$$\begin{aligned} \partial_{\tilde{x}} \hat{H}(\mathbf{w}(\tilde{\mathbf{x}}, t), t) \cdot E \partial_{\tilde{x}} f(\mathbf{w}(\tilde{\mathbf{x}}, t), t) &= J^T \partial_x \hat{H} \cdot E J^T \partial_x f \\ &= -J E J^T \partial_x \hat{H} \cdot \partial_x f \ , \end{aligned}$$

pertanto la (A9) prende la forma

$$J E J^T \partial_x \hat{H} \cdot \partial_x f = E \partial_x H \cdot \partial_x f \ .$$

Per l'arbitrarietà di $f$, ciò implica che sia $J E J^T \partial_x \hat{H} = E \partial_x H$, ovvero

$$\partial_x \hat{H} = A \partial_x H \ ,$$

con $A = -(E J E J^T)^{-1}$, o più precisamente $A = -(E J E J^T)^{-1} \circ \tilde{\mathbf{w}}$. Si vede così che la matrice $A$ ha la proprietà che, comunque si prenda $H$, $A \partial_x H$ è irrotazionale (è un gradiente). Dal Lemma 4.2 si deduce allora che $A$ è proporzionale all'identità, e dunque si ottiene la relazione

$$J E J^T = c^{-1} E \ , \tag{A10}$$

Con ulteriori considerazioni si potrebbe anche vedere che la costante $c$ non può dipendere neanche da $t$.

Abbiamo così ottenuto $\partial_x \hat{H} = c \partial_x H$, e dunque (trascurando l'inessenziale costante additiva) $\hat{H} = cH$, cioè $\hat{K} = cH \circ \mathbf{w}$. La (A10) è poi equivalente alla (4.81) $\big($o se vogliamo, la (A8), con $\hat{K} = cH \circ \mathbf{w}$ e $H$, $f$ arbitrari, coincide con la (4.81)$\big)$. *Q.E.D.*