

DIPARTIMENTO
DI MATEMATICA



SAPIENZA
UNIVERSITÀ DI ROMA

Appunti di meccanica hamiltoniana per il corso di IFM 2024-2025

© 2025 di Dario Benedetto con licenza attribuzione - non commerciale
- condividi allo stesso modo 4.0 internazionale CC BY-NC-SA 4.0 

Dario Benedetto - <http://brazil.mat.uniroma1.it/dario>

Sapienza Università di Roma
Dipartimento di Matematica
Piazzale Aldo Moro n. 5, 00185 Roma
www.mat.uniroma1.it

Appunti di meccanica hamiltoniana 2024/2025

1 aprile 2025

Indice

1	EDO, flussi e equazione di Liouville	4
1.1	EDO e flussi	4
1.2	L'equazione di Liouville	6
1.3	Il pushforward	7
2	Flussi a divergenza nulla	8
2.1	Il caso dei sistemi meccanici conservativi	9
3	Dal metodo delle caratteristiche alle equazioni di Hamilton	9
3.1	Metodo delle caratteristiche	9
3.2	L'equazione di Hamilton-Jacobi e le equazioni di Hamilton	12
3.3	Prime proprietà dei sistemi hamiltoniani	13
3.4	Un principio variazionale per le equazioni di Hamilton	15
3.5	Dalla lagrangiana all'hamiltoniana, versione differenziale	17
3.6	Lagrangiane e hamiltoniane naturali	18
3.7	Variabili cicliche e riduzione dei gradi di libertà	19
4	Trasformazioni canoniche e simplettiche	21
4.1	Trasformazioni di coordinate	21
4.2	$\mathbf{p} \cdot d\mathbf{q} - H dt$	22
4.3	Trasformazioni simplettiche	24
5	Parentesi di Poisson	26
5.1	Integrali primi	33
5.2	Il teorema di Noether per sistemi hamiltoniani	35
6	L'equazione di Hamilton-Jacobi	36
6.1	Funzioni generatrici	36
6.2	Il metodo di Hamilton-Jacobi	39
6.3	Esempio: moto centrale con termine di dipolo	41
7	Sistemi integrabili	43
7.1	Sistemi integrabili	44
7.2	Geometria simplettica	45
7.3	Integrabilità locale	47

7.4	Integrabilità globale	49
7.5	Moti quasi periodici	52
7.6	Variabili azione-angolo	53

8 Il teorema del ritorno di Poincaré **55**

Queste note sono “appunti” e non “dispense”: vuol dire che trovate gli argomenti e le dimostrazioni, ma le motivazioni sono piuttosto stringate. Inoltre, presuppongono una buona conoscenza del formalismo lagrangiano, che dovrebbe essere noto dai corsi della triennale. Per approfondire i temi di questi appunti suggerisco la lettura di:

A V.I. Arnold: **Metodi matematici della meccanica classica** Editori Riuniti

BN P. Buttà, P. Negrini: **Note del corso di sistemi dinamici**

www1.mat.uniroma1.it/~butta/didattica/sisdin.pdf

E R. Esposito: **Appunti delle lezioni di meccanica razionale**, Aracne 1999.

C E. Caglioti: **Dispense di Meccanica Razionale** che trovate su

<https://sites.google.com/site/ecaglioti/didattica/MR>

EV L. C. Evans **Partial Differential Equation** AMS Graduate Studies in Mathematics, vol. 19

1 EDO, flussi e equazione di Liouville

1.1 EDO e flussi

Introdurrò l'equazione di Liouville discutendo la descrizione statistica di un moto governato da una EDO.

Indico con $\Phi^{t,s}(\mathbf{x})$ il flusso associato all'EDO

$$\dot{\mathbf{x}} = \mathbf{u}(t, \mathbf{x}),$$

cioè $\Phi^{t,s}(\mathbf{x})$ risolve l'equazione differenziale come funzione di t , con dato al tempo s uguale a \mathbf{x} : $\Phi^{s,s}(\mathbf{x})$:

$$\begin{aligned} \frac{d}{dt}\Phi^{t,s}(\mathbf{x}) &= \mathbf{u}(t, \Phi^{t,s}(\mathbf{x})) \\ \Phi^{s,s}(\mathbf{x}) &= \mathbf{x}. \end{aligned} \tag{1.1}$$

Suppongo per semplicità di essere nelle condizioni che garantiscono esistenza globale nel tempo, unicità, e regolarità delle soluzioni in tutto \mathbb{R}^d . È facile provare che dalla definizione e dalle assunzioni fatte, il flusso è una famiglia di diffeomorfismi a due parametri che verifica, per ogni $r, s, t, \in \mathbb{R}$:

$$\begin{aligned} \Phi^{t,s} \circ \Phi^{s,r} &= \Phi^{t,r} \\ \Phi^{s,s} &= \mathbf{id} \end{aligned} \tag{1.2}$$

e in particolare

$$(\Phi^{t,s})^{-1} = \Phi^{s,t}.$$

Il flusso non contiene un maggior numero di informazioni rispetto al sistema differenziale, però la descrizione a livello di flusso ci permette di fare nuove domande. Qui in particolare ci interessiamo della descrizione del sistema in presenza di incertezza nel dato iniziale. Tenete presente che questo è in effetti quello che accade nella realtà, infatti non è possibile conoscere il dato iniziale con infinita precisione.

Immaginiamo dunque di assegnare una densità di probabilità per i dati iniziali al tempo $t = 0$, che indicherò con $f_0(x)$ (per semplicità, considero il tempo $s = 0$ come tempo iniziale ma nulla cambia nel caso il tempo iniziale sia un altro). Questa incertezza sul dato iniziale si traduce in una incertezza sulla soluzione al tempo t . L'equazione di Liouville è proprio l'equazione che governa il cambiamento nel tempo di questa distribuzione di probabilità.

Per esempio, supponiamo che il dato iniziale sia uniformemente distribuito in un dominio B , quindi

$$f_0(\mathbf{x}) = \frac{1}{|B|} \mathcal{X}\{\mathbf{x} \in B\}$$

dove $|B|$ è la misura di B . Cosa sappiamo del sistema al tempo t ? Visto che l'evoluzione è **deterministica** (è solo il dato iniziale che ha una descrizione probabilistica), sicuramente la probabilità di trovare il sistema in

$$B_t = \Phi^{t,0}(B) = \{\Phi^{t,0}(\mathbf{x}) \mid \mathbf{x} \in B\}$$

sarà 1, ma non abbiamo motivo di pensare che la distribuzione resti uniforme. È necessaria dunque un'analisi più approfondita per capire come evolve f . Per fare questa analisi ci servono alcune proprietà del flusso.

Proposizione 1.1. Sia $\Phi^{t,s}$ il flusso associato al campo $\mathbf{u}(t, \mathbf{x})$ e sia $J(\mathbf{x}, t, s) = \det \frac{\partial \Phi^{t,s}}{\partial \mathbf{x}}(\mathbf{x})$ il determinante jacobiano associato al flusso. Considerando noto il flusso, lo jacobiano del flusso verifica la seguente equazione lineare a coefficienti non costanti (in forma matriciale)

$$\frac{d}{dt} \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}) = \partial_{\mathbf{x}} \mathbf{u}|_{t, \Phi^{t,s}(\mathbf{x})} \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}) \quad (1.3)$$

e il determinante jacobiano verifica la seguente equazione differenziale lineare a coefficienti non costanti

$$\frac{d}{dt} J(\mathbf{x}, t, s) = \operatorname{div} \mathbf{u}|_{t, \Phi^{t,s}(\mathbf{x})} J(\mathbf{x}, t, s) \quad (1.4)$$

La prima equazione si ottiene semplicemente scambiando gli ordini di derivazione e usando l'equazione che definisce il flusso. La seconda equazione richiede un lavoro maggiore.

Il cuore della dimostrazione è nel seguente lemma

Lemma 1.1.

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det \partial_{\mathbf{x}} \Phi^{t+\varepsilon, t}(\mathbf{x}) = \operatorname{div} \mathbf{u}(t, \mathbf{x})$$

Il testo un po' barocco dell'affermazione è indispensabile perché voglio derivare $\partial_{\mathbf{x}} \Phi^{t,t}$ rispetto alla prima t . Una volta dimostrato questo lemma, si può procedere al calcolo di $\frac{d}{dt} J(\mathbf{x}, t, s)$:

$$\begin{aligned} \frac{d}{dt} \det \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}) &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det \partial_{\mathbf{x}} (\Phi^{t+\varepsilon, t} \circ \Phi^{t,s}(\mathbf{x})) \\ &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det \partial_{\mathbf{x}} \Phi^{t+\varepsilon, t}|_{\Phi^{t,s}(\mathbf{x}), t} \det \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}) \\ &= \operatorname{div} \mathbf{u}|_{t, \Phi^{t,s}(\mathbf{x})} \det \partial_{\mathbf{x}} \Phi^{t,s}(\mathbf{x}). \end{aligned} \quad (1.5)$$

Rimane da dimostrare il lemma, che è conseguenza della seguente affermazione sul determinante di una perturbazione della matrice identità.

Lemma 1.2. Sia A una matrice quadrata, allora

$$\det(\mathbf{I} + \varepsilon A) = 1 + \varepsilon \operatorname{Tr} A + O(\varepsilon^2)$$

Dimostrazione. Sia $M = \mathbf{I} + \varepsilon A$.

$$\det M = \sum_{\sigma} (-1)^{\sigma} \prod_{i=1}^n M_{i\sigma_i}$$

dove la somma è sulle permutazioni degli indici da 1 a n . Scegliendo la permutazione identica si ottiene il termine

$$\prod_{i=1}^n M_{ii} = \prod_{i=1}^n (1 + \varepsilon A_{ii}) = 1 + \varepsilon \operatorname{Tr} A + O(\varepsilon^2).$$

La prova si conclude notando che ogni altra permutazione dà un termine di ordine maggiore o uguale a 2 (perché?).

Dimostriamo ora il lemma 1.1 Dal fatto che la derivata del flusso nel tempo è il campo \mathbf{u} , segue che

$$\Phi^{t+\varepsilon, t}(\mathbf{x}) = \mathbf{x} + \varepsilon \mathbf{u}(t, \mathbf{x}) + O(\varepsilon^2)$$

Derivo in \mathbf{x} e calcolo il determinante:

$$\det \partial_{\mathbf{x}} \Phi^{t+\varepsilon, t}(\mathbf{x}) = \det (\mathbf{I} + \varepsilon \partial_{\mathbf{x}} \mathbf{u}(t, \mathbf{x}) + O(\varepsilon^2)) = 1 + \varepsilon \operatorname{div} \mathbf{u}(t, \mathbf{x}) + O(\varepsilon^2)$$

(dove ho usato che la traccia dello jacobiano di un campo è proprio la divergenza del campo). A questo punto, usando la definizione di derivata, si ottiene la tesi.

1.2 L'equazione di Liouville

Consideriamo una generica distribuzione di probabilità iniziale f_0 (per la precisione, f_0 è una densità di probabilità rispetto a Lebesgue). Sia $f(t, \mathbf{x})$ la densità di probabilità al tempo t . Sicuramente, per ogni dominio A , deve valere

$$\int_{A_t} d\mathbf{x} f(t, \mathbf{x}) = \int_A d\mathbf{x} f_0(\mathbf{x}) \quad (1.6)$$

Questa identità ci permette di identificare f al tempo t . Infatti, cambiando variabile di integrazione $\mathbf{x} = \Phi^{t,0}(\mathbf{y})$, (e poi chiamando di nuovo \mathbf{x} la variabile), si ha

$$\int_A d\mathbf{x} f_0(\mathbf{x}) = \int_{A_t} d\mathbf{x} f(t, \mathbf{x}) = \int_A d\mathbf{x} f(t, \Phi^{t,0}(\mathbf{x})) J(\mathbf{x}, t, 0)$$

Poiché questa identità vale per ogni insieme A , gli integrandi devono essere uguali quasi ovunque:

$$f(t, \Phi^{t,0}(\mathbf{x})) J(\mathbf{x}, t, 0) = f_0(\mathbf{x}). \quad (1.7)$$

Sto usando il lemma

Lemma 1.3. *Sia $g(\mathbf{x})$ una funzione localmente integrabile. Se per ogni misurabile A vale*

$$\int_A g(\mathbf{x}) = 0$$

allora g è nulla quasi ovunque.

La dimostrazione è semplice: sia A^+ l'insieme su cui $g > 0$; per definizione è misurabile, e, usando l'ipotesi, si ottiene che ha misura nulla (ha misura nulla $A^+ \cap B$ per ogni palla B , e dunque A^+ ha misura nulla). Analogamente si procede per $g < 0$, e si ottiene che la misura degli \mathbf{x} per cui g è non nulla è zero, cioè g è nulla quasi ovunque.

L'uguaglianza (1.7) ci dice chi è f al tempo t :

$$f(t, \mathbf{x}) = f_0(\Phi^{0,t}(\mathbf{x})) / J(\Phi^{0,t}(\mathbf{x}), t, 0).$$

Va messo in luce il fatto che (1.7) è la versione locale della conservazione della probabilità imposta nella (1.6). Consideriamo infatti $f_0(\mathbf{x}) d\mathbf{x}$ come la probabilità di scegliere il dato iniziale in un volume infinitesimo $d\mathbf{x}$ intorno a \mathbf{x} . Infatti stiamo moltiplicando il volume della base $d\mathbf{x}$ per l'altezza $f_0(\mathbf{x})$. Il flusso deforma i volumi, ma a livello infinitesimo la trasformazione è lineare ed è data da $\partial \Phi^{t,0}$. In particolare il volume diventa $J(\mathbf{x}, t, 0) d\mathbf{x}$. Dunque la (1.7) dice esattamente come conservare la probabilità infinitesima: devo cambiare l'altezza perché il volume di base è cambiato.

Resta da determinare, infine, l'equazione soddisfatta dalla densità di probabilità $f(\mathbf{x}, t)$. Deriviamo l'identità (1.7). A destra abbiamo derivata nulla, a sinistra ci sono due contributi:

$$\partial_t(f(t, \Phi^{t,0})J) + f \partial_t J = 0$$

Il primo contiene la **derivata lungo il flusso di f** :

$$\partial_t(f(t, \Phi^{t,0}(\mathbf{x}))) = (\partial_t f + \mathbf{u} \cdot \nabla f)|_{t, \Phi^{t,0}(\mathbf{x})}$$

Nel secondo compare la derivata di J che abbiamo già trovato. Mettendo insieme i due termini e passando da \mathbf{x} a $\Phi^{0,t}(\mathbf{x})$ otteniamo

$$(\partial_t f + \mathbf{u} \cdot \nabla f)J + f \operatorname{div} \mathbf{u} J = 0$$

Si noti che l'equazione per J ci dice che J non si può annullare in tempo finito (ha dato iniziale 1, ed è lineare anche se a coefficienti non costanti). Dunque f soddisfa l'equazione

$$\partial_t f + \mathbf{u} \cdot \nabla f + f \operatorname{div} \mathbf{u} = 0.$$

Il primo termine è dovuto al trasporto di f , il secondo alla variazione di volume infinitesimo. Con un po' di algebra si ottiene infine l'equazione di Liouville:

$$\partial_t f + \operatorname{div}(\mathbf{u}f) = 0. \tag{1.8}$$

L'equazione di Liouville è una **legge di conservazione** in forma di divergenza (*vedi* anche Buttà pagg. 10–11). Infatti, se f soddisfa l'equazione di Liouville, per ogni A vale

$$\frac{d}{dt} \int_A f(\mathbf{x}, t) \, d\mathbf{x} = - \int_A \operatorname{div}(\mathbf{u}f)(\mathbf{x}, t) \, d\mathbf{x} = - \int_{\partial A} f(\mathbf{x}, t) \mathbf{u}(\mathbf{x}, t) \cdot \mathbf{n} \, \sigma(d\mathbf{x})$$

Questa uguaglianza afferma che l'integrale di f in un dominio fissato varia solo perché f viene trasportata dentro o fuori dal campo \mathbf{u} . In questo senso, si dice che $\mathbf{u}f$ è la **corrente** di f .

Per esercizio, si ricavi l'identità generale $J(\mathbf{x}, t, s)f(t, \Phi^{t,s}(\mathbf{x})) = f(\mathbf{x}, s)$ mostrando che la derivata del primo membro è nulla. Si mostri anche che

$$J(\Phi^{s,t}(\mathbf{x}), t, s) J(\mathbf{x}, s, t) = 1.$$

Si provi inoltre che

$$f(t, \Phi^{t,0}(\mathbf{x}), t) = f_0(\mathbf{x}) \exp(-\operatorname{div} \mathbf{u}(\mathbf{x}, \Phi^{s,0}(\mathbf{x})))$$

1.3 Il pushforward

Rileggiamo la soluzione trovata in termini di **pushforward** di misure di probabilità, concetto che ci sarà utile nel seguito.

Supponiamo di considerare una variabile aleatoria \mathbf{X} che prende valori su V (per i nostri scopi un sottoinsieme di \mathbb{R}^d , ma può essere più generale). La sua legge è una misura di probabilità μ su V che permette di esprimere

$$\mathbb{P}(\mathbf{X} \in A) = \int_A \mu(d\mathbf{x}).$$

Inoltre, se α è una funzione continua e limitata, il valore atteso dell'osservabile α è per definizione

$$\mathbb{E}(\alpha(\mathbf{X})) = \int_V \alpha(\mathbf{x})\mu(d\mathbf{x}).$$

Ricordo che assegnare μ è equivalente ad assegnare i valori attesi di qualunque funzione $\alpha \in C_b(\mathbb{R}^d)$ (con C_b intendo le funzioni continue e limitate), perché in tal modo si individua la misura μ .

Sia ora Φ una funzione continua da V a W (per fissare le idee, un sottoinsieme di \mathbb{R}^m). In tal modo si definisce $\mathbf{Y} := \Phi(\mathbf{X})$ variabile aleatoria su W . Cerchiamo la legge di \mathbf{Y} che indicheremo con ν . Sia $\beta \in C_b(W)$. Per definizione

$$\int \beta(\mathbf{y})\nu(d\mathbf{y}) = \mathbb{E}(\beta(\mathbf{Y})) = \mathbb{E}(\beta(\Phi(\mathbf{X}))) = \int \beta(\Phi(\mathbf{x}))\mu(d\mathbf{x}).$$

Poiché le misure sono individuate dai valori degli integrali degli osservabili, la relazione precedente definisce esattamente la legge ν . Formalizzo.

Si chiama **pushforward** di μ mediante $\Phi : V \rightarrow W$ la misura su W , indicata con $\Phi\#\mu$, definita dalla seguente relazione: qualunque sia $\beta \in C_b(\mathbb{R}^m)$

$$\int_W \beta(\mathbf{y})\Phi\#\mu(d\mathbf{y}) := \int_V \beta(\Phi(\mathbf{x}))\mu(d\mathbf{x}).$$

Teorema 1.1. *La soluzione dell'equazione di Liouville di dato iniziale f_0 è il pushforward di f_0 mediante il flusso:*

$$f_t = \Phi^{t,0}\#f_0.$$

Si noti che dato il flusso, posso fare pushforward di ogni misura iniziale, non solo quelle assolutamente continue rispetto a Lebesgue, in tal caso, però, otterremo soluzioni deboli perché non possiamo dare senso puntuale a $\text{div}(\mathbf{u}f)$. Per esercizio, si provi che se μ è una misura, allora

$$\mu_t := \Phi^{t,0}\#\mu$$

risolve nel senso debole delle distribuzioni l'equazione di Liouville.

2 Flussi a divergenza nulla

Vale

Corollario 2.1.

$$\frac{d}{dt}|A_t| = \int_{A_t} \text{div} \mathbf{u}(t, \mathbf{x}) dx$$

In particolare, $\text{div} \mathbf{u} \equiv 0$ se e solo se il flusso conserva la misura.

La dimostrazione si ottiene calcolando la derivata dopo aver cambiato variabile nell'integrale

$$|A_t| = \int_{A_t} dx = \int_A J(\mathbf{x}, t) dx$$

Si noti che se il flusso di campo \mathbf{u} conserva la misura, per ogni misura μ assolutamente continua rispetto a Lebesgue, di densità f_0 , si ha

$$\Phi^{t,0}\#\mu(d\mathbf{x}) = f(t, \mathbf{x}) d\mathbf{x} = f_0(\Phi^{0,t}(\mathbf{x})) d\mathbf{x}$$

cioè f è **trasportata** da \mathbf{u} . L'equazione di Liouville diventa l'**equazione del trasporto**

$$\partial_t f + \mathbf{u} \cdot \nabla f = 0.$$

2.1 Il caso dei sistemi meccanici conservativi

I sistemi meccanici conservativi sono governati da EDO del secondo ordine, in cui l'accelerazione è pari alla forza \mathbf{F} , che non dipende dalla velocità, ed è l'opposto del gradiente dell'energia potenziale. In questo caso, la densità di probabilità nello spazio delle fasi (velocità e posizione) $f(\mathbf{x}, \mathbf{v}, t)$ soddisfa l'equazione di Liouville

$$\partial_t f(\mathbf{x}, \mathbf{v}, t) + \operatorname{div}_{\mathbf{x}}(\mathbf{v}f(\mathbf{x}, \mathbf{v}, t)) + \operatorname{div}_{\mathbf{v}}(\mathbf{F}(\mathbf{x}, t)f(\mathbf{x}, \mathbf{v}, t)) = 0$$

che è equivalente all'**equazione di trasporto**

$$\partial_t f(\mathbf{x}, \mathbf{v}, t) + \mathbf{v} \cdot \nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{v}, t) + \mathbf{F}(\mathbf{x}, t) \cdot \nabla_{\mathbf{v}} f(\mathbf{x}, \mathbf{v}, t) = 0$$

Infatti, il campo $\left(\frac{\mathbf{v}}{\mathbf{F}}\right)$ ha divergenza nulla nella coppia di variabili (\mathbf{x}, \mathbf{v}) . In altre parole, i sistemi meccanici conservano la misura nello spazio delle fasi, e l'equazione di Liouville coincide con la corrispondente equazione del trasporto (si noti che queste affermazioni non sono invarianti per la scelta di coordinate lagrangiane diverse da quelle rettangolari, ma sono sempre vere se il sistema viene descritto nel formalismo hamiltoniano, come vedremo in seguito).

Infine, il più semplice sistema meccanico è quello costituito da una particella libera. In questo caso la forza è nulla e l'equazione di Liouville si riduce a

$$\partial_t f(\mathbf{x}, \mathbf{v}, t) + \mathbf{v} \cdot \nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{v}, t) = 0$$

che è risolta da

$$f(\mathbf{x}, \mathbf{v}, t) = f_0(\mathbf{x} - t\mathbf{v}, \mathbf{v})$$

che descrive il trasporto del **flusso libero** di una distribuzione di probabilità di dati iniziali.

3 Dal metodo delle caratteristiche alle equazioni di Hamilton

3.1 Metodo delle caratteristiche

Rileggendo il paragrafo precedente “al contrario”, si prova che l'equazione alle derivate parziali

$$\partial_t f + \operatorname{div}(\mathbf{u}f) = 0$$

si risolve trovando il flusso per l'EDO $\dot{\mathbf{x}} = \mathbf{u}(\mathbf{x}, t)$. In questo contesto, chiameremo *caratteristiche* le traiettorie individuate dal flusso. Anche l'equazione del trasporto

$$\partial_t f + \mathbf{u} \cdot \nabla f = 0$$

si risolve per caratteristiche, in modo anche più semplice dell'equazione di Liouville. Infatti $f(\Phi^{t,s}(\mathbf{x}), t)$ è costante in t (esercizio). In particolare, per definizione, l'equazione del trasporto è l'equazione che è verificata dagli *integrali primi* dell'EDO considerata:

$$\frac{d}{dt}G(\mathbf{x}(t), t) = (\partial_t G + \mathbf{u} \cdot \nabla G)|_{\mathbf{x}(t), t}$$

che è 0 se e solo se G soddisfa l'equazione del trasporto.

Introduco ora un ulteriore esempio, il caso di una generica EDP del I ordine (potete leggere questo argomento sull'Evans). Indico con $\mathbf{x} \in \mathbb{R}^d$ la variabile indipendente, con $f(\mathbf{x})$ la funzione incognita, e con

$$F: \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$$

una funzione assegnata. L'EDP del I ordine che voglio risolvere è

$$F(\nabla_x f(\mathbf{x}), \mathbf{x}, f(\mathbf{x})) = 0. \quad (3.1)$$

Tentiamo di capire che informazioni dà questa equazione su f . Indico con γ le prime d variabili di F , e noto che se $\nabla_\gamma F$ è non nullo, posso al massimo sperare di ottenere il valore di una delle derivate di f in funzione delle altre, di f stessa, e di \mathbf{x} . Evidentemente questo non può bastare a identificare f , neanche nell'intorno di un punto. Supponiamo invece che f sia nota su una superficie Σ di dimensione $d-1$. Ne segue che sono anche note le componenti di ∇f tangenti a Σ . Dunque possiamo leggere la (3.1) come un'equazione per la componente normale di ∇f . Conoscendo $\mathbf{n} \cdot \nabla f$ su Σ , conosciamo la velocità di variazione di f nella direzione della normale a Σ , e dunque è possibile conoscere f fuori da Σ (si immagini per esempio uno schema iterativo approssimato). Ci si aspetta dunque un teorema di esistenza e unicità in un intorno di Σ , nota f su Σ , a meno di qualche condizione di non degenerazione. Il metodo delle caratteristiche formalizza questo auspicio, in particolare la caratteristica deve essere una traiettoria $\mathbf{x}(s)$ che esce da Σ , e che ci permette di determinare $f(\mathbf{x}(s))$, valore che chiameremo $z(s)$. D'altra parte, non conoscendo preliminarmente f , anche ∇f sarà un'incognita del nostro problema.

Il sistema alle caratteristiche è fatto dunque di tre incognite $\mathbf{x}(s)$, $\gamma(s)$, $z(s)$, che una volta risolto ci permette di determinare f attraverso l'uguaglianza $f(\mathbf{x}(s)) = z(s)$, e inoltre sarà vero che $\nabla f(\mathbf{x}(s), s) = \gamma(s)$.

Ora si tratta di costruire un sistema differenziale per $\mathbf{x}(s)$, $\mathbf{z}(s) = f(\mathbf{x}(s))$, $\gamma(s) = \nabla_x f(\mathbf{x}(s))$. Notiamo intanto che per definizione di z :

$$\dot{z}(s) = \nabla f(\mathbf{x}(s)) \cdot \dot{\mathbf{x}}(s) = \gamma(\mathbf{x}(s)) \cdot \dot{\mathbf{x}}(s). \quad (3.2)$$

Deriviamo $\gamma(s)$:

$$\dot{\gamma} = \frac{d}{ds} \nabla f(\mathbf{x}(s)) = \partial^2 f \dot{\mathbf{x}}(s). \quad (3.3)$$

D'altra parte, se $F(\nabla f, \mathbf{x}, f) = 0$, il suo gradiente in \mathbf{x} deve essere nullo, dunque

$$\partial^2 f \nabla_\gamma F + \nabla_x F + \partial_z F \nabla f = 0 \quad (3.4)$$

Si tratta ora di imporre delle equazioni per $\dot{\mathbf{x}}$ e $\dot{\gamma}$ compatibili con le (3.2), (3.3), (3.4).

Scegliamo $\dot{\mathbf{x}} = \nabla_\gamma F$; inserendo questa espressione nella (3.3), possiamo riconoscere in $\dot{\gamma}$ il primo addendo della (3.4) e ricavare quindi:

$$\dot{\gamma} = \partial^2 f \nabla_\gamma F = -\partial_x F - \partial_z F \gamma$$

Questo è il sistema differenziale delle caratteristiche associato alla EDP:

$$\begin{aligned} \dot{\mathbf{x}}(s) &= \nabla_\gamma F(\gamma(s), \mathbf{x}(s), z(s)) \\ \dot{\gamma}(s) &= -\nabla_x F(\gamma(s), \mathbf{x}(s), z(s)) - \partial_z F(\gamma(s), \mathbf{x}(s), z(s)) \gamma(s) \\ \dot{z}(s) &= \gamma(\mathbf{x}(s)) \cdot \nabla_\gamma F(\gamma(s), \mathbf{x}(s), z(s)) \end{aligned}$$

Abbiamo ottenuto un sistema di equazioni differenziali ordinarie, che con ipotesi ragionevoli su F ha soluzione unica. Vediamo come la possibilità di risolvere questo sistema ci permette di determinare f in un intorno di Σ . Fissiamo i dati iniziali per il sistema. Per ogni $\mathbf{x}_0 \in \Sigma$, sia $z_0 = f(\mathbf{x}_0)$, e sia $\boldsymbol{\gamma}_0$ un vettore tale che

$$F(\boldsymbol{\gamma}_0, \mathbf{x}_0, z_0) = 0$$

e per ogni $\boldsymbol{\tau}$ vettore tangente a Σ in \mathbf{x}_0 ,

$$\boldsymbol{\gamma}_0 \cdot \boldsymbol{\tau} = \nabla f(\mathbf{x}_0) \cdot \boldsymbol{\tau}$$

(si noti che il membro di destra è noto, perché è nota f su Σ). In particolare stiamo ipotizzando che l'equazione sia risolubile nella componente normale, e $\boldsymbol{\gamma}_0$ è proprio il gradiente di f in $\mathbf{x}_0 \in \Sigma$. Ipotizziamo inoltre regolarità continua per $f(\mathbf{x}_0)$ e $\boldsymbol{\gamma}_0(\mathbf{x}_0)$ al variare di $\mathbf{x}_0 \in \Sigma$. A questo punto risolviamo il sistema delle caratteristiche ottenendo

$$\mathbf{x}(s, \mathbf{x}_0), \boldsymbol{\gamma}(s, \mathbf{x}_0), z(s, \mathbf{x}_0)$$

(indico solo la dipendenza da \mathbf{x}_0 , perché la dipendenza da z_0 e da $\boldsymbol{\gamma}_0$ si traduce, per le imposizioni precedenti, in una dipendenza da \mathbf{x}_0). In questo modo otteniamo una funzione da $\mathbb{R} \times \Sigma$ in \mathbb{R}^d . Supponendo che su Σ valga $\boldsymbol{\gamma}_0 \cdot \mathbf{n} \neq 0$, questa funzione sarà suriettiva a valori in un intorno di Σ , come vedremo subito. Sia \mathbf{x} un punto "vicino" a Σ . Per trovare $f(\mathbf{x})$ è necessario trovare \mathbf{x}_0 e s tale che

$$\mathbf{x} = \mathbf{x}(s, \mathbf{x}_0),$$

e ottenere così $f(\mathbf{x}) = z(s, \mathbf{x}_0)$. Se $\mathbf{x} \in \Sigma$ l'equazione ha soluzione $\mathbf{x}_0 = \mathbf{x}$ e $s = 0$. Per assicurarci l'esistenza in un intorno di \mathbf{x}_0 è sufficiente usare il teorema della funzione implicita, imponendo che lo jacobiano del membro di destra sia non singolare in \mathbf{x}_0 e $s = 0$.

Consideriamo prima un caso molto semplice, quello in cui Σ è l'iperpiano $\Sigma = \{\mathbf{x} | x_d = 0\}$. Un punto su Σ è del tipo $(\mathbf{y}, 0)$, con $\mathbf{y} \in \mathbb{R}^{d-1}$. La traiettoria che parte da questo dato è dunque funzione solo di s e di \mathbf{y} , e la scrivo come $\mathbf{x}(s, \mathbf{y})$. Il suo jacobiano rispetto a Per $s = 0$, $\mathbf{x}(s, \mathbf{y}) = \mathbf{x}(0, \mathbf{y}) = \mathbf{y}$, dunque le derivate rispetto a \mathbf{y} delle prime $d - 1$ componenti sono date dalla matrice identica $(d - 1) \times (d - 1)$. La d -esima componente di \mathbf{x} è $x_d(s, \mathbf{y})$ che in $s = 0$ è 0, e dunque ha gradiente nullo in \mathbf{y} . Infine, la derivata in s calcolata in $s = 0$ è $\nabla_{\boldsymbol{\gamma}} F$. Dunque lo jacobiano di $\mathbf{x}(s, \mathbf{y})$ nelle variabili (s, \mathbf{y}) calcolato per $s = 0$ è

$$\begin{pmatrix} \partial_{\boldsymbol{\gamma}_1} F & 1 & 0 & \dots & 0 \\ \partial_{\boldsymbol{\gamma}_2} F & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \partial_{\boldsymbol{\gamma}_{d-1}} F & 0 & 0 & \dots & 1 \\ \partial_{\boldsymbol{\gamma}_d} F & 0 & 0 & \dots & 0 \end{pmatrix}$$

Il suo determinante a meno del segno è $\partial_{\boldsymbol{\gamma}_d} F$. Quindi l'invertibilità è garantita intorno a Σ se $\partial_{\boldsymbol{\gamma}_d} F \neq 0$.

Consideriamo ora il caso in Σ non sia l'iperpiano, ma sia tangente all'iperpiano nel punto $(\mathbf{y}_0, 0)$. I punti su Σ possono essere individuati, localmente intorno al punto di tangenza, da una funzione $\mathbf{y} \rightarrow \mathbf{w}(\mathbf{y})$ con $w_n(\mathbf{y}_0) = 0$. La condizione di tangenza impone che $\partial_{y_i} w_n(\mathbf{y}_0) = 0$ per ogni $i = 1, \dots, d - 1$. Consideriamo ora le caratteristiche $\mathbf{x}(s, \mathbf{y})$ dove \mathbf{y} individua il dato iniziale mediante la funzione $\mathbf{w}(\mathbf{y})$. Poiché $\mathbf{x}(0, \mathbf{y}) = (\mathbf{y}, w_n(\mathbf{y}))$, anche in questo caso le

derivate in \mathbf{y} delle prime $d-1$ componenti sono date dalla matrice identica. La condizione di tangenza permette di concludere che le derivate della d -esima componente sono tutte nulle, tranne quella in s che è di nuovo $\partial_{\gamma_d} F$. Quindi la condizione di invertibilità locale intorno a $(\mathbf{y}_0, 0)$ è sempre $\partial_{\gamma_d} F$.

Possiamo sempre ridurci a questo caso in ogni punto di Σ considerando un sistema di coordinate cartesiane con Σ tangente all'iperpiano delle prime $d-1$ coordinate e x_d coordinata normale a Σ nel punto. In conclusione, la condizione di invertibilità è $\mathbf{n} \cdot \nabla_{\gamma} F \neq 0$ nei punti di Σ .

Costruita f , per provare l'esistenza della soluzione rimane da provare che effettivamente $\gamma = \nabla f$. Ometto la prova, che richiede un po' di lavoro e potete trovare sull'Evans [EV].

3.2 L'equazione di Hamilton-Jacobi e le equazioni di Hamilton

Specializzo l'analisi precedente a un caso particolare. Indico con S la funzione incognita e con $\mathbf{x} = (t, \mathbf{q})$, con $\mathbf{q} \in \mathbb{R}^n$ le variabili. Suppongo inoltre che la superficie Σ coincida con l'iperpiano $t = 0$. La condizione di non degenerazione permette di esplicitare $\partial_t S$, dunque l'equazione assume la forma

$$\partial_t S(t, \mathbf{q}) + H(t, \mathbf{q}, \nabla_{\mathbf{q}} S(t, \mathbf{q}), S(t, \mathbf{q})) = 0.$$

Semplifico inoltre il problema ipotizzando che H non dipenda da S . In questo caso, H prende il nome di hamiltoniana, e l'equazione prende il nome di **equazione di Hamilton-Jacobi**.

Vediamo come si adatta la costruzione delle caratteristiche a questo caso. Procedendo come nel paragrafo precedente, ma tenendo presente il ruolo particolare della variabile t . Dunque definiamo

$$\begin{aligned} \mathbf{x}(s) &= (t(s), \mathbf{q}(s)) \\ \gamma(s) &= (r(s), \mathbf{p}(s)) \\ z(s) &= S(t(s), \mathbf{q}(s)) \end{aligned}$$

dove r è la prima componente di γ , e corrisponde alla derivata temporale di S . La funzione $F(\gamma, \mathbf{x}, \mathbf{z})$ con $\gamma = (r, \mathbf{p})$, $\mathbf{x} = (t, \mathbf{q})$ è

$$F = r + H(\mathbf{q}, \mathbf{p}, t).$$

L'equazione per la variabile t è

$$\dot{t}(s) = 1$$

dunque possiamo identificare s con t . Inoltre possiamo notare che le equazioni per \mathbf{q} e \mathbf{p} formano un sistema differenziale in $\mathbb{R}^d \times \mathbb{R}^d$:

$$\begin{cases} \dot{\mathbf{q}}(t) = \partial_{\mathbf{p}} H(\mathbf{q}(t), \mathbf{p}(t), t) \\ \dot{\mathbf{p}}(t) = -\partial_{\mathbf{q}} H(\mathbf{q}(t), \mathbf{p}(t), t) \end{cases}$$

Queste sono le **equazioni di Hamilton** per l'hamiltoniana H .

L'ultima equazione per il sistema delle caratteristiche è quella per $\dot{z}(t)$. Poiché $z(t) = S(\mathbf{q}(t), t)$, possiamo ricavare direttamente l'equazione, usando la definizione di \mathbf{p} e l'equazione di HJ:

$$\frac{d}{dt} S(\mathbf{q}(t), t) = \nabla_{\mathbf{q}} S(\mathbf{q}(t), t) \cdot \dot{\mathbf{q}}(t) + \partial_t S(\mathbf{q}(t), t) = \mathbf{p}(t) \cdot \dot{\mathbf{q}}(t) - H(\mathbf{q}(t), \mathbf{p}(t), t),$$

da cui

$$S(\mathbf{q}(t), t) = S(\mathbf{q}(t), 0) + \int_0^t (\mathbf{p} \cdot \dot{\mathbf{q}} - H) dt.$$

Nel contesto dei sistemi hamiltoniani, la funzione S viene detta **azione**.

Un commento importante: lo scopo iniziale era risolvere una EDP per l'azione, e per farlo ho costruito un sistema di EDO, che una volta risolto permette di trovare l'azione. Però, come nel caso generale, l'EDP è corredata di un dato iniziale, la funzione $S(\mathbf{q}, 0)$. Il sistema differenziale definisce il flusso sullo spazio delle \mathbf{q}, \mathbf{p} , che è in $\mathbb{R}^n \times \mathbb{R}^n$. Per risolvere realmente l'EPD bisogna considerare solo l'evoluzione dei dati

$$(\mathbf{q}, \mathbf{p}) = (\mathbf{q}_0, \partial_q S(\mathbf{q}_0, 0)).$$

Indicando con $\mathbf{q}(t, \mathbf{q}_0)$ la componente \mathbf{q} della soluzione delle equazioni di Hamilton, l'azione verifica

$$S(\mathbf{q}(t, \mathbf{q}_0), t) = S(\mathbf{q}_0, 0) + \int_0^t (\mathbf{p} \cdot \dot{\mathbf{q}} - H) dt$$

dove l'integrando è calcolato sulla soluzione delle equazioni di Hamilton. Per trovare S al tempo t è necessario che $\mathbf{q}(t, \mathbf{q}_0)$ sia invertibile in \mathbf{q}_0 . Non è difficile trovare una condizione sufficiente per l'invertibilità a t piccolo. Infatti, per definizione di derivata, e per le equazioni di Hamilton

$$\mathbf{q}(t, \mathbf{q}_0) = \mathbf{q}_0 + t \partial_p H(\mathbf{q}_0, \partial_q S(\mathbf{q}_0, 0)) + o(t^2),$$

dunque per t piccolo si ha l'invertibilità attraverso il teorema della funzione implicita.

3.3 Prime proprietà dei sistemi hamiltoniani

Le equazioni di Hamilton sono un sistema di EDO in forma normale, nello spazio delle fasi $\mathbb{R}^d \times \mathbb{R}^d$. È facile verificare che si tratta di un sistema a divergenza nulla, dunque i flussi hamiltoniani (cioè i flussi dei sistemi hamiltoniani) conservano la misura nello spazio delle fasi. Ne segue anche che l'equazione di Liouville e l'equazione del trasporto associate al flusso.

Nel caso autonomo, cioè se H non dipende esplicitamente dal tempo, H stessa è un integrale primo del moto, e viene chiamata anche "energia" o "energia generalizzata" a seconda dei contesti. Infatti, derivando e usando le equazioni di Hamilton

$$\frac{d}{dt} H(\mathbf{q}, \mathbf{p}) = \partial_q H \cdot \dot{\mathbf{q}} + \partial_p H \cdot \dot{\mathbf{p}} = 0.$$

Questa proprietà rende particolarmente semplice studiare i sistemi a un grado di libertà, cioè quelli con \mathbf{q}, \mathbf{p} che variano in \mathbb{R} .

Riassumo le proprietà che illustro a lezione,

- a. Le orbite vivono sulle curve di livello di H
- b. I punti stazionari del sistema sono tutti e soli i punti singolari di H , cioè i punti in cui il suo gradiente è nullo. In particolare, i punti di massimo e minimo isolati sono stabili, e in genere gli altri punti sono instabili.

c. Il moto si può **ridurre alle quadrature**, cioè è risolubile in forma implicita attraverso integrali e inversioni.

Lascio al lettore la prova di **a** e **b** (si noti che H è una funzione di Lyapunov per il sistema). Provo **c**. Consideriamo il dato iniziale (q_0, p_0) di energia $H_0 = H(q_0, p_0)$, e supponiamo che il gradiente di H nel dato iniziale sia non nullo, in particolare $\partial_p H(q_0, p_0) \neq 0$. Per il teorema della funzione implicita, in un intorno del dato iniziale si può ricavare $p = p(H_0, q)$. L'equazione del moto in q è dunque

$$\dot{q}(t) = \partial_p H(q(t), p(t)) = \partial_p H(q(t), p(H_0, q(t)))$$

che è un'equazione autonoma in una variabile, risolubile per quadrature

$$t = \int_{q_0}^{q(t)} \frac{dq}{\partial_p H(q, p(H_0, q))}$$

(adattare la dimostrazione nel caso in cui $\partial_p H = 0$ ma $\partial_q H \neq 0$). Si rifletta sulla differenza qualitativa e quantitativa della risoluzione di un moto per quadrature o attraverso i teoremi di esistenza (e connessi schemi numerici).

Un altro commento importante. Localmente, in un punto in cui il campo non è nullo, tutti i sistemi differenziali automi sono qualitativamente indistinguibili, e localmente esiste un integrale primo del moto e il sistema si muove sulle sue curve di livello. Le differenze si notano nel caso di una analisi "globale". Si confrontino per esempio un sistema hamiltoniano e un sistema hamiltoniano con attrito intorno a un punto di equilibrio instabile: non esiste un integrale primo del moto nel caso con attrito.

D'altra parte l'esistenza dell'integrale primo permette di portare il moto alle quadrature, che sembra una proprietà algebrica (e in realtà lo è). Vedremo come si realizza la possibilità di portare i moti hamiltoniani alle quadrature e le conseguenze qualitative sul moto.

Si provi per esercizio che

a. Il sistema differenziale autonomo

$$\begin{aligned}\dot{q} &= f(q, p) \\ \dot{p} &= g(q, p)\end{aligned}$$

conserva la misura se e solo se è (localmente) hamiltoniano, cioè se il campo è a divergenza nulla.

b. Supponi che il campo non sia a divergenza nulla e sia non nullo. Mostra che se esiste un integrale primo non banale $u(q, p)$, allora la funzione

$$\alpha(q, p) = \partial_q u / g = -\partial_p u / f$$

è ben definita. Mostra che il campo $(\alpha f, \alpha g)$ è localmente hamiltoniano. In questo senso, l'esistenza di integrali primi non banali è legata all'hamiltonianità del sistema. Osserva inoltre che l'esistenza di u integrale primo garantisce la riduzione alle quadrature del moto.

Nota che la condizione di divergenza nulla corrisponde alla chiusura della forma differenziale

$$-g dq + f dp.$$

Una funzione α tale che la forma

$$\alpha(-g dq + f dp)$$

è chiusa, è detta **fattore integrante** della forma differenziale.

3.4 Un principio variazionale per le equazioni di Hamilton

Teorema 3.1. Principio variazionale per le equazioni di Hamilton, II forma

Il moto hamiltoniano di hamiltoniana H rende stazionario il funzionale di azione

$$\mathcal{S} = \int_0^T (\mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{q}, \mathbf{p}, t)) dt$$

nello spazio dei moti

$$M_{II} = \{(\mathbf{q}(t), p(\mathbf{t})), t \in [0, T] | (\mathbf{q}(0), \mathbf{p}(0)) = (\mathbf{q}_0, \mathbf{p}_0), (\mathbf{q}(T), \mathbf{p}(T)) = (\mathbf{q}_1, \mathbf{p}_1)\}.$$

La verifica è semplice. Sia $\delta\mathbf{q}$, $\delta\mathbf{p}$ una variazione del moto che soddisfa le condizioni iniziali e finali, cioè sia nulla per $t = 0$ e $t = T$. La variazione dell'azione è

$$\delta\mathcal{S} = \int_0^T (\delta\mathbf{p} \cdot \dot{\mathbf{q}} + \mathbf{p} \cdot \delta\dot{\mathbf{q}} - \partial_{\mathbf{q}}H \cdot \delta\mathbf{q} - \partial_{\mathbf{p}}H \cdot \delta\mathbf{p}) dt$$

Integrando per parti il termine in $\delta\dot{\mathbf{q}} = \frac{d}{dt}\delta\mathbf{q}$ si ottiene

$$\delta\mathcal{S} = \int_0^T (\dot{\mathbf{q}} - \partial_{\mathbf{p}}H) \cdot \delta\mathbf{p} dt - \int_0^T (\dot{\mathbf{p}} + \partial_{\mathbf{q}}H) \cdot \delta\mathbf{q} dt$$

Dunque $\delta\mathcal{S}$ è nulla per ogni variazione del moto che soddisfi i dati iniziali se e solo se valgono le equazioni di Hamilton.

Nel principio variazionale per le equazioni di Hamilton abbiamo fissato sia i momenti che gli impulsi al tempo 0 e al tempo T . Ma se si fissa (\mathbf{q}, \mathbf{p}) al tempo 0, esiste una sola soluzione, dunque non si può fissare anche (\mathbf{q}, \mathbf{p}) al tempo T . D'altra parte, è facile verificare che per ottenere le equazioni di Hamilton abbiamo integrato per parti il solo termine in $\delta\dot{\mathbf{q}}$, dunque non è necessario fissare \mathbf{p} al bordo, ma basta fissare solo \mathbf{q} . Possiamo dunque riformularlo in modo più corretto.

Teorema 3.2. Principio variazionale per le equazioni di Hamilton, I forma

Il moto hamiltoniano di hamiltoniana H rende stazionaria il funzionale di azione

$$\mathcal{S} = \int_0^T (\mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{q}, \mathbf{p}, t)) dt$$

nello spazio dei moti

$$M_I = \{(\mathbf{q}(t), p(\mathbf{t})), t \in [0, T] | \mathbf{q}(0) = \mathbf{q}_0, \mathbf{q}(T) = \mathbf{q}_1\}.$$

Il principio variazione permette di comprendere il legame tra formalismo hamiltoniano e formalismo lagrangiano. Osserviamo intanto che cercare i punti stazionari di una funzione $F(\mathbf{x}, \mathbf{y})$ può essere fatto “in sequenza”. Imponendo che $\partial_x F = 0$ si ottiene un’equazione per \mathbf{x} , che ipotizziamo sia risolvibile per ogni \mathbf{y} , e dunque otteniamo $\mathbf{x}(\mathbf{y})$. Consideriamo ora la funzione $G(\mathbf{y}) = F(\mathbf{x}(\mathbf{y}), \mathbf{y})$. Se troviamo i suoi punti stazionari avremo anche trovato quello di F (controllate i dettagli per esercizio).

Usiamo questa procedura per \mathcal{S} , imponendo prima che la variazione in \mathbf{p} sia nulla. Si ottiene ovviamente $\dot{\mathbf{q}} = \partial_p H(\mathbf{q}, \mathbf{p}, t)$. Supponiamo di poter risolvere questa equazione in \mathbf{p} , cioè di poter trovare una funzione regolare $\mathbf{p}(\mathbf{q}, \dot{\mathbf{q}}, t)$ che la rende identicamente verificata. Possiamo ora definire la funzione L come

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = \mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{q}, \mathbf{p}, t),$$

dove \mathbf{p} è la funzione di \mathbf{q} , $\dot{\mathbf{q}}$ e t che abbiamo trovato prima. Resta dunque da stazionarizzare il funzionale del solo moto $\mathbf{q}(t)$ dato da

$$\mathcal{S}_L = \int_0^T L(\mathbf{q}, \dot{\mathbf{q}}, t) dt$$

nello spazio dei moti

$$M_L = \{\mathbf{q}(t), t \in [0, T] | \mathbf{q}(0) = \mathbf{q}_0, \mathbf{q}(T) = \mathbf{q}_1\}.$$

Questo funzionale è esattamente il funzionale di azione per un sistema lagrangiano, ed è stazionarizzato dalle soluzioni delle corrispondenti equazioni di Eulero-Lagrange:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}.$$

Abbiamo dunque dimostrato il seguente fatto.

Teorema 3.3. Sistemi hamiltoniani e sistemi lagrangiani - I

Se l’equazione $\dot{\mathbf{q}} = \partial_p H(\mathbf{q}, \mathbf{p}, t)$ può essere risolta \mathbf{p} come funzione regolare delle altre variabili, il moto $(\mathbf{q}(t), \mathbf{p}(t))$ risolve le equazioni di Hamilton di hamiltoniana H se e solo se il moto $\mathbf{q}(t)$ risolve le equazioni di Eulero-Lagrange di lagrangiana $L = \mathbf{p} \cdot \dot{\mathbf{q}} - H$.

Per discutere il viceversa, cioè associare un sistema hamiltoniano a uno lagrangiano, è il caso introdurre la **trasformata di Legendre** (vedi Arnold pp 65-66).

Data una funzione $f(\mathbf{p})$ si definisce la sua trasformata di Legendre

$$f^*(\mathbf{x}) = \sup_{\mathbf{p}} (\mathbf{p} \cdot \mathbf{x} - f(\mathbf{p}))$$

Poiché $\mathbf{x} \rightarrow \mathbf{p} \cdot \mathbf{x} - f(\mathbf{p})$ è una famiglia di funzioni convesse in \mathbf{x} , indicizzata in \mathbf{p} , la funzione $f^*(\mathbf{x})$ è convessa (il sup di funzioni convesse è una funzione convessa). Si può provare che se f è strettamente convessa lo è anche f^* e che f^{**} è proprio f (cioè la trasformata di Legendre è involutiva). Inoltre, per definizione, vale la disuguaglianza di Young

$$\mathbf{x} \cdot \mathbf{p} \leq f(\mathbf{p}) + f^*(\mathbf{x})$$

Per esercizio, si calcoli la trasformata di Legendre di p^a/a , con $a > 1$ (per assicurare la convessità) e si riconosca la disuguaglianza di Young nota dalla teoria degli spazi L^p .

Nel caso di f regolare, il sup è un massimo, che si ottiene in \mathbf{p} tale che

$$\mathbf{x} = \partial_{\mathbf{p}} f(\mathbf{p})$$

e \mathbf{p} è univocamente determinato da questa espressione.

Applicando questa definizione al caso dell'azione, si comprende come, data l'hamiltoniana H , la lagrangiana altro non è che la sua trasformata di Legendre: $L = H^*$. Sia ora data una lagrangiana L , ipotizzando che verifichi le condizioni per cui la trasformata di Legendre in $\dot{\mathbf{q}}$ sia involutiva: definiamo $H = L^*$ e usiamo il teorema precedente notando che $H^* = L^{**} = L$. Otteniamo così il seguente teorema.

Teorema 3.4. Sistemi hamiltoniani e sistemi lagrangiani - II

*Sia data L tale che $L^{**} = L$. Allora il moto $\mathbf{q}(t)$ risolve le equazioni di Eulero-Lagrange di lagrangiana L se e solo se il moto (\mathbf{q}, \mathbf{p}) , dove $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$, risolve il moto di hamiltoniana $H = L^*$.*

Un'osservazione: si mostra che per t piccoli, i punti stazionari dell'azione per lagrangiane associate a sistemi meccanici sono minimi dell'azione. Ne segue che i punti stazionari dell'azione associata a un sistema hamiltoniano sono dei punti di mini-max (massimi in \mathbf{p} , minimi di \mathbf{q}).

3.5 Dalla lagrangiana all'hamiltoniana, versione differenziale

Per gli scopi di questo paragrafo, è importante evidenziare la distinzione tra le variabili in cui viene descritto il moto (cioè $\mathbf{q}(t)$ e la sua derivata temporale $\dot{\mathbf{q}}(t)$) e le variabili in cui è definita la lagrangiana. Dopo questo paragrafo, tornerò ad indicare con $\dot{\mathbf{q}}$ sia la velocità del moto $\mathbf{q}(t)$, sia la variabile nella lagrangiana.

Sia dunque $L = L(\mathbf{q}, \boldsymbol{\eta}, t)$, con $\mathbf{q} \in \mathbb{R}^n$, $\boldsymbol{\eta} \in \mathbb{R}^n$, $t \in \mathbb{R}$. Le equazioni di Eulero-Lagrange associate al L sono:

$$\frac{d}{dt} \left[\frac{\partial L}{\partial \boldsymbol{\eta}} \Big|_{\mathbf{q}(t), \dot{\mathbf{q}}(t), t} \right] = \frac{\partial L}{\partial \mathbf{q}} \Big|_{\mathbf{q}(t), \dot{\mathbf{q}}(t), t} \quad (3.5)$$

Queste equazioni sono un sistema di equazioni del secondo ordine, in forma non esplicita, nella variabile $\mathbf{q}(t) \in \mathbb{R}^n$ (non esplicita vuol dire che non è del tipo $\ddot{\mathbf{q}} = \dots$).

L'hamiltoniana è la trasformata di Legendre di L nella variabile $\boldsymbol{\eta}$:

$$H = \boldsymbol{\eta} \cdot \mathbf{p} - L(\mathbf{q}, \boldsymbol{\eta}, t).$$

Il suo differenziale è

$$dH = \boldsymbol{\eta} \cdot d\mathbf{p} + \mathbf{p} \cdot d\boldsymbol{\eta} - \partial_{\mathbf{q}} L \cdot d\mathbf{q} - \partial_{\boldsymbol{\eta}} L \cdot d\boldsymbol{\eta} - \partial_t L dt$$

Poiché $\mathbf{p} = \partial_{\boldsymbol{\eta}} L$, i termini in $d\boldsymbol{\eta}$ si cancellano, e si ottiene

$$dH = \boldsymbol{\eta} \cdot d\mathbf{p} - \partial_{\mathbf{q}} L \cdot d\mathbf{q} - \partial_t L dt$$

da cui

$$\begin{aligned} \partial_{\mathbf{q}} H &= -\partial_{\mathbf{q}} L \\ \partial_{\mathbf{p}} H &= \boldsymbol{\eta} \\ \partial_t H &= -\partial_t L \end{aligned}$$

Usando le prime due uguaglianze, si ottiene che le equazioni di Eulero-Lagrange, sono equivalenti alle equazioni di Hamilton.

3.6 Lagrangiane e hamiltoniane naturali

Una lagrangiana che si ottiene da un sistema fisico conservativo in un sistema di riferimento inerziale, con forze puramente posizionali e vincoli perfetti bilateri è sempre del tipo

$$L(\dot{\mathbf{q}}, \mathbf{q}) = \frac{1}{2} \dot{\mathbf{q}} \cdot T(\mathbf{q}) \dot{\mathbf{q}} - V(\mathbf{q}),$$

dove $T(\mathbf{q})$ è una matrice simmetrica e definita positiva, e $\mathbf{q}, \dot{\mathbf{q}}$ sono in \mathbb{R}^n . Chiamerò lagrangiane *naturali* quelle di questa forma.

Il passaggio all'hamiltoniana è semplice. Infatti il vettore degli impulsi coniugati è dato da:

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}} = T(\mathbf{q}) \dot{\mathbf{q}},$$

ed essendo T definita positiva, in particolare è invertibile. Dunque

$$\dot{\mathbf{q}} = T(\mathbf{q})^{-1} \mathbf{p}.$$

Ma allora l'hamiltoniana è data da:

$$H = \mathbf{p} \cdot \dot{\mathbf{x}} - \frac{1}{2} \dot{\mathbf{q}} \cdot T(\mathbf{q}) \dot{\mathbf{q}} + V(\mathbf{q}) = \mathbf{p} \cdot T(\mathbf{q})^{-1} \mathbf{p} - \frac{1}{2} (T(\mathbf{q})^{-1} \mathbf{p}) \cdot T(\mathbf{q}) T(\mathbf{q})^{-1} \mathbf{p} + V(\mathbf{q}) = \frac{1}{2} \mathbf{p} \cdot T(\mathbf{q})^{-1} \mathbf{p} + V(x).$$

Quindi per il calcolo dell'hamiltoniana è sufficiente calcolare l'inversa della matrice T .

Il caso dei vincoli olonomi dipendenti dal tempo è un po' diverso. Se $\mathbf{x} \in \mathbb{R}^m$ è la configurazione non vincolata e $\mathbf{q} \in \mathbb{R}^n$ sono le coordinate vincolari, la lagrangiana si trova a partire da

$$\frac{1}{2} \dot{\mathbf{x}}^2 - V(\mathbf{x})$$

usando che

$$\mathbf{x} = \mathbf{x}(\mathbf{q}, t), \quad \dot{\mathbf{x}} = \partial_t \mathbf{x}(\mathbf{q}, t) + \partial_{\mathbf{q}} \mathbf{x}(\mathbf{q}) \dot{\mathbf{q}}$$

(per semplicità ho considerato la matrice cinetica unitaria nelle coordinate \mathbf{x} , cioè masse unitarie). Sostituendo, si ottiene

$$L = \frac{1}{2} \partial_{\mathbf{q}} \mathbf{x}^t \partial_{\mathbf{q}} \mathbf{x} \dot{\mathbf{q}} \cdot \dot{\mathbf{q}} + \partial_{\mathbf{q}} \mathbf{x}^t \partial_t \mathbf{x} \cdot \dot{\mathbf{q}} + \frac{1}{2} \partial_t \mathbf{x} \cdot \partial_t \mathbf{x} - V(\mathbf{x})$$

Dunque la lagrangiana è della forma

$$L = \frac{1}{2} T(\mathbf{q}, t) \dot{\mathbf{q}} \cdot \dot{\mathbf{q}} + \mathbf{b}(\mathbf{q}, t) \cdot \dot{\mathbf{q}} - U(\mathbf{q}, t) \quad (3.6)$$

È da notare che in alcuni casi, anche se il vincolo dipende dal tempo, la lagrangiana nelle coordinate vincolari può non dipendere dal tempo; è questo il caso di vincoli in rotazione uniforme intorno a un asse, in cui compaiono termini dovuti alle "forze apparenti".

Anche la lagrangiana per il moto di una particella di massa m e carica e , in un campo elettromagnetico di potenziale $V(\mathbf{x}, t)$ e di potenziale vettore $\mathbf{A}(\mathbf{x}, t)$ ha questa forma, infatti è

$$L = \frac{1}{2} \dot{\mathbf{x}}^2 + \frac{e}{c} \mathbf{A}(\mathbf{x}, t) \cdot \dot{\mathbf{x}} - eV(\mathbf{x}, t) \quad (3.7)$$

dove c è la velocità della luce. Infatti, il moto è governato dall'equazione

$$m\ddot{\mathbf{x}} = e\mathbf{E} + \frac{e}{c} \dot{\mathbf{x}} \wedge \mathbf{B}$$

dove (\mathbf{E}, \mathbf{B}) è il campo elettromagnetico \wedge è il prodotto vettoriale, e c è la velocità della luce.

Esercizio 1. Particella carica

Verificare che, nel caso di campi indipendenti dal tempo, se $\nabla V = -\mathbf{E}$ e $\nabla \times \mathbf{A} = \mathbf{B}$, le equazioni di Eulero-Lagrange per (3.7) coincidono con l'equazione di Newton. Si estenda al caso di campi dipendenti dal tempo, caso in cui $\mathbf{E} = -\nabla V - \frac{1}{c} \partial_t \mathbf{A}$.

Esercizio 2. Hamiltoniana per la particella carica

Per esercizio, si provi che se la lagrangiana è data da (3.6) allora

$$\mathbf{p} = T\dot{\mathbf{q}} + \mathbf{b}$$

e

$$H = \frac{1}{2}(\mathbf{p} - \mathbf{b}) \cdot T^{-1}(\mathbf{p} - \mathbf{b}) + U$$

In particolare, si scriva l'hamiltoniana per il moto della particella carica.

3.7 Variabili cicliche e riduzione dei gradi di libertà

Mostrerò con un esempio il diverso comportamento dei sistemi lagrangiani e di quelli hamiltoniani in presenza di variabili cicliche. Consideriamo la lagrangiana del moto centrale piano

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}\dot{\mathbf{x}}^2 - V(|\mathbf{x}|)$$

con $V(r) = \frac{1}{r}$, le equazioni del moto sono le corrispondenti equazioni di Eulero-Lagrange:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{x}}} = \ddot{\mathbf{x}} = \frac{\partial L}{\partial \mathbf{x}} = -\nabla V$$

Per ottenere le equazioni in coordinate polari è sufficiente considerare il cambiamento di coordinate

$$x_1 = \rho \cos \vartheta \quad x_2 = \rho \sin \vartheta$$

che genera il corrispondente cambiamento di variabili nelle velocità:

$$\dot{x}_1 = \dot{\rho} \cos \vartheta - \rho \dot{\vartheta} \sin \vartheta \quad \dot{x}_2 = \dot{\rho} \sin \vartheta + \rho \dot{\vartheta} \cos \vartheta$$

Infine si calcola la lagrangiana nelle nuove variabili. Si ottiene

$$L(\rho, \vartheta, \dot{\rho}, \dot{\vartheta}) = \frac{1}{2}\dot{\rho}^2 + \frac{1}{2}\rho^2\dot{\vartheta}^2 - V(\rho)$$

Le equazioni del moto in coordinate polari sono esattamente le equazioni di Eulero Lagrange che si ottengono da questa Lagrangiana:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\rho}} = \ddot{\rho} = \frac{\partial L}{\partial \rho} = \rho \dot{\vartheta}^2 - V'(\rho)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\vartheta}} = \frac{d}{dt}(\rho^2 \dot{\vartheta}) = 0$$

La seconda equazione indica che il $\rho^2 \dot{\vartheta}$, il **momento coniugato** alla variabile ϑ , si conserva lungo il moto (infatti ϑ è una variabile ciclica, cioè L non dipende esplicitamente da ϑ).

Si può trarre vantaggio dalla conservazione di questa quantità, riducendo il sistema a un solo grado di libertà, sostituendo il momento con una costante nell'espressione dell'energia meccanica (si riveda, sui testi di Meccanica, come si porta alle quadrature il moto centrale). Noto che per ottenere questa riduzione si esce dal formalismo lagrangiano (non si può infatti sostituire il momento dentro la lagrangiana, verrebbero equazioni errate).

La corrispondente hamiltoniana è

$$H = \frac{1}{2}p_\rho^2 + \frac{1}{2\rho^2}p_\vartheta^2 + V(\rho)$$

dove $p_\rho = \partial_{\dot{\rho}}L = \dot{\rho}$ è il momento coniugato alla variabile ρ e $p_\vartheta = \partial_{\dot{\vartheta}}L = \rho^2\dot{\vartheta}$ è il momento coniugato alla variabile ϑ .

Le equazioni di Hamilton corrispondenti sono

$$\begin{aligned}\dot{\rho} &= \frac{\partial H}{\partial p_\rho} = p_\rho \\ \dot{p}_\rho &= -\frac{\partial H}{\partial \rho} = p_\vartheta^2/\rho^3 - V'(\rho) \\ \dot{\vartheta} &= \frac{\partial H}{\partial p_\vartheta} = p_\vartheta/\rho^2 \\ \dot{p}_\vartheta &= -\frac{\partial H}{\partial \vartheta} = 0\end{aligned}$$

Questo sistema di 4 equazioni è un sistema a due gradi di libertà, con un variabile ciclica, infatti H non dipende da ϑ . Il corrisponde impulso p_ϑ si conserva, come afferma l'ultima equazione. Ma allora, le prime due equazioni, in ρ e p_ρ , sono un sistema hamiltoniano a un solo grado di libertà, in cui l'impulso p_ϑ è un parametro. Il fatto che la variabile ϑ sia ciclica, ha dunque una conseguenza importante: le altre equazioni sono automaticamente le equazioni del moto di un sistema con un grado di libertà in meno. Questo è un fatto generale: nel formalismo hamiltoniano, a ogni variabile ciclica corrisponde la riduzione del sistema di un grado di libertà, e non sono necessari passaggi ulteriori rispetto alla scrittura delle equazioni del moto.

Ricordo che questo non accade nel formalismo Lagrangiano: la ciclicità di una variabile garantisce la conservazione del momento coniugato, ma la riduzione di un grado di libertà non è contenuta nel formalismo.

Considero, come ulteriore esempio, la lagrangiana della trottola pesante

$$L = \frac{I}{2}(\dot{\vartheta}^2 + \dot{\phi}^2 \sin^2 \vartheta) + \frac{J}{2}(\dot{\psi} + \dot{\phi} \cos \theta)^2 - mgl \cos \theta$$

dove ϑ è l'angolo tra l'asse della trottola e l'asse verticale, ϕ è un angolo che esprime la rotazione intorno all'asse verticale, ψ è un angolo che esprime la rotazione intorno all'asse della trottola; J è il momento di inerzia rispetto all'asse della trottola, I è quello rispetto a un qualunque asse ortogonale che passa per il punto di appoggio, l è la distanza del baricentro dal punto di appoggio.

L'energia cinetica è

$$\frac{1}{2} \begin{pmatrix} \dot{\vartheta} \\ \dot{\phi} \\ \dot{\psi} \end{pmatrix} \cdot \begin{pmatrix} I & 0 & 0 \\ 0 & I \sin^2 \vartheta + J \cos^2 \vartheta & J \cos \theta \\ 0 & J \cos \theta & J \end{pmatrix} \begin{pmatrix} \dot{\vartheta} \\ \dot{\phi} \\ \dot{\psi} \end{pmatrix}$$

L'inversa della matrice cinetica è

$$\frac{1}{IJ \sin^2 \vartheta} \begin{pmatrix} J \sin^2 \vartheta & 0 & 0 \\ 0 & J & -J \cos \theta \\ 0 & -J \cos \theta & I \sin^2 \vartheta + J + J \cos^2 \vartheta \end{pmatrix}$$

dunque l'hamiltoniana è

$$\begin{aligned} H &= \frac{1}{2I} p_\vartheta^2 + \frac{1}{2IJ \sin^2 \vartheta} (J p_\phi^2 - 2J \cos \varphi p_\phi p_\psi + (I \sin^2 \vartheta + J \cos^2 \vartheta) p_\psi^2) + mgl \cos \vartheta \\ &= \frac{1}{2I} p_\vartheta^2 + \frac{1}{2I \sin^2 \vartheta} (p_\phi - p_\psi \cos \vartheta)^2 + \frac{1}{2J} p_\psi^2 + mgl \cos \vartheta \end{aligned}$$

Anche in questo caso, si può considerare questa come l'hamiltoniana di un sistema a un grado di libertà, in cui p_ϕ e p_ψ sono integrali primi fissati dai dati iniziali.

4 Trasformazioni canoniche e simplettiche

4.1 Trasformazioni di coordinate

Le equazioni di Eulero-Lagrange sono equazioni del secondo ordine in n variabili, e sono **invarianti in forma**: se $L = L(\mathbf{q}, \dot{\mathbf{q}}, t)$ è la lagrangiana e $\tilde{\mathbf{q}} = \tilde{\mathbf{q}}(\mathbf{q}, t)$ sono delle nuove variabili, allora il moto nelle variabili $\tilde{\mathbf{q}}$ è governato dalle equazioni di Eulero-Lagrange per la lagrangiana \tilde{L} , che è esattamente la lagrangiana L scritta nelle nuove variabili, tenendo conto che

$$\dot{\tilde{\mathbf{q}}} = \partial_{\mathbf{q}} \tilde{\mathbf{q}} \dot{\mathbf{q}} + \partial_t \tilde{\mathbf{q}}$$

Ricordo che questa proprietà di invarianza è una conseguenza immediata del fatto che le equazioni di Eulero-Lagrange sono le equazioni che esprimono la stazionarietà dell'azione; più avanti studieremo nello stesso modo l'invarianza in forma delle equazioni di Hamilton.

L'invarianza in forma è il motivo del "successo" del formalismo lagrangiano: permette infatti di ottenere facilmente le equazioni del moto, scegliendo il sistema di coordinate più opportuno. È utile fare questa analisi anche nel caso hamiltoniano, dunque studieremo le trasformazioni che garantiscono l'invarianza in forma delle equazioni di Hamilton.

Definizione di trasformazione canonica

Una trasformazione

$$\begin{aligned} \tilde{\mathbf{q}} &= \tilde{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) \\ \tilde{\mathbf{p}} &= \tilde{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) \end{aligned}$$

è detta **canonica** se per ogni funzione $H(\mathbf{q}, \mathbf{p}, t)$ esiste una funzione $K(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}, t)$ tale che $(\mathbf{q}(t), \mathbf{p}(t))$ verifica le equazioni di Hamilton di hamiltoniana H se e solo se $(\tilde{\mathbf{q}}(t), \tilde{\mathbf{p}}(t))$ verifica le equazioni di Hamilton di hamiltoniana K .

Ogni trasformazione di coordinate conserva la natura lagrangiana di un moto, ma non tutte le trasformazioni di coordinate e impulsi conservano la natura hamiltoniana del moto. Vedremo però che la classe di trasformazioni che conservano la natura hamiltoniana del moto è più ampia delle sole trasformazioni di coordinate, e questo fatto è il primo vero vantaggio del formalismo hamiltoniano su quello lagrangiano. L'esempio più semplice che si può fare è

questo: data $H = H(q, p)$, si consideri la trasformazione che scambia, a meno di un segno, momento e coordinata:

$$\begin{aligned}\tilde{p} &= -q \\ \tilde{q} &= p\end{aligned}$$

è facile verificare che se $K(\tilde{q}, \tilde{p}) = H(-\tilde{p}, \tilde{q})$, e quindi $H(q, p) = K(p, -q)$, allora

$$\begin{aligned}\dot{q} &= \partial_p H(q, p) & \text{se e solo se} & & \dot{\tilde{q}} &= \dot{p} = -\partial_q H(q, p) = \partial_{\tilde{p}} K(\tilde{q}, \tilde{p}) \\ \dot{p} &= -\partial_q H(q, p) & & & \dot{\tilde{p}} &= -\dot{q} = -\partial_p H(q, p) = -\partial_{\tilde{q}} K(\tilde{q}, \tilde{p})\end{aligned}$$

Dunque abbiamo operato una trasformazione che scambia il ruolo di coordinate e impulsi, cosa evidentemente impossibile da farsi nel formalismo lagrangiano, dove le trasformazioni delle velocità $\dot{\mathbf{q}}$ sono determinate dalle trasformazioni delle coordinate.

4.2 $\mathbf{p} \cdot d\mathbf{q} - H dt$

Il formalismo hamiltoniano sprigiona tutta la sua potenza attraverso la definizione di un metodo generale per la ricerca di soluzioni. Per introdurlo, devo ulteriormente approfondire la struttura delle trasformazioni canoniche, attraverso lo studio del principio variazionale per le equazioni di Hamilton.

Nella sua seconda formulazione (quella puramente interpretativa) possiamo pensare all'azione S come all'integrale su una qualunque curva γ nello **spazio delle fasi esteso** $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ (cioè lo spazio prodotto dello spazio delle fasi e dell'asse temporale), che unisce $(\mathbf{q}_0, \mathbf{p}_0, 0)$ e $(\mathbf{q}_1, \mathbf{p}_1, T)$ della forma differenziale

$$\mathbf{p} \cdot d\mathbf{q} - H dt$$

Infatti, se $[0, \bar{\lambda}] \ni \lambda \rightarrow (\mathbf{q}(\lambda), \mathbf{p}(\lambda), t(\lambda))$ è un cammino siffatto

$$\int_{\gamma} (\mathbf{p} \cdot d\mathbf{q} - H dt) = \int_0^{\bar{\lambda}} (\mathbf{p}(\lambda) \cdot \partial_{\lambda} \mathbf{q}(\lambda) - H(\mathbf{q}(\lambda), \mathbf{p}(\lambda), t(\lambda)) \partial_{\lambda} t) d\lambda$$

Ma se la relazione tra λ e t è invertibile, riparametrizzando in t si ottiene di nuovo

$$S = \int_0^T (\mathbf{p} \cdot \dot{\mathbf{q}} - H) dt$$

Dunque il principio variazionale può essere riformulato dicendo che il moto $t \rightarrow (\mathbf{q}(t), \mathbf{p}(t))$ è hamiltoniano se e solo se l'azione $\int_{\gamma} (\mathbf{p} \cdot d\mathbf{q} - H dt)$ definita sulle curve γ in $\mathbb{R}^d \times \mathbb{R}^d \times [0, T]$ è stazionaria sulla curva $t \rightarrow (\mathbf{q}(t), \mathbf{p}(t), t)$.

Il principio variazionale pensato in questo modo permette di dare un'elegante condizione di canonicità di una trasformazione. Infatti, se, mediante un cambiamento di variabili, il sistema hamiltoniano di hamiltoniana H si trasforma nel sistema hamiltoniano di hamiltoniana K , allora i due seguenti funzionali

$$S = \int_0^T (\mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{q}, \mathbf{p}, t)) dt, \quad \tilde{S} = \int_0^T (\mathbf{P} \cdot \dot{\mathbf{Q}} - K(\mathbf{Q}, \mathbf{P}, t)) dt$$

devono avere gli stessi punti stazionari, una volta fissati i valori delle variabili agli estremi temporali. Una condizione sufficiente perché ciò accada è che le due forme differenziali

$$\mathbf{p} \cdot d\mathbf{q} - H dt, \quad \mathbf{P} \cdot d\mathbf{Q} - K dt,$$

dove \mathbf{P} e \mathbf{Q} sono scritte in termini di (\mathbf{q}, \mathbf{p}) , differiscano per un differenziale esatto dG . In tal caso, infatti, il valore dell'azione su una traiettoria nelle nuove variabili è

$$\tilde{S} = \int_0^T (\mathbf{P} \cdot \dot{\mathbf{Q}} - K(\mathbf{Q}, \mathbf{P}, t)) dt = \int_{\tilde{\gamma}} (\mathbf{P} \cdot d\mathbf{Q} - K dt)$$

dove $\tilde{\gamma}$ è la corrispondente curva nello spazio delle fasi esteso. Traducendo questo integrale nelle variabili (\mathbf{q}, \mathbf{p}) si ottiene

$$\int_{\tilde{\gamma}} (\mathbf{P} \cdot d\mathbf{Q} - K dt) = \int_{\gamma} (\mathbf{p} \cdot d\mathbf{q} - H dt) + \int_{\gamma} dG$$

dove γ è la curva $\tilde{\gamma}$ nelle variabili (\mathbf{q}, \mathbf{p}) . Il primo integrale è proprio l'azione S calcolata sul cammino nelle variabili (\mathbf{q}, \mathbf{p}) , mentre

$$\int_{\gamma} dG = G(\mathbf{q}_1, \mathbf{p}_1, T) - G(\mathbf{q}_0, \mathbf{p}_0, 0)$$

Ma allora S e \tilde{S} differiscono per una costante che dipende solo dai valori (fissati) agli estremi, dunque una traiettoria rende stazionaria S se e solo se rende stazionaria \tilde{S} nelle variabili (\mathbf{Q}, \mathbf{P}) . In tal modo, abbiamo dimostrato il seguente teorema.

Teorema 4.1. *Se la trasformazione $(\mathbf{q}, \mathbf{p}) \rightarrow (\mathbf{Q}, \mathbf{P})$ è tale che data H esiste K e una funzione G tali che*

$$\mathbf{p} \cdot d\mathbf{q} - H dt = \mathbf{P} \cdot d\mathbf{Q} - K dt + dG$$

allora la trasformazione manda il sistema hamiltoniano di hamiltoniana H nelle variabili (\mathbf{q}, \mathbf{p}) nel sistema hamiltoniano di hamiltoniana K nelle variabili (\mathbf{Q}, \mathbf{P}) .

Teorema 4.2. Canonicità attraverso i differenziali

Se la trasformazione $(\mathbf{q}, \mathbf{p}) \rightarrow (\mathbf{Q}, \mathbf{P})$ è tale che per ogni t fissato

$$\mathbf{p} \cdot d\mathbf{q} = \mathbf{P} \cdot d\mathbf{Q} + dG$$

allora la trasformazione è canonica.

Assegnata H , chiediamoci se esiste K tale che

$$\mathbf{p} \cdot d\mathbf{q} - H dt = \mathbf{P} \cdot d\mathbf{Q} - K dt + dG.$$

In questa espressione è tutto noto, tranne G . Isolo i coefficienti di dt

$$-H = \mathbf{P} \cdot \partial_t \mathbf{Q} - K + \partial_t G$$

Questa relazione definisce K . Usando il teorema precedente si ottiene la tesi.

Si noti che se la trasformazione non dipende dal tempo, la nuova hamiltoniana si ottiene dalla vecchia scrivendola nelle nuove variabili.

Si può dare una caratterizzazione più algebrica di queste trasformazioni.

4.3 Trasformazioni simplettiche

Indicherò con \mathbf{z} il complesso delle variabili in \mathbb{R}^{2n} :

$$\mathbf{z} = \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}, \quad \partial_{\mathbf{z}} = \begin{pmatrix} \partial_{\mathbf{q}} \\ \partial_{\mathbf{p}} \end{pmatrix}$$

Dunque

$$\dot{\mathbf{z}} = \begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \partial_{\mathbf{p}} H \\ -\partial_{\mathbf{q}} H \end{pmatrix} = J \partial_{\mathbf{z}} H$$

dove J è la **matrice simplettica fondamentale**

$$J = \begin{pmatrix} 0 & \mathbf{I}_n \\ -\mathbf{I}_n & 0 \end{pmatrix}$$

e \mathbf{I}_n è la matrice identità in \mathbb{R}^n . Consideriamo ora una trasformazione di coordinate indipendente dal tempo

$$\tilde{\mathbf{z}} = \tilde{\mathbf{z}}(\mathbf{z}),$$

e sia $\tilde{H} = H(\mathbf{z}(\tilde{\mathbf{z}}))$, così che $\partial_{\mathbf{z}} H = \left(\frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} \right)^t \partial_{\tilde{\mathbf{z}}} \tilde{H}$. Il sistema nelle nuove variabili è

$$\dot{\tilde{\mathbf{z}}} = \frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} \dot{\mathbf{z}} = \frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} J \partial_{\mathbf{z}} H = \frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} J \left(\frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} \right)^t \partial_{\tilde{\mathbf{z}}} \tilde{H}$$

che coincide con

$$\dot{\tilde{\mathbf{z}}} = J \partial_{\tilde{\mathbf{z}}} \tilde{H}$$

se e solo se

$$\frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} J \left(\frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} \right)^t = J$$

Se questa condizione è verificata per ogni \mathbf{z} , la trasformazione è canonica.

È utile dunque dare una definizione: una matrice A si dice **simplettica** se e solo se

$$A J A^t = J. \tag{4.1}$$

Inoltre, chiameremo simplettica qualunque trasformazione che ha jacobiano simplettico. Abbiamo dimostrato il seguente fatto

Teorema 4.3. Trasformazioni simplettiche e trasformazioni canoniche - I

Se

$$\tilde{\mathbf{z}} = \tilde{\mathbf{z}}(\mathbf{z}),$$

è simplettica, allora è canonica, e la nuova hamiltoniana \tilde{H} è semplicemente H nelle nuove variabili.

Per considerare il caso di trasformazioni dipendenti dal tempo, è necessario studiare un po' le matrici simplettiche.

Proposizione 4.1. Matrici simplettiche

- $J^2 = -\mathbf{I}_{2n}$, quindi $J^{-1} = -J$.

- Calcolando il determinante, si ottiene $\det J^2 = 1$, dunque $|\det J| = 1$ (in realtà è 1, come si può calcolare direttamente).
- Se A è simplettica, allora, passando ai determinanti, si ha che $\det A^2 = 1$, dunque A è invertibile
- A è simplettica se e solo se A^t è simplettica. Infatti, moltiplicando a destra per JA la (4.1) si ha

$$AJA^tJA = J^2A = -A$$

moltiplicando a sinistra per A^{-1} si ha

$$JA^tJA = -I$$

moltiplicando a sinistra per $-J$ si ottiene

$$A^tJA = J$$

che dimostra la tesi.

- A è simplettica se e solo se A^{-1} è simplettica. Infatti, passando agli inversi nell'ultima equazione del punto precedente, si ha

$$A^{-1}(-J)A^{-1t} = -J$$

che dà la tesi.

- Se A e B sono simplettiche, allora AB è simplettica (esercizio). Dunque le matrici simplettiche formano un sottogruppo del gruppo delle trasformazioni non singolari.

Diremo che una trasformazione è simplettica se il suo jacobiano è una matrice simplettica in ogni punto.

Teorema 4.4. $\mathbf{p} \cdot d\mathbf{q}$ e trasformazioni simplettiche

La trasformazione $\mathbf{Q} = \mathbf{Q}(\mathbf{q}, \mathbf{p}, t)$, $\mathbf{P} = \mathbf{P}(\mathbf{q}, \mathbf{p}, t)$ è una trasformazione simplettica a t fissato se e solo se $\mathbf{p} \cdot d\mathbf{q} - \mathbf{P} \cdot d\mathbf{Q}$ è un differenziale esatto a t fissato.

Osservo che

$$\mathbf{p} \cdot d\mathbf{q} = d(\mathbf{p} \cdot \mathbf{q}) - \mathbf{q} \cdot d\mathbf{p} = \frac{1}{2}(\mathbf{p} \cdot d\mathbf{q} - \mathbf{q} \cdot d\mathbf{p}) + \frac{1}{2}d(\mathbf{p} \cdot \mathbf{q})$$

e posso ottenere una analoga espressione per $\mathbf{P} \cdot d\mathbf{Q}$. Ne segue che $\mathbf{p} \cdot d\mathbf{q} - \mathbf{P} \cdot d\mathbf{Q}$ è un differenziale esatto se e solo se

$$\mathbf{p} \cdot d\mathbf{q} - \mathbf{q} \cdot d\mathbf{p} - (\mathbf{P} \cdot d\mathbf{Q} - \mathbf{Q} \cdot d\mathbf{P})$$

è un differenziale esatto. In termini di $\mathbf{z} = (\mathbf{q}, \mathbf{p})$ e $\mathbf{Z} = (\mathbf{Q}, \mathbf{P})$ questa espressione è

$$J\mathbf{z} \cdot d\mathbf{z} - J\mathbf{Z} \cdot d\mathbf{Z}$$

Ma

$$d\mathbf{Z} = \partial_{\mathbf{z}}\mathbf{Z} d\mathbf{z}$$

dunque la forma differenziale è

$$(J\mathbf{z} - \partial_{\mathbf{z}}\mathbf{Z}^t J\mathbf{Z}) \cdot d\mathbf{z}$$

La forma è localmente esatta se e solo se è localmente chiusa, cioè

$$\partial_i(J\mathbf{z} - \partial_{\mathbf{z}}\mathbf{Z}^t J\mathbf{Z})_j = \partial_j(J\mathbf{z} - \partial_{\mathbf{z}}\mathbf{Z}^t J\mathbf{Z})_i$$

Ora

$$\begin{aligned}\partial_i(J\mathbf{z})_j &= \sum_k \partial_i(J_{jk}z_k) \sum_k J_{jk}\delta_{ik} = J_{ji} \\ \partial_i(\partial_{\mathbf{z}}\mathbf{Z}^t J\mathbf{Z})_j &= \sum_{h,k} \partial_i(\partial_j Z_h J_{hk} Z_k) = \sum_{h,k} \partial_{ij}^2 Z_h J_{hk} Z_k + \sum_{h,k} \partial_j Z_h J_{hk} \partial_i Z_k \\ &= \sum_{h,k} \partial_{ij}^2 Z_h J_{hk} Z_k + (\partial_{\mathbf{z}}\mathbf{Z}^t J \partial_{\mathbf{z}}\mathbf{Z})_{ji}\end{aligned}$$

Sviluppando nello stesso modo il membro di destra, si ottiene la condizione di chiusura

$$J_{ji} - (\partial_{\mathbf{z}}\mathbf{Z}^t J \partial_{\mathbf{z}}\mathbf{Z})_{ji} = J_{ij} - (\partial_{\mathbf{z}}\mathbf{Z}^t J \partial_{\mathbf{z}}\mathbf{Z})_{ij}$$

(i termini con le derivate seconde sono uguali e si cancellano). Ma sia la matrice J che la matrice $\partial_{\mathbf{z}}\mathbf{Z}^t J \partial_{\mathbf{z}}\mathbf{Z}$ sono antisimmetriche, e per una matrice antisimmetrica $A_{ij} = -A_{ji}$ se e solo se $A_{ij} = 0$. Dunque la condizione di chiusura è proprio la condizione di simpletticità dello jacobiano della trasformazione

$$J = \partial_{\mathbf{z}}\mathbf{Z}^t J \partial_{\mathbf{z}}\mathbf{Z}.$$

Una conseguenza di questo teorema è che le trasformazioni simplettiche dipendenti dal tempo sono canoniche.

Teorema 4.5. Trasformazioni simplettiche e trasformazioni canoniche - II

Se la trasformazione $\mathbf{Q} = \mathbf{Q}(\mathbf{q}, \mathbf{p}, t)$, $\mathbf{P} = \mathbf{P}(\mathbf{q}, \mathbf{p}, t)$, anche dipendente dal tempo, è simplettica per ogni t fissato, allora esiste $G(\mathbf{q}, \mathbf{p}, t)$ tale che,

$$\mathbf{p} \cdot d\mathbf{q} - \mathbf{P} \cdot d\mathbf{Q} = dG, \quad \text{a } t \text{ fissato, per ogni } t$$

In tal caso, se H è un'hamiltoniana, scegliendo

$$K = H + \mathbf{P} \cdot \partial_t \mathbf{Q} + \partial_t G \tag{4.2}$$

si ottiene che

$$\mathbf{p} \cdot d\mathbf{q} - H dt = \mathbf{P} \cdot d\mathbf{Q} - K dt + dG$$

Quindi una trasformazione simplettica dipendente dal tempo è una trasformazione canonica e la nuova hamiltoniana si calcola come in (4.2). In particolare, se la trasformazione simplettica non dipende dal tempo la nuova hamiltoniana è la vecchia hamiltoniana espressa in funzione delle nuove variabili.

5 Parentesi di Poisson

Le parentesi di Poisson sono uno dei degli strumenti chiave del formalismo hamiltoniano. La forma bilineare antisimmetrica in \mathbb{R}^{2n}

$$[\mathbf{v}, \mathbf{w}]_J = \mathbf{v} \cdot J\mathbf{w}$$

è detta **prodotto simplettico**.

È facile verificare che una matrice A è simplettica se e solo se, per ogni \mathbf{v}, \mathbf{w} vale

$$[A\mathbf{v}, A\mathbf{w}]_J = [\mathbf{v}, \mathbf{w}]_J.$$

Infatti questa condizione equivale a

$$\mathbf{v} \cdot A^t J A \mathbf{w} = \mathbf{v} \cdot J \mathbf{w}, \quad \forall \mathbf{v}, \mathbf{w}$$

e questo può accadere se e solo se $A^t J A = J$.

In prodotto simplettico è legato alle “parentesi di Poisson”, che sono una operazione sugli “osservabili”, cioè sulle funzioni definite nello spazio delle fasi $\mathbf{z} = (\mathbf{q}, \mathbf{p})$.

Le **parentesi di Poisson** sono un’operazione che associa a due funzioni f, g , la funzione

$$\{f, g\} = [\partial_{\mathbf{z}} f, \partial_{\mathbf{z}} g]_J = \partial_{\mathbf{z}} f \cdot J \partial_{\mathbf{z}} g = \partial_{\mathbf{q}} f \cdot \partial_{\mathbf{p}} g - \partial_{\mathbf{p}} f \cdot \partial_{\mathbf{q}} g$$

Consideriamo ora un cambiamento di variabili, e, con un punto di vista “geometrico”, indichiamo con f sia $f(\mathbf{q}, \mathbf{p})$, sia $f(\mathbf{q}(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}), \mathbf{p}(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}))$, cioè f come funzione delle nuove variabili tramite le vecchie. Se si trasformano le variabili, cioè si pensano f e g funzioni delle nuove variabili tramite le vecchie

$$\partial_{\mathbf{z}} f = \partial_{\tilde{\mathbf{z}}} \tilde{\mathbf{z}}^t \partial_{\tilde{\mathbf{z}}} f$$

Dunque

$$\{f, g\}_{\mathbf{q}, \mathbf{p}} = [\partial_{\tilde{\mathbf{z}}} \tilde{\mathbf{z}}^t \partial_{\tilde{\mathbf{z}}} f, \partial_{\tilde{\mathbf{z}}} \tilde{\mathbf{z}}^t \partial_{\tilde{\mathbf{z}}} g]_J = \partial_{\tilde{\mathbf{z}}} f \cdot \partial_{\tilde{\mathbf{z}}} \tilde{\mathbf{z}} J \partial_{\tilde{\mathbf{z}}} \tilde{\mathbf{z}}^t \partial_{\tilde{\mathbf{z}}} g$$

quindi

$$\forall f, g, \quad \{f, g\}_{\mathbf{q}, \mathbf{p}} = \{f, g\}_{\tilde{\mathbf{q}}, \tilde{\mathbf{p}}}$$

se e solo se la trasformazione è simplettica.

Le parentesi di Poisson delle coppie di variabili danno le **regole di commutazione canoniche**:

$$\{q_i, q_j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{q_i, p_j\} = \delta_{ij} \quad (5.1)$$

Se considero una trasformazione è canonica, queste relazioni devono valere anche per le nuove variabili rispetto alle nuove variabili, ma, per l’invarianza appena dimostrata, devono valere anche per le nuove variabili rispetto alle vecchie variabili:

$$\{\tilde{q}_i, \tilde{q}_j\} = 0, \quad \{\tilde{p}_i, \tilde{p}_j\} = 0, \quad \{\tilde{q}_i, \tilde{p}_j\} = \delta_{ij} \quad (5.2)$$

Queste condizioni sono del tutto equivalenti alla simpletticità della trasformazione. Dimostriamolo. La condizione di simpletticità è:

$$\begin{aligned} J &= \frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} J \left(\frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} \right)^t = \frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} J \begin{pmatrix} \partial_{\mathbf{q}} \tilde{\mathbf{q}}^t & \partial_{\mathbf{q}} \tilde{\mathbf{p}}^t \\ \partial_{\mathbf{p}} \tilde{\mathbf{q}}^t & \partial_{\mathbf{p}} \tilde{\mathbf{p}}^t \end{pmatrix} = \frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{z}} \begin{pmatrix} \partial_{\mathbf{p}} \tilde{\mathbf{q}}^t & \partial_{\mathbf{p}} \tilde{\mathbf{p}}^t \\ -\partial_{\mathbf{q}} \tilde{\mathbf{q}}^t & -\partial_{\mathbf{q}} \tilde{\mathbf{p}}^t \end{pmatrix} \\ &= \begin{pmatrix} \partial_{\mathbf{q}} \tilde{\mathbf{q}} \partial_{\mathbf{p}} \tilde{\mathbf{q}}^t - \partial_{\mathbf{p}} \tilde{\mathbf{q}} \partial_{\mathbf{q}} \tilde{\mathbf{q}}^t & \partial_{\mathbf{q}} \tilde{\mathbf{q}} \partial_{\mathbf{p}} \tilde{\mathbf{p}}^t - \partial_{\mathbf{p}} \tilde{\mathbf{q}} \partial_{\mathbf{q}} \tilde{\mathbf{p}}^t \\ \partial_{\mathbf{q}} \tilde{\mathbf{p}} \partial_{\mathbf{p}} \tilde{\mathbf{q}}^t - \partial_{\mathbf{p}} \tilde{\mathbf{p}} \partial_{\mathbf{q}} \tilde{\mathbf{q}}^t & \partial_{\mathbf{q}} \tilde{\mathbf{p}} \partial_{\mathbf{p}} \tilde{\mathbf{p}}^t - \partial_{\mathbf{p}} \tilde{\mathbf{p}} \partial_{\mathbf{q}} \tilde{\mathbf{p}}^t \end{pmatrix} \end{aligned}$$

(si ricordi che si tratta di prodotti a blocchi di matrici). Si noti ora che se $\mathbf{f}(\mathbf{q}, \mathbf{p})$ e $\mathbf{g}(\mathbf{q}, \mathbf{p})$ sono due campi vettoriali a valori in \mathbb{R}^n , allora

$$(\partial_{\mathbf{q}} \mathbf{f} (\partial_{\mathbf{p}} \mathbf{g})^t)_{ij} = \partial_{\mathbf{q}} f_i \cdot \partial_{\mathbf{p}} g_j.$$

Con questa osservazione è semplice verificare che l'identità tra J e l'ultima matrice è equivalente alle condizioni (5.2).

Consideriamo un esempio. Sia

$$P = \frac{1}{2}(q^2 + p^2)$$

$$Q = \arctan \frac{q}{p}$$

In questo caso è molto semplice verificare la canonicità della trasformazione mediante le parentesi di Poisson. Infatti, per definizione, $\{Q, Q\} = 0 = \{P, P\}$, dunque resta solo da verificare che

$$\{Q, P\} = 1$$

Il semplice calcolo delle derivate mostra che effettivamente questa condizione è verificata (completare per esercizio).

Consideriamo ora l'hamiltoniana dell'oscillatore armonico $H = (p^2 + q^2)/2$. L'hamiltoniana nelle nuove variabili è

$$K = P$$

per cui le equazioni del moto diventano

$$\dot{Q} = \partial_P K = \partial_P P = 1$$

$$\dot{P} = -\partial_Q K = -\partial_Q P = 0$$

che sono di facile soluzione: P è costante e pari all'energia del moto, mentre $Q(t) = Q_0 + t$. Ne segue che il moto è risolto dalle uguaglianze

$$\frac{1}{2}(p^2(t) + q^2(t)) = E$$

$$\arctan \frac{q(t)}{p(t)} = Q_0 + t$$

Dove E e Q_0 si determinano a partire dal dato iniziale.

In questo esempio si **porta alle quadrature** (cioè si risolve il moto in termini di integrali di funzioni elementari) il moto di un oscillatore armonico (naturalmente questo moto si risolve anche utilizzando la teoria delle equazioni differenziali lineari). Esiste un metodo generale per provare a portare alle quadrature un sistema hamiltoniano mediante una trasformazione canonica che renda semplice il sistema nelle nuove variabili. Per poterlo illustrare serve però introdurre un metodo che permette di ottenere abbastanza facilmente delle trasformazioni canoniche, come mostreremo tra qualche pagina.

Come operazione tra funzioni, le parentesi di Poisson verificano le seguenti proprietà.

Teorema 5.1. Proprietà delle parentesi di Poisson.

1. Sono bilineari (verificare per esercizio).
2. Sono antisimmetriche:

$$\{f, g\} = -\{g, f\}$$

e quindi $\{f, f\} = 0$ (verificare per esercizio).

3. Vale la formula di Leibniz

$$\{fg, h\} = f\{g, h\} + g\{f, h\}$$

(verificare per esercizio).

4. Vale l'identità di Jacobi

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$$

Uno spazio vettoriale (reale o complesso), dotato di un prodotto interno che verifica le proprietà **1,2,4** è detto **algebra di Lie**, dunque lo spazio delle funzioni regolari in \mathbb{R}^{2n} con le parentesi di Poisson è un'algebra di Lie. Da questa definizione perché incontreremo altri casi di algebre di Lie. In particolare, lo spazio vettoriale di operatori lineari su uno spazio Hilbert H (per esempio \mathbb{R}^m o L^2) diventano algebre di Lie considerando come operazione interna il **commutatore** tra operatori: se \mathbf{v} è un elemento dello spazio H , e A e B sono due operatori lineari,

$$[A, B]\mathbf{v} = AB\mathbf{v} - BA\mathbf{v}.$$

La bilinearità e l'antisimmetria del commutatore sono di verifica immediata, l'identità di Jacobi si può mostrare facilmente sviluppando tutti i termini, ma si può abbreviare con un minimo di riflessione: tutti i termini dell'espressione

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]]$$

sono formati da una permutazione del prodotto tra le tre matrici A, B, C . Isoliamo i termini che iniziano per A . Indicando con \dots i termini che non iniziano per H , si ha

$$\begin{aligned} [A, [B, C]] &= A[B, C] = ABC - ACB \\ [B, [C, A]] &= -[C, A]B + \dots = ACB + \dots \\ [C, [A, B]] &= -[A, B]C + \dots = -ABC + \dots \end{aligned}$$

Dunque la somma di tutti i termini che iniziano per A è nulla. Si può ripetere lo stesso ragionamento per i termini che iniziano per B e C (l'espressione dell'identità di Jacobi è invariante per le permutazioni degli argomenti), dunque la somma dei tre termini è effettivamente nulla.

Ho premesso la prova dell'identità di Jacobi per il commutatore perché fa da traccia per la prova dell'identità di Jacobi per le parentesi di Poisson. Serve però qualche utile passaggio intermedio. Dato il campo vettoriale \mathbf{v} , indico la derivata di una funzione f lungo \mathbf{v} con il simbolo

$$L_{\mathbf{v}}f = \mathbf{v} \cdot \nabla f$$

Abbiamo così definito un operatore (differenziale) associato al campo vettoriale \mathbf{v} . Si noti che ogni operatore differenziale del primo ordine ha questa forma. Inoltre è evidente che il campo \mathbf{v} è nullo se e solo se il corrispondente operatore $L_{\mathbf{v}}$ è nullo.

Introduco ora un'operazione tra campi vettoriali che discende dal commutatore dei rispettivi operatori associati. Siano \mathbf{v} e \mathbf{w} due campi vettoriali, Considero il commutatore tra gli operatori $L_{\mathbf{v}}$ e $L_{\mathbf{w}}$, su una funzione f :

$$[L_{\mathbf{v}}, L_{\mathbf{w}}]f = \mathbf{v} \cdot \nabla(\mathbf{w} \cdot \nabla f) - \mathbf{w} \cdot \nabla(\mathbf{v} \cdot \nabla f)$$

L'espressione a destra sembra contenere derivate prime e seconde di f , ma a una più attenta analisi si scopre che le derivate seconde non ci sono, e dunque il commutatore dei due operatori di derivazione è anch'esso un operatore del primo ordine. Infatti, i termini nelle derivate seconde sono:

$$\sum_{i,j} v_i w_j \partial_{ij}^2 f - \sum_{i,j} w_i v_j \partial_{ji}^2 f$$

che dunque si cancellano. Il campo vettoriale \mathbf{u} tale che

$$[L_{\mathbf{v}}, L_{\mathbf{w}}] = L_{\mathbf{u}}$$

è il **commutatore** $[\mathbf{v}, \mathbf{w}]$ dei due campi \mathbf{v} e \mathbf{w} (detto anche **prodotto di Lie** dei due campi). Usando la definizione si ottiene

$$[\mathbf{v}, \mathbf{w}] = (\mathbf{v} \cdot \nabla) \mathbf{w} - (\mathbf{w} \cdot \nabla) \mathbf{v}$$

L'identità di Jacobi per operatori $L_{\mathbf{u}}, L_{\mathbf{v}}, L_{\mathbf{w}}$, si riscrive facilmente come

$$L_{[\mathbf{u},[\mathbf{v},\mathbf{w}]]} + L_{[\mathbf{v},[\mathbf{w},\mathbf{u}]]} + L_{[\mathbf{w},[\mathbf{u},\mathbf{v}]]} = 0$$

Poiché $L : \mathbf{u} \rightarrow L_{\mathbf{u}}$ è lineare in \mathbf{u} , l'identità precedente si scrive come

$$L_{[\mathbf{u},[\mathbf{v},\mathbf{w}]]+[\mathbf{v},[\mathbf{w},\mathbf{u}]]+[\mathbf{w},[\mathbf{u},\mathbf{v}]]} = 0$$

che è possibile se e solo se anche il commutatore dei campi vettoriali verifica l'identità di Jacobi.

Torniamo alle parentesi di Poisson. Usando la definizione, si vede che

$$\{f, g\} = \partial_{\mathbf{z}} f \cdot J \partial_{\mathbf{z}} g = L_{J \partial_{\mathbf{z}} g} f = -L_{J \partial_{\mathbf{z}} f} g$$

Riscrivo i tre termini nel membro di sinistra dell'identità di Jacobi per le parentesi di Poisson come operatori che agiscono su h . Il primo si riscrive come

$$-L_{J \partial_{\mathbf{z}} f} (-L_{J \partial_{\mathbf{z}} g} h) = L_{J \partial_{\mathbf{z}} f} L_{J \partial_{\mathbf{z}} g} h$$

Il secondo è

$$-L_{J \partial_{\mathbf{z}} g} L_{J \partial_{\mathbf{z}} f} h$$

Il terzo è

$$L_{J \partial_{\mathbf{z}} \{f, g\}} h$$

La somma dei tre termini è dunque

$$([L_{J \partial_{\mathbf{z}} f}, L_{J \partial_{\mathbf{z}} g}] + L_{J \partial_{\mathbf{z}} \{f, g\}}) h$$

che è un'espressione lineare nelle derivate prime di h ma non contiene derivate seconde di h . Ripetendo il ragionamento, si prova che non c'è dipendenza dalle derivate seconde di nessuna delle tre funzioni. D'altra parte ogni termine dello sviluppo dell'identità di Jacobi è lineare nelle derivate seconde di una delle tre funzioni. Ne segue che tutti i termini sono nulli, e dunque vale l'identità di Jacobi.

Come corollario, segue che l'identità di Jacobi è equivalente a

$$[L_{J \partial_{\mathbf{z}} f}, L_{J \partial_{\mathbf{z}} g}] = -L_{J \partial_{\mathbf{z}} \{f, g\}}$$

cioè, in termini di prodotti di Lie,

$$[J \partial_{\mathbf{z}} f, J \partial_{\mathbf{z}} g] = -J \partial_{\mathbf{z}} \{f, g\}.$$

D'ora in poi chiamerò **campo vettoriale hamiltoniano** associato alla funzione f , il campo $J \partial_{\mathbf{z}} f$. L'identità precedente afferma che il campo vettoriale hamiltoniano associato alle parentesi di Poisson delle due funzioni f e g è meno il commutatore dei due campi hamiltoniani associati a f e g .

La commutatività dei campi vettoriali è equivalente alla commutatività dei flussi generati. Lavorerò in coordinate rettangolari, ma usando carte locali questa trattazione si estende al caso di flussi e campi su varietà.

Siano dati due campi vettoriali regolari $\mathbf{v}(\mathbf{x})$ e $\mathbf{w}(\mathbf{x})$. Definisco i due flussi associati:

$$\begin{cases} \frac{d}{dt} \Phi^t(\mathbf{x}) = \mathbf{v}(\Phi^t(\mathbf{x})) \\ \Phi^0(\mathbf{x}) = \mathbf{x} \end{cases} \quad \begin{cases} \frac{d}{dt} \Psi^t(\mathbf{x}) = \mathbf{w}(\Psi^t(\mathbf{x})) \\ \Psi^0(\mathbf{x}) = \mathbf{x} \end{cases}$$

Teorema 5.2. Flussi commutanti.

Φ^t e Ψ^s commutano, cioè

$$\Phi^t(\Psi^s(\mathbf{x})) = \Psi^s(\Phi^t(\mathbf{x}))$$

per ogni s, t, \mathbf{x} , se e solo se i corrispondenti campi commutano.

Dimostro questo risultato in \mathbb{R}^m , ma la tesi rimane valida anche per flussi su varietà.

Osservo preliminarmente che, sviluppando in $t = 0$:

$$\Phi^t(\mathbf{x}) = \mathbf{x} + t\mathbf{v}(\mathbf{x}) + \frac{t^2}{2} \mathbf{v} \cdot \nabla \mathbf{v} + O(t^3)$$

Analogamente,

$$\Psi^s(\mathbf{x}) = \mathbf{x} + s\mathbf{w}(\mathbf{x}) + \frac{s^2}{2} \mathbf{w} \cdot \nabla \mathbf{w} + O(s^3)$$

Lemma 5.1. Vale

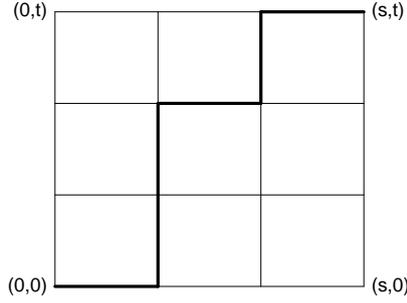
$$\Phi^t(\Psi^s(\mathbf{x})) - \Psi^s(\Phi^t(\mathbf{x})) = st[\mathbf{w}, \mathbf{v}] + O_3 \tag{5.3}$$

dove con O_3 intendo termini di ordine superiore al secondo.

Prova completata a lezione (si fa sostituendo un'espressione nell'altra). Ci sono due punti importanti in questo enunciato: non ci sono termini del primo ordine, e quelli del secondo sono rappresentati dal solo termine misto st di coefficiente pari al commutatore dei campi, mentre sono assenti i termini del secondo ordine in t^2 e s^2 .

Da questo lemma segue facilmente che se Φ^t e Ψ^s commutano, allora $[\mathbf{v}, \mathbf{w}]$ è nullo, infatti lo sviluppo in serie di potenze di s e t del membro di destra della (5.3) deve essere identicamente nullo, e quindi deve essere nullo il coefficiente del termine in st .

Vale anche il viceversa, come ora dimostreremo. La dimostrazione è concettualmente semplice, ma richiede un po' di notazioni e di grafici. Fissiamo $s, t > 0$, m intero, $\delta s = s/m$, $\delta t = t/m$.



Considera il rettangolo $[0, s] \times [0, t]$ nel piano (s, t) , diviso in m^2 rettangolini di lati $\delta s, \delta t$ (in figura $m = 3$). Consideriamo un cammino γ da $(0, 0)$ a (s, t) fatto di segmenti dei rettangolini, ma con s e t non decrescenti (il cammino può solo andare a destra o in alto). Fissato γ , indichiamo con $T_\gamma(\mathbf{x})$ il punto di \mathbb{R}^n che otteniamo da \mathbf{x} evolvendo con la composizione di $\Phi^{\delta t}$ e $\Psi^{\delta s}$ nell'ordine con cui compaiono in γ i tratti orizzontali e i tratti verticali, rispettivamente. Per esempio, al cammino γ in figura corrisponde il punto

$$T_\gamma(\mathbf{x}) = \Psi^{\delta s} \circ \Phi^{\delta t} \circ \Psi^{\delta s} \circ \Phi^{\delta t} \circ \Phi^{\delta t} \circ \Psi^{\delta s}(\mathbf{x})$$

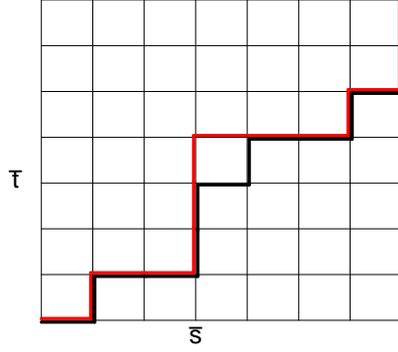
(nota che la sequenza dei tratti di γ si ritrova al contrario nell'espressione di T_γ). Dunque i cammini γ sono in biezione con le sequenze composte di m flussi $\Phi^{\delta t}$ e m flussi $\Psi^{\delta s}$. È facile convincersi che esiste una sequenza di cammini $\{\gamma_k\}_{k=0\dots m^2}$ tale che

- $T_{\gamma_{m^2}} = (\Phi^{\delta t})^m \circ (\Psi^{\delta s})^m = \Phi^t \circ \Psi^s$;
- $T_{\gamma_0} = (\Psi^{\delta s})^m \circ (\Phi^{\delta t})^m = \Psi^s \circ \Phi^t$;
- γ_{k+1} differisce da γ_k per un solo rettangolino, cioè per lo sostituzione di un movimento “a destra, poi in alto” con un movimento “in alto, poi a destra”, e quindi la corrispondente sequenza di flussi differisce per uno scambio di $\Psi^{\delta s}\Phi^{\delta t}$ con $\Phi^{\delta t}\Psi^{\delta s}$;

•

$$\Phi^t(\Psi^s(\mathbf{x})) - \Psi^s(\Phi^t(\mathbf{x})) = \sum_{k=0}^{m^2-1} (T_{\gamma_{k+1}}(\mathbf{x}) - T_{\gamma_k}(\mathbf{x}))$$

Proveremo che ogni termine della sommatoria è di ordine $1/m^3$, dunque passando al limite $m \rightarrow +\infty$ si ottiene la tesi.



Considera i due cammini γ' e γ in figura, che differiscono per un solo rettangolino. Indico con T^- la sequenza di flussi da $(0,0)$ al punto (\bar{s}, \bar{t}) , e con T^+ la sequenza da $(\bar{s} + \delta s, \bar{t} + \delta t)$ a (s, t) . Dunque

$$T_{\gamma'}(\mathbf{x}) = T^+ \circ \Psi^{\delta s} \circ \Phi^{\delta t} \circ T^-(\mathbf{x}), \quad T_{\gamma}(\mathbf{x}) = T^+ \circ \Phi^{\delta t} \circ \Psi^{\delta s} \circ T^-(\mathbf{x}),$$

Indicando con $\mathbf{y} = T^-(\mathbf{x})$.

$$|T_{\gamma'}(\mathbf{x}) - T_{\gamma}(\mathbf{x})| = |T^+ \circ \Phi^{\delta t} \circ \Psi^{\delta s}(\mathbf{y}) - T^+ \circ \Psi^{\delta s} \circ \Phi^{\delta t}(\mathbf{y})| \leq c |\Phi^{\delta t} \circ \Psi^{\delta s}(\mathbf{y}) - \Psi^{\delta s} \circ \Phi^{\delta t}(\mathbf{y})|$$

dove c è la costante di Lipschitz per T^+ . Usando il lemma e la commutatività dei campi, si ottiene che ogni termine è di ordine $1/m^3$. Quindi, poiché i termini della somma che stiamo considerando sono solo m^2 , si ha

$$\Phi^t(\Psi^s(\mathbf{x})) - \Psi^s(\Phi^t(\mathbf{x})) = \sum_{k=0}^{m^2-1} (\varphi(T_{\gamma_{k+1}}(\mathbf{x})) - \varphi(T_{\gamma_k}(\mathbf{x}))) = O\left(\frac{1}{m}\right)$$

che tende a 0 per $n \rightarrow +\infty$.

Esercizio 3.

Nella dimostrazione ho usato che le costanti di Lipschitz dei flussi T_{γ} sono limitate uniformemente in γ . Dimostralo trovando una costante L tale che

$$|\partial_{\mathbf{x}} T_{\gamma}(\mathbf{x})| \leq L$$

indipendentemente da γ e supponendo \mathbf{x} in un compatto.

5.1 Integrali primi

Le parentesi di Poisson permettono di esprimere le equazioni di Hamilton in termini degli osservabili. Infatti, se $f(\mathbf{q}, \mathbf{p}, t)$ è una funzione regolare data,

$$\frac{df}{dt} = \partial_t f + \partial_{\mathbf{q}} f \cdot \dot{\mathbf{q}} + \partial_{\mathbf{p}} f \cdot \dot{\mathbf{p}} = \partial_t f + \partial_{\mathbf{q}} f \cdot \partial_{\mathbf{p}} H - \partial_{\mathbf{p}} f \cdot \partial_{\mathbf{q}} H = \partial_t f + \{f, H\}$$

In particolare si ottengono le equazioni di Hamilton:

$$\begin{aligned}\dot{q}_i &= \frac{dq_i}{dt} = \{q_i, H\} \\ \dot{p}_i &= \frac{dp_i}{dt} = \{p_i, H\}\end{aligned}$$

Inoltre è facile scrivere la definizione di integrale primo del moto in termini di parentesi di Poisson. La funzione f è costante lungo il moto se e solo se $\frac{df}{dt} = 0$, cioè se

$$\partial_t f + \{f, H\} = 0.$$

Si noti che questa è anche l'equazione di Liouville, poiché i campi hamiltoniani sono campi a divergenza nulla.

Nel caso di funzioni indipendenti dal tempo, vale

$$\frac{df}{dt} = \{H, f\}$$

che è un'equazione lineare che esprime l'evoluzione temporale degli osservabili. Infine, f indipendente dal tempo, è un integrale primo del moto se e solo $\{f, H\} = 0$. Usando l'identità di Jacobi, si dimostri per esercizio il seguente teorema

Teorema 5.3. Parentesi di Poisson di due integrali primi.

Se f e g sono due integrali primi del moto, allora anche $\{f, g\}$ lo è (si deve usare l'identità di Jacobi).

Per esempio, sia $\mathbf{q} \in \mathbb{R}^3$, e sia $\frac{m}{2}\dot{\mathbf{q}}^2$ l'energia cinetica. Allora $\mathbf{p} = m\dot{\mathbf{q}}$ e il momento della quantità di moto è $\ell = \mathbf{q} \wedge \mathbf{p}$. Mostrare che

$$\{\ell_1, \ell_2\} = \ell_3$$

Quindi se si conservano le prime due componenti del momento della quantità di moto si conserva anche la terza (in generale $\{\ell_i, \ell_j\} = \varepsilon_{ijk}\ell_k$ dove ε_{ijk} è il tensore completamente antisimmetrico).

Definizione: due funzioni H e K sono in **involuzione** se $\{H, K\} = 0$.

Ricordando che il commutatore dei campi hamiltoniani associati a due funzioni H e K è meno il campo hamiltoniano generato da $\{H, K\}$, si ottiene che due flussi hamiltoniani di hamiltoniane H e K commutano se e solo se

$$J \partial_{\mathbf{z}} \{H, K\} = 0$$

cioè se e sole se $\{H, K\} = \text{costante}$, e questo accade in particolare se H e K sono in involuzione. Osservo anche che essere in involuzione garantisce che K è un integrale primo del moto di hamiltoniana H , e viceversa (cosa che non accade se $\{H, K\}$ è una costante non nulla).

Riassumendo: se $\{H, K\} = 0$, i flussi hamiltoniani di hamiltoniane H e K commutano, e H e K sono integrali primi per entrambi i flussi.

5.2 Il teorema di Noether per sistemi hamiltoniani

Per un sistema lagrangiano, l'esistenza di un gruppo a un parametro di simmetrie nelle variabili è equivalente all'esistenza di una combinazione lineare dei momenti che si conserva. Per i sistemi hamiltoniani si possono considerare gruppi di simmetrie che coinvolgono le $2n$ variabili, ma è falso che a ogni gruppo di simmetria nelle $2n$ variabili corrisponde un integrale primo. La maggiore generalità delle simmetrie considerabili va limitata dalla condizione di simpletticità.

Teorema 5.4. Generatori dei flussi

Sia $\mathbf{x} \in \mathbb{R}^m$, e sia $\Psi_t(\mathbf{x})$ un gruppo a un parametro di diffeomorfismi, cioè

$$\Psi_0(\mathbf{x}) = \mathbf{x}, \quad e \quad \forall s, t \quad \Psi_t \circ \Psi_s = \Psi_{t+s}$$

Allora esiste un campo vettoriale \mathbf{u} tale che Ψ_t è il flusso generato da \mathbf{u} , cioè

$$\frac{d}{dt} \Psi_t(\mathbf{x}) = \mathbf{u}(\Psi_t(\mathbf{x}))$$

Se esiste \mathbf{u} , in particolare $\mathbf{u}(\mathbf{x})$ deve coincidere con la derivata in t al tempo $t = 0$ di $\Psi_t(\mathbf{x})$. Definisco dunque

$$\mathbf{u}(\mathbf{x}) = \frac{d}{dt} \Psi_t(\mathbf{x}).$$

Verifico che genera Ψ_t .

$$\frac{d}{dt} \Psi_t(\mathbf{x}) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\Psi_{t+\varepsilon}(\mathbf{x}) - \Psi_t(\mathbf{x})) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\Psi_\varepsilon(\Psi_t(\mathbf{x})) - \Psi_t(\mathbf{x})) = \mathbf{u}(\Psi_t(\mathbf{x}))$$

Nella penultima uguaglianza ho usato la proprietà di gruppo, nell'ultima ho usato la definizione di \mathbf{u} .

Teorema 5.5. Flussi simplettici

Sia $\mathbf{z} \in \mathbb{R}^{2n}$ e sia Ψ_t un gruppo a un parametro, generato dal campo vettoriale \mathbf{u} . Allora $\partial_z \Psi_t$ è simplettica se e solo se esiste (localmente) $K = K(\mathbf{z})$ tale che

$$\mathbf{u} = J \partial_z K$$

Chiamerò flussi simplettici i flussi che hanno jacobiano simplettico. Il teorema asserisce che i flussi simplettici sono tutti e soli i flussi hamiltoniani.

Dimostro il teorema. Derivando in \mathbf{z} la relazione

$$\frac{d}{dt} \Psi_t(\mathbf{z}) = \mathbf{u}(\Psi_t(\mathbf{z}))$$

si ottiene

$$\frac{d}{dt} \partial_z \Psi_t(\mathbf{z}) = \partial_z \mathbf{u}(\Psi_t(\mathbf{z})) \partial_z \Psi_t(\mathbf{z})$$

Suppongo ora che $\partial_z \Psi_t$ sia simplettica per ogni t , cioè

$$(\partial_z \Psi_t)^t J \partial_z \Psi_t = J$$

Derivo in t . Il secondo membro è nullo, mentre

$$\begin{aligned} \frac{d}{dt} ((\partial_z \Psi_t)^t J \partial_z \Psi_t) &= (\partial_z \Psi_t)^t (\partial_z \mathbf{u})^t J \partial_z \Psi_t + (\partial_z \Psi_t)^t J \partial_z \mathbf{u} \partial_z \Psi_t \\ &= (\partial_z \Psi_t)^t ((\partial_z \mathbf{u})^t J + J \partial_z \mathbf{u}) \partial_z \Psi_t \end{aligned}$$

dove $\partial_{\mathbf{z}}\mathbf{u}$ è calcolato in $\Psi^t(\mathbf{z})$. Poiché il flusso è invertibile, ottengo

$$(\partial_{\mathbf{z}}\mathbf{u})^t J + J \partial_{\mathbf{z}}\mathbf{u}$$

che riscrivo come

$$J \partial_{\mathbf{z}}\mathbf{u} = -(\partial_{\mathbf{z}}\mathbf{u})^t J$$

e calcolo al tempo $t = 0$, in modo da avere informazioni su $\partial_{\mathbf{z}}u$. Poiché $-J = J^t$, la matrice $J \partial_{\mathbf{z}}\mathbf{u}$ è simmetrica, cioè la forma differenziale associata è chiusa, dunque localmente esiste K tale che

$$J\mathbf{u} = -\partial_{\mathbf{z}}K$$

e da questa relazione, moltiplicando per $-J$ ottengo proprio che \mathbf{u} è il campo di hamiltoniana K .

Viceversa, se $\mathbf{u} = J \partial_{\mathbf{z}}K$, $(\partial_{\mathbf{z}}\mathbf{u})^t J + J \partial_{\mathbf{z}}\mathbf{u} = 0$ calcolato in $\Psi^t(z)$, per ogni t e z . Ne segue che la derivata nel tempo di $A_t = (\partial_{\mathbf{z}}\Psi_t)^t J \partial_{\mathbf{z}}\Psi_t$ è nulla. La tesi segue facilmente notando che al tempo 0 $\partial_{\mathbf{z}}\Psi_t$ è la matrice identità, che è simplettica, dunque $A_0 = J$, e che $\frac{d}{dt}A_t = 0$, e dunque $A_t = J$ per ogni t .

Siamo dunque in grado di enunciare e dimostrare il teorema di Noether per i sistemi hamiltoniani. Per semplicità considero solo il caso indipendente dal tempo.

Teorema 5.6. Teorema di Noether

L'hamiltoniana H è invariante per un gruppo a un parametro di diffeomorfismi simplettici, se e solo se esiste una funzione K tale che il gruppo è il flusso di hamiltoniana K e K è un integrale primo per H .

La prima parte del teorema è il contenuto del teorema precedente. Devo solo provare che H è invariante per il flusso di hamiltoniana K se e solo se $\{H, K\} = 0$. Infatti, detto Ψ_t il flusso

$$0 = \partial_t H(\Psi_t(\mathbf{z})) = \partial_{\mathbf{z}}H \cdot J \partial_{\mathbf{z}}K \text{ sse } \{H, K\} = 0$$

6 L'equazione di Hamilton-Jacobi

6.1 Funzioni generatrici

La caratterizzazione delle trasformazioni simplettiche attraverso la chiusura della forma $\mathbf{p} \cdot d\mathbf{q} - \mathbf{P} \cdot d\mathbf{Q}$ permette di **costruire** trasformazioni canoniche. Vediamo come.

Consideriamo, come esempio, la funzione $F = F(\mathbf{q}, \mathbf{Q}) = \mathbf{q} \cdot \mathbf{Q}$, che ha differenziale

$$dF = \mathbf{Q} \cdot d\mathbf{q} + \mathbf{q} \cdot d\mathbf{Q}$$

Chiediamoci ora se esistono due funzioni di $\mathbf{p}(\mathbf{q}, \mathbf{Q})$ e $\mathbf{P}(\mathbf{q}, \mathbf{Q})$ tali che

$$\mathbf{p} \cdot d\mathbf{q} - \mathbf{P} \cdot d\mathbf{Q} = dF = \mathbf{Q} \cdot d\mathbf{q} + \mathbf{q} \cdot d\mathbf{Q} \tag{6.1}$$

La risposta è evidentemente sì, e

$$\mathbf{p} = \mathbf{Q}, \quad \mathbf{P} = -\mathbf{q}$$

In questo modo risulta definita la trasformazione

$$\mathbf{P} = -\mathbf{q}, \quad \mathbf{Q} = \mathbf{p}$$

che è in effetti la trasformazione canonica che scambia coordinate e momenti (a meno di un segno). D'altra parte, l'uguaglianza (6.1) è verificata anche pensando che le variabili indipendenti siano (\mathbf{q}, \mathbf{p}) mentre \mathbf{Q}, \mathbf{P} sono date dal cambiamento di variabile. Infatti $F(\mathbf{q}, \mathbf{Q}(\mathbf{q}, \mathbf{p})) = \mathbf{q} \cdot \mathbf{p}$ e

$$\mathbf{p} \cdot d\mathbf{q} - \mathbf{P} \cdot d\mathbf{Q} = \mathbf{p} \cdot d\mathbf{q} + \mathbf{q} \cdot d\mathbf{p} = d(\mathbf{p} \cdot \mathbf{q}) = dF$$

Questo esempio dovrebbe rendere evidente che vale il seguente teorema

Teorema 6.1. Funzione generatrice $F(\mathbf{q}, \mathbf{Q}, t)$

Sia $F = F(\mathbf{q}, \mathbf{Q}, t)$ una funzione regolare, con la matrice $\partial_{\mathbf{q}\mathbf{Q}}^2 F$ non singolare. Allora le equazioni

$$\begin{aligned}\mathbf{p} &= \partial_{\mathbf{q}} F(\mathbf{q}, \mathbf{Q}, t) \\ -\mathbf{P} &= \partial_{\mathbf{Q}} F(\mathbf{q}, \mathbf{Q}, t)\end{aligned}$$

definiscono un cambiamento di variabili che è *simplettico* per ogni t .

Inoltre, data $H = H(\mathbf{q}, \mathbf{p}, t)$, l'hamiltoniana per il sistema nelle nuove variabili è

$$K = H + \partial_t F$$

Infatti, la prima equazione permette di ottenere \mathbf{Q} in funzione di \mathbf{q} (a questo serve l'ipotesi di non singolarità della matrice delle derivate seconde incrociate). La seconda equazione permette di determinare \mathbf{P} in funzione di \mathbf{q} e \mathbf{Q} . La canonicità è garantita dal fatto che, in variabili \mathbf{q}, \mathbf{Q} , ovviamente vale

$$\mathbf{p} \cdot d\mathbf{q} - \mathbf{P} \cdot d\mathbf{Q} = dF$$

e questa stessa relazione rimane naturalmente vera anche se viene espressa in funzione di \mathbf{q} e \mathbf{p} . Poiché una trasformazione simplettica ha jacobiano simplettico, e le matrici simplettiche hanno determinate 1, segue che la mappa $\mathbf{Q} = \mathbf{Q}(\mathbf{q}, \mathbf{p}, t)$ e $\mathbf{P} = \mathbf{P}(\mathbf{q}, \mathbf{p}, t)$ è non singolare, dunque definisce effettivamente una trasformazione di coordinate.

Data H , la nuova hamiltoniana K si definisce imponendo l'uguaglianza

$$\mathbf{p} \cdot d\mathbf{q} - H dt = \mathbf{P} \cdot d\mathbf{Q} - K dt + dF$$

Dunque

$$K = H + \partial_t F$$

Notate che questa relazione tra H e K è più semplice rispetto a quella espressa nella (4.2), il motivo è che

$$\mathbf{p} \cdot d\mathbf{q} - H dt = \mathbf{P} \cdot d\mathbf{Q} - K dt + dF$$

letta nelle variabili \mathbf{q}, \mathbf{Q} afferma che

$$\partial_t F = K - H,$$

letta invece nelle variabili \mathbf{q}, \mathbf{p} afferma che

$$\partial_t F = K - H - \mathbf{P} \cdot \partial_t \mathbf{Q}$$

che è appunto la (4.2)

È possibile anche scegliere altre coppie di variabili indipendenti tra le $\mathbf{q}, \mathbf{p}, \mathbf{Q}, \mathbf{P}$. Infatti, sia $\mathbf{Q} = (\mathbf{q}, \mathbf{p}, t)$, $\mathbf{P}(\mathbf{q}, \mathbf{p}, t)$ simplettica per ogni t . Allora, per i teoremi precedenti, esiste $G = G(\mathbf{q}, \mathbf{p}, t)$ tale che a t fissato

$$\mathbf{p} \cdot d\mathbf{q} - \mathbf{P} \cdot d\mathbf{Q} = dG \quad (6.2)$$

Se si possono considerare indipendenti \mathbf{q}, \mathbf{P} , allora, aggiungendo $d(\mathbf{P} \cdot \mathbf{Q})$ a entrambi i membri della (6.2), e definendo

$$S(\mathbf{q}, \mathbf{P}, t) = G(\mathbf{q}, \mathbf{p}(\mathbf{q}, \mathbf{P}), t) + \mathbf{P} \cdot \mathbf{Q}(\mathbf{q}, \mathbf{P}, t)$$

si ha che

$$\mathbf{p} \cdot d\mathbf{q} + \mathbf{Q} \cdot d\mathbf{P} = dS, \quad \text{cioè} \quad \begin{cases} \mathbf{p} = \frac{\partial S}{\partial \mathbf{q}} \\ \mathbf{Q} = \frac{\partial S}{\partial \mathbf{P}} \end{cases} \quad \det \partial_{\mathbf{q}\mathbf{P}} S \neq 0$$

Se si possono considerare indipendenti \mathbf{p}, \mathbf{P} , allora, aggiungendo $d(\mathbf{P} \cdot \mathbf{Q} - \mathbf{p} \cdot \mathbf{q})$ a entrambi i membri della (6.2), e definendo

$$F_3(\mathbf{p}, \mathbf{P}, t) = G(\mathbf{q}(\mathbf{p}, \mathbf{P}, t), \mathbf{p}, t) + \mathbf{P} \cdot \mathbf{Q}(\mathbf{p}, \mathbf{P}, t) - \mathbf{p} \cdot \mathbf{q}(\mathbf{p}, \mathbf{P}, t)$$

si ha che

$$-\mathbf{p} \cdot d\mathbf{q} + \mathbf{Q} \cdot d\mathbf{P} = dF_3, \quad \text{cioè} \quad \begin{cases} \mathbf{q} = -\partial_{\mathbf{p}} F_3 \\ \mathbf{Q} = \partial_{\mathbf{P}} F_3 \end{cases} \quad \det \partial_{\mathbf{p}\mathbf{P}} F_3 \neq 0$$

Infine, se si possono considerare indipendenti \mathbf{p}, \mathbf{Q} , allora, aggiungendo $-d(\mathbf{p} \cdot \mathbf{q})$ a entrambi i membri della (6.2), e definendo

$$F_4(\mathbf{p}, \mathbf{Q}, t) = G(\mathbf{q}(\mathbf{p}, \mathbf{Q}, t), \mathbf{p}, t) - \mathbf{p} \cdot \mathbf{q}(\mathbf{p}, \mathbf{Q}, t)$$

si ha che

$$-\mathbf{q} \cdot d\mathbf{p} - \mathbf{P} \cdot d\mathbf{Q} = dF_4, \quad \text{cioè} \quad \begin{cases} \mathbf{q} = -\partial_{\mathbf{p}} F_4 \\ \mathbf{P} = -\partial_{\mathbf{Q}} F_4 \end{cases} \quad \det \partial_{\mathbf{p}\mathbf{Q}} F_4 \neq 0$$

Si noti che in dimensione maggiore di 1, è possibile considerare scelte diverse per ogni coppia di variabili coniugate. Per esempio, provate a scrivere qual è la trasformazione indotta da una funzione generatrice $F = F(q_1, Q_1, q_2, P_2)$.

Osservazione: le varie funzioni generatrici non sono equivalenti. Per esempio, la trasformazione identica

$$\mathbf{Q} = \mathbf{q}, \quad \mathbf{P} = \mathbf{p}$$

è generata da $S = S(\mathbf{q}, \mathbf{P}) = \mathbf{q} \cdot \mathbf{P}$, ma non può essere generata da una funzione generatrice del tipo $F = F(\mathbf{q}, \mathbf{Q})$ (infatti non si possono scegliere \mathbf{q} e \mathbf{Q} come variabili indipendenti).

Usando le funzioni generatrici, è semplice mostrare come un cambiamento di coordinate induca un cambiamento negli impulsi.

Esempio: sia $\mathbf{x}(\mathbf{q})$ un diffeomorfismo da un dominio di \mathbb{R}^n in un dominio di \mathbb{R}^n , cioè $\mathbf{Q}(\mathbf{q}) = \mathbf{x}(\mathbf{q})$ sia cambiamento regolare delle variabili \mathbf{q} . Sia

$$S(\mathbf{q}, \mathbf{P}) = \mathbf{P} \cdot \mathbf{x}(\mathbf{q})$$

Poiché $\partial_{\mathbf{q}, \mathbf{P}} S = \partial_{\mathbf{q}} \mathbf{x}$, che per ipotesi è non singolare, S definisce la trasformazione canonica

$$\mathbf{Q} = \partial_{\mathbf{P}} S = \mathbf{x}(\mathbf{q}), \quad \mathbf{p} = (\partial_{\mathbf{q}} \mathbf{x})^t \mathbf{P}$$

Questa trasformazione solleva la trasformazione delle sole coordinate in una trasformazione nello spazio delle fasi.

È da notare che la funzione generatrice

$$S(\mathbf{q}, \mathbf{P}) = \mathbf{P} \cdot \mathbf{x}(\mathbf{q}) + g(\mathbf{q}),$$

dove g è una funzione scalare delle \mathbf{q} , genera la trasformazione canonica

$$\mathbf{Q} = \partial_{\mathbf{P}} S = \mathbf{x}(\mathbf{q}), \quad \mathbf{p} = (\partial_{\mathbf{q}} \mathbf{x})^t \mathbf{P} + \partial_{\mathbf{q}} g$$

che non coincide con la precedente. Dunque, come preannunciato, le trasformazioni canoniche sono “più numerose” delle trasformazioni indotte da trasformazioni delle sole \mathbf{q} .

6.2 Il metodo di Hamilton-Jacobi

Ho mostrato un metodo generale, quello delle funzioni generatrici, per trovare trasformazioni simplettiche. Il problema ora è trovare trasformazioni simplettiche **utili!**. Ci viene in aiuto una semplice considerazione: supponiamo di avere una funzione generatrice $S(\mathbf{q}, \mathbf{P}, t)$. Allora sappiamo che $\mathbf{p} = \partial_{\mathbf{q}} S$, e che la nuova hamiltoniana è $K = H + \partial_t S$. Possiamo provare a cercare S tale che la nuova hamiltoniana è nulla: in tal caso S deve soddisfare l'equazione di Hamilton-Jacobi:

$$\partial_t S(\mathbf{q}, \mathbf{P}, t) + H(\mathbf{q}, \partial_{\mathbf{q}} S(\mathbf{q}, \mathbf{P}, t), t) = 0. \quad (6.3)$$

Se una S siffatta esiste, nelle nuove variabili \mathbf{Q}, \mathbf{P} le equazioni di hamilton sono

$$\begin{aligned} \dot{\mathbf{Q}} &= \mathbf{0} \\ \dot{\mathbf{P}} &= \mathbf{0}, \end{aligned} \quad (6.4)$$

dunque nota la trasformazione si può risalire al moto nelle variabili q, p . Infatti, $(\mathbf{q}(t), \mathbf{p}(t))$, moto di hamiltoniana H , è determinato, in forma implicita, dal seguente sistema:

$$\begin{aligned} \mathbf{Q}(\mathbf{q}(t), \mathbf{p}(t), t) &= \mathbf{Q}(\mathbf{q}_0, \mathbf{p}_0, 0) \\ \mathbf{P}(\mathbf{q}(t), \mathbf{p}(t), t) &= \mathbf{P}(\mathbf{q}_0, \mathbf{p}_0, 0) \end{aligned}$$

Per Hamiltoniane indipendenti dal tempo, si può cercare una soluzione **separando le variabili**, cioè cercando la soluzione nella forma $S(\mathbf{q}, \mathbf{P}, t) = W(\mathbf{q}, \mathbf{P}) - P_n t$, dove P_n è l'ultimo impulso. L'equazione di HJ diventa l'**equazione caratteristica di HJ**:

$$H(\mathbf{q}, \partial_{\mathbf{q}} W(\mathbf{q}, \mathbf{P}, t)) = P_n. \quad (6.5)$$

Si noti che P_n è pari all'hamiltoniana, cioè all'energia del sistema.

In pratica, invece di cercare S dipendente dal tempo che renda nulla l'hamiltoniana, cerchiamo W , indipendente da t , che rende costante l'hamiltoniana. Infatti, se W risolve l'equazione caratteristica di Hamilton, con P_n ultimo nuovo impulso, W genera una trasformazione canonica nelle nuove variabili (Q, P) , per le quali l'hamiltoniana è P_n . Nelle nuove variabili le equazioni del moto sono banali:

$$\begin{aligned} \dot{Q}_i &= \partial_{P_i} K = 0 \quad i = 1, \dots, n-1 \\ \dot{Q}_n &= \partial_{P_n} K = 1 \\ \dot{P}_i &= -\partial_{Q_i} K = 0 \quad i = 1, \dots, n. \end{aligned} \quad (6.6)$$

Come si procede in pratica? Sempre e solo per separazione di variabili, cioè cercando la soluzione come somma di n funzioni ognuna delle quali dipende solo da una delle vecchie coordinate. In alcuni rari casi ci si riesce (sistemi integrabili) in generale no. Il fatto che si riesca a trovare la soluzione dipende dal fatto che dentro H la dipendenza delle variabili è *separata*.

Prima di vedere come funziona la separazione di variabili, guardiamo come funziona il metodo in un caso che sappiamo già risolvere, quello dell'oscillatore armonico.

Considero un oscillatore armonico di hamiltoniana

$$H = \frac{1}{2}(p^2 + q^2).$$

L'equazione caratteristica di HJ è

$$\frac{1}{2} \left(\frac{\partial W}{\partial q} + q \right)^2 = E$$

dove E (l'energia) sarà il nuovo impulso. Esplicitando rispetto alla derivata di W si ottiene

$$W(q, E) = \pm \int^q dq \sqrt{2E - q^2}$$

L'integrale si può calcolare esplicitamente, ma questo calcolo non è necessario per trovare la soluzione delle equazioni del moto. Infatti, se Q è la nuova variabile, la soluzione è

$$Q(t) = c + t,$$

dove c dipende dal dato iniziale, mentre la relazione tra Q e le vecchie variabili è data da

$$Q = \frac{\partial W}{\partial E} = \pm \int^q \frac{dq}{\sqrt{2E - q^2}}$$

Dunque la soluzione delle equazioni del moto si esprime mediante la **formula di quadratura**:

$$\pm \int^{q(t)} \frac{dq}{\sqrt{2E - q^2}} = t + c$$

dove il segno e la costante c dipendo da dato iniziale. Scegliendo $t = 0$ si determina c , dunque l'identità precedente si può scrivere

$$\pm \int_{q(0)}^{q(t)} \frac{dq}{\sqrt{2E - q^2}} = t.$$

Se p_0 è positivo va scelto il $+$, e l'identità vale fino a che $p(t)$ non si annulla, se p_0 è negativo va scelto il $-$. Dopo tempo di annullamento t_0 di $p(t)$, si riparte con la formula corretta, a partire da $q(t_0)$.

Si noti che in questo caso l'integrale si può calcolare esplicitamente, e vale $\arcsin q(t)/\sqrt{2E}$. Sostituendo questa espressione e invertendo rispetto a q si ottiene $q = \sqrt{2E} \sin(t + c)$.

Il valore del metodo di HJ si comprende nei rari casi di sistemi che si risolvono per quadrature, senza che esista un'evidente simmetria che renda ciclica qualche variabile. Faccio un esempio di questo fatto.

6.3 Esempio: moto centrale con termine di dipolo

Premetto qualche breve considerazione su cariche e dipoli. Ricordo che $-1/(4\pi|\mathbf{x}|)$ è la funzione di Green del laplaciano, cioè risolve il problema

$$\Delta - \frac{1}{4\pi|\mathbf{x}|} = \delta_0(d\mathbf{x}).$$

Sia $f(\mathbf{x})$ una densità di carica a supporto compatto, concentrata vicino all'origine. L'energia potenziale elettrostatica di f è

$$V(\mathbf{x}) = \frac{1}{4\pi} \int \frac{1}{|\mathbf{x} - \mathbf{y}|} f(\mathbf{y}) d\mathbf{y}$$

Supponiamo di osservare il sistema da molto lontano, o, equivalentemente, che f sia molto concentra, cioè analizziamo il caso $|\mathbf{y}|/|\mathbf{x}|$ piccolo sul supporto di f .

Sviluppando in serie di Taylor $1/|\mathbf{x} - \mathbf{y}|$, si ha, al primo ordine,

$$\frac{1}{|\mathbf{x} - \mathbf{y}|} = \frac{1}{|\mathbf{x}|} \frac{1}{|\hat{\mathbf{x}} - \mathbf{y}/|\mathbf{x}||} \approx \frac{1}{|\mathbf{x}|} + \frac{1}{|\mathbf{x}|^2} \hat{\mathbf{x}} \cdot \mathbf{y}$$

Sostituendo,

$$V(\mathbf{x}) \approx \frac{1}{4\pi|\mathbf{x}|} M_0 + \frac{1}{4\pi|\mathbf{x}|^2} \hat{\mathbf{x}} \cdot \mathbf{M}_1$$

dove M_0 è la carica totale, mentre \mathbf{M}_1 è il **momento di dipolo** di f .

$$M_0 = \int f(\mathbf{y}) d\mathbf{y}, \quad \mathbf{M}_1 = \int \mathbf{y} f(\mathbf{y}) d\mathbf{y}.$$

Quindi, all'approssimazione di ordine 0, è come se tutta la carica M_0 fosse concentrata nell'origine. Il termine successivo è il termine di dipolo. Si noti che se f ha segno costante, come nel caso gravitazionale, si può assumere che f abbia baricentro nell'origine, dunque il termine di dipolo non c'è. Compare invece se f ha segno variabile. Consideriamo un caso semplice: sia \mathbf{n} un versore, e mettiamo una carica positiva c in $\varepsilon\mathbf{n}$ e una negativa $-c$ in $-\varepsilon\mathbf{n}$. Il momento di dipolo viene $2c\varepsilon\mathbf{n}$. Questo valore rimane limitato per $\varepsilon \rightarrow 0$ solo se la carica diverge come $1/\varepsilon$. Per esempio, se $2c\varepsilon = 1$, per $\varepsilon \rightarrow 0$ la distribuzione di carica

$$\frac{1}{2\varepsilon} \delta_{\varepsilon\mathbf{n}}(d\mathbf{x}) - \frac{1}{2\varepsilon} \delta_{-\varepsilon\mathbf{n}}(d\mathbf{x})$$

converge alla distribuzione

$$\mathbf{n} \cdot \nabla \delta_0(d\mathbf{x})$$

Il campo di dipolo è esattamente quello generato dalla derivata direzionale della δ .

Fatta questa premessa, consideriamo un moto centrale kepleriano, perturbato da un dipolo nell'origine con momento diretto come x ; consideriamo cioè la lagrangiana

$$L = \frac{1}{2} \dot{\mathbf{x}}^2 + \frac{1}{|\mathbf{x}|} + \frac{x_1}{|\mathbf{x}|^3}$$

(sto pensando al moto di una carica negativa in orbita). In coordinate polari è

$$L = \frac{1}{2} (\dot{\rho}^2 + \rho^2 \dot{\vartheta}^2) + \frac{1}{\rho} + \frac{1}{\rho^2} \cos \vartheta$$

La corrispondente hamiltoniana è

$$H = \frac{p_\rho^2}{2} + \frac{p_\vartheta^2}{2\rho^2} - \frac{1}{\rho} - \frac{1}{\rho^2} \cos \vartheta$$

Come si nota facilmente, non ci sono variabili cicliche. L'equazione caratteristica di HJ è

$$\frac{1}{2}(\partial_\rho W)^2 + \frac{1}{2\rho^2}(\partial_\vartheta W)^2 - \frac{1}{\rho} - \frac{1}{\rho^2} \cos \vartheta = E$$

dove E sarà uno dei nuovi impulsi. Cerco la soluzione nella forma

$$W(\rho, \vartheta) = A(\rho) + B(\vartheta)$$

Otengo

$$\frac{1}{2}(\partial_\rho A(\rho))^2 - \frac{1}{\rho} + \frac{1}{2\rho^2}(\partial_\vartheta B(\theta))^2 - \frac{1}{\rho^2} \cos \vartheta = E$$

che si può riscrivere come

$$\frac{1}{2}(\partial_\rho A(\rho))^2 - \frac{1}{\rho} + \frac{1}{\rho^2} \left(\left(\frac{1}{2} \partial_\vartheta B(\theta) \right)^2 - \cos \vartheta \right) = E$$

È abbastanza evidente che questa equazione si può risolvere solo ipotizzando che

$$\frac{1}{2}(\partial_\vartheta B(\theta))^2 - \cos \vartheta = J$$

con J che non dipende da ρ e da ϑ . In tal caso l'equazione per A diventa

$$\frac{1}{2}(\partial_\rho A(\rho))^2 - \frac{1}{\rho} + \frac{J}{2\rho^2} = E$$

Le due equazioni sono risolte da

$$B(\vartheta, J) = \pm \int^\vartheta d\vartheta \sqrt{2(J + \cos \vartheta)}$$

$$A(\rho, J, E) = \pm \int^\rho d\rho \sqrt{2 \left(E + \frac{1}{\rho} - \frac{J}{\rho^2} \right)}$$

Le nuove variabili sono

$$Q_J = \frac{\partial W}{\partial J} = \frac{\partial B}{\partial J} + \frac{\partial A}{\partial J} = \pm \int^\vartheta \frac{d\vartheta}{\sqrt{2(J + \cos \vartheta)}} \mp \int^\rho \frac{d\rho}{2\rho^2 \sqrt{2E - \frac{J}{\rho^2}}}$$

$$Q_E = \frac{\partial W}{\partial E} = \pm \int^\rho \frac{d\rho}{\sqrt{2 \left(E + \frac{1}{\rho} - \frac{J}{\rho^2} \right)}}$$

Poiché la soluzione delle equazioni del moto è data da $Q_E = t + c_1$ e $Q_J = c_2$, le formule di quadratura sono

$$\pm \int_{\vartheta(0)}^{\vartheta(t)} \frac{d\vartheta}{2\sqrt{J + 2 \cos \vartheta}} \mp \int_{\rho(0)}^{\rho(t)} \frac{d\rho}{2\rho^2 \sqrt{2E - \frac{J}{\rho^2}}} = 0$$

$$\pm \int_{\rho(0)}^{\rho(t)} \frac{d\rho}{\sqrt{2E - \frac{J}{\rho^2}}} = t$$

Abbiamo potuto portare il moto alle quadrature perché il metodo di HJ ci ha permesso di notare l'esistenza di un integrale primo, diverso dall'energia:

$$J = \frac{1}{2}p_\vartheta^2 - \cos \vartheta.$$

L'esistenza di questo integrale primo, insieme a $E = \frac{1}{2}p_\rho^2 - \frac{1}{\rho} + \frac{J}{\rho^2}$, permette anche di fare l'analisi qualitativa del moto. In particolare, disegnando "l'energia potenziale" $-\cos \vartheta$ è facile comprendere che $J \geq -1$, e le sue curve di livello nel piano (ϑ, p_ϑ) sono

- $J = -1$: il punto $\vartheta = 0$, $p_\vartheta = 0$;
- $J \in (-1, 1)$: una curva chiusa centrata nell'origine;
- $J = 1$: il punto $\vartheta = \pm\pi$, $p_\vartheta = 0$, e due curve aperte che sono asintoticamente vicine a questo punto, una con $p_\vartheta > 0$ e una con $p_\vartheta < 0$.
- $J > 1$: due curve chiuse, una con $p_\vartheta > 0$ e una con $p_\vartheta < 0$, in cui tutti i valori di ϑ sono assunti.

Se $J \leq 0$, il valore di E è illimitato dal basso, e si distinguono orbite in cui il punto cade nell'origine se $E \leq 0$, e orbite in cui il punto si allontana indefinitamente se $E > 0$. Se invece $J > 0$, detto $E_J = -2/J$,

- se $E = E_J$ si ha che $\rho = J/2$ e $p_\rho = 0$;
- se $E \in (E_J, 0)$ si ha una curva chiusa;
- se $E > 0$ si hanno curve illimitate.

7 Sistemi integrabili

Formalizzo i concetti espressi nel paragrafo precedente. Si chiama **integrale completo** dell'equazione caratteristica di Hamilton-Jacobi

$$H(\mathbf{q}, \partial_{\mathbf{q}}W) = E$$

una funzione $W = W(\mathbf{q}, \tilde{\mathbf{p}})$, dove $\tilde{\mathbf{p}}$ sono n parametri indipendenti da cui dipende la soluzione, con $E = E(\tilde{\mathbf{p}})$, e tali che

$$\det \partial_{\mathbf{q}\tilde{\mathbf{p}}}^2 W \neq 0$$

Se una tale funzione esiste, allora è ben definito il cambiamento di variabili dato, implicitamente, da

$$\begin{cases} \tilde{\mathbf{q}} = \partial_{\tilde{\mathbf{p}}}W(\mathbf{q}, \tilde{\mathbf{p}}) \\ \mathbf{p} = \partial_{\mathbf{q}}W(\mathbf{q}, \tilde{\mathbf{p}}) \end{cases}$$

L'hamiltoniana nelle nuove variabili è data da $E = E(\tilde{\mathbf{p}})$, le equazioni del moto diventano

$$\begin{cases} \dot{\tilde{\mathbf{p}}} = 0 \\ \dot{\tilde{\mathbf{q}}} = \partial_{\tilde{\mathbf{p}}}E(\tilde{\mathbf{p}}) \end{cases}$$

Poiché i nuovi impulsi sono costanti, anche $\partial_{\tilde{\mathbf{p}}}E(\tilde{\mathbf{p}})$ è un vettore costante, dunque il moto nelle nuove variabili è

$$\begin{cases} \tilde{\mathbf{p}}(t) = \tilde{\mathbf{p}}(0) \\ \tilde{\mathbf{q}}(t) = \tilde{\mathbf{q}}(0) + t \partial_{\tilde{\mathbf{p}}}E(\tilde{\mathbf{p}}) \end{cases}$$

7.1 Sistemi integrabili

Se l'equazione caratteristica di HJ

$$H(\mathbf{q}, \partial_{\mathbf{q}}W) = E$$

ha un integrale completo W , i nuovi impulsi sono degli integrali primi indipendenti del moto. Ne segue che se un sistema non ha n integrali primi indipendenti, non è possibile trovare un integrale completo delle equazioni di HJ.

Inoltre, se HJ ha soluzione, il moto nelle nuove variabili è particolarmente semplice, infatti i nuovi impulsi sono costanti, e nelle nuove coordinate il moto è rettilineo uniforme:

$$\begin{cases} \tilde{\mathbf{p}}(t) = \tilde{\mathbf{p}}(0) \\ \tilde{\mathbf{q}}(t) = \tilde{\mathbf{q}}(0) + t \partial_{\tilde{\mathbf{p}}}E(\tilde{\mathbf{p}}(0)) \end{cases}$$

Queste stesse formule indicano che il moto è stato **ridotto alle quadrature**, cioè il moto è descritto attraverso inversioni di funzioni (che in genere sono espresse mediante integrali):

$$\begin{cases} \tilde{\mathbf{p}}(\mathbf{q}(t), \mathbf{p}(t)) = \tilde{\mathbf{p}}(0) \\ \tilde{\mathbf{q}}(\mathbf{q}(t), \mathbf{p}(t)) = \tilde{\mathbf{q}}(\mathbf{q}(0), \mathbf{p}(0)) + t \partial_{\tilde{\mathbf{p}}}E(\tilde{\mathbf{p}}(0)) \end{cases}$$

Un sistema che si possa ridurre alle quadrature mediante HJ è detto **sistema integrabile**.

I sistemi integrabili sono sistemi con specifiche proprietà. Si noti, infatti, che i nuovi impulsi sono integrali primi del moto indipendenti (cioè non si può determinarne uno in funzione degli altri). Inoltre, per le regole di commutazione canonica,

$$\{\tilde{p}_i, \tilde{p}_j\}_{\tilde{\mathbf{q}}, \tilde{\mathbf{p}}} = 0$$

Dimostriamo l'esistenza di n integrali primi indipendenti in involuzione è anche una condizione sufficiente all'integrabilità. In particolare dimostreremo l'integrabilità locale di un sistema siffatto (teorema di Liouville), per poi descrivere la versione globale (teorema di Arnold-Liouville), che dà anche importanti informazioni qualitative sul moto,

È necessario fare una osservazione generale sugli integrali primi di un sistema di equazioni differenziali. Consideriamo il flusso Φ^t associato a un campo vettoriale \mathbf{u} in \mathbb{R}^m , e sia \mathbf{x}_0 un punto non stazionario, cioè $\mathbf{u}(\mathbf{x}_0) \neq \mathbf{0}$. Sia Π l'iperpiano che passa per \mathbf{x}_0 ed è ortogonale a $\mathbf{u}(\mathbf{x}_0)$. Consideriamo un intorno B di raggio r di \mathbf{x}_0 . Se r è abbastanza piccolo, per ogni $\mathbf{x} \in B$ esistono $t^- < 0 < t^+$ tale che per $t \in (t^-, t^+)$ si ha $\Phi^t(\mathbf{x}) \in B$ mentre $\Phi^{t^\pm}(\mathbf{x}) \in \partial B$. Inoltre, esiste $t_0 \in (t^-, t^+)$ tale che $\Phi^{t_0}(\mathbf{x}) \in \Pi$. Risulta dunque ben definita la funzione

$$\mathbf{F} : B \rightarrow \Pi$$

che a \mathbf{x} associa $\Phi^{t_0}(\mathbf{x})$, cioè che a \mathbf{x} associa il punto di intersezione della sua orbita con Π . È evidente che la funzione vettoriale a $m - 1$ componenti \mathbf{F} è conservata dal flusso. Quindi, intorno a un punto non stazionario, esistono $m - 1$ integrali primi del moto. Questa affermazione, però, è puramente locale, e non dà nessuna informazione in più rispetto all'esistenza locale delle soluzioni. Al contrario, gli integrali primi che permettono la risolubilità dei sistemi hamiltoniani sono globali, e possono esistere o meno. L'integrabilità è dunque una proprietà in un certo senso algebrica (vedi l'Arnold su questo punto) di alcuni (pochi) moti. Nei sistemi hamiltoniani questa proprietà è legata all'esistenza di n integrali primi indipendenti in involuzione.

7.2 Geometria simplettica

Premetto una breve sezione geometrica che mi servirà per generalizzare alcuni risultati.

In questa sezione chiamo **prodotto antiortogonale** il prodotto $\mathbf{z} \cdot J\mathbf{w}$ (che prima chiamavo prodotto simplettico).

Un sottospazio V è detto **nullo** se tutti i suoi vettori sono antiortogonali tra loro. Si dimostri per esercizio che V è nullo se e solo se V è ortogonale a

$$JV = \{J\mathbf{z} \mid \mathbf{z} \in V\}.$$

Teorema 7.1. Sulla dimensione dei sottospazi nulli

Se V è nullo, allora ha dimensione $\leq n$.

Infatti V e JV sono ortogonali, e hanno la stessa dimensione, che deve dunque essere al più n .

Se \mathbf{u}_i sono n vettori antiortogonali, allora $V = \text{span}\{\mathbf{u}_i\}_{i=1}^n$ è uno spazio nullo, come anche JV .

Indico con $\hat{\mathbf{q}}_i$ e $\hat{\mathbf{p}}_j$ i versori degli assi q_i e p_j . Chiamo i -esimo **piano coordinato** il piano generato da $\hat{\mathbf{q}}_i$ e $\hat{\mathbf{p}}_i$. Noto che l' i -esimo piano coordinato è invariante per J , perché

$$J\hat{\mathbf{q}}_i = -\hat{\mathbf{p}}_i \quad \text{e} \quad J\hat{\mathbf{p}}_i = \hat{\mathbf{q}}_i$$

Si consideri ora una qualunque scelta dei versori $\boldsymbol{\sigma}_i = \hat{\mathbf{q}}_i$ oppure $\boldsymbol{\sigma}_i = \hat{\mathbf{p}}_i$. Sia

$$\Pi_{\boldsymbol{\sigma}} = \text{span}\{\boldsymbol{\sigma}_i\}_{i=1}^n$$

Poiché $J\boldsymbol{\sigma}_i$ è ortogonale a $\boldsymbol{\sigma}_i$ (vedi sopra), $\Pi_{\boldsymbol{\sigma}}$ è nullo. Lo chiamerò “sottospazio coordinato nullo n -dimensionale”.

Posso ora enunciare e dimostrare il seguente risultato.

Teorema 7.2. *Sia V un sottospazio nullo n -dimensionale. Esiste almeno un sottospazio coordinato nullo n -dimensionale Π tale che*

$$V \cap \Pi = \{\mathbf{0}\}$$

e dunque

$$\mathbb{R}^{2n} = V \oplus \Pi$$

Considero il sottospazio nullo Π_0 generato da $\{\hat{\mathbf{p}}_i\}_{i=1}^n$. Sia k la dimensione di $V \cap \Pi_0$. Senza mancare di generalità (rinominando gli indici) posso supporre che

$$V \cap \Pi_0 = \text{span}\{\hat{\mathbf{p}}_i\}_{i=1}^k$$

Ciò è equivalente a dire che $\hat{\mathbf{p}}_i \in V$ se $i \leq k$, mentre $\hat{\mathbf{p}}_i \notin V$ se $i > k$. Poiché se $i \leq k$, $\hat{\mathbf{p}}_i \in V$ e V è nullo, $-\hat{\mathbf{q}}_i = J\hat{\mathbf{p}}_i$ è ortogonale a V , e dunque non gli appartiene. In questo modo abbiamo provato che se

$$\Pi = \text{span}\{\{\hat{\mathbf{q}}_i\}_{i=1}^k, \{\hat{\mathbf{p}}_i\}_{i=k+1}^n\}$$

allora $V \cap \Pi = \{\mathbf{0}\}$.

Teorema 7.3. *Sia V un sottospazio nullo n dimensionale generato da $\{\mathbf{u}_i\}_{i=1}^n$, vettori indipendenti e antiortogonali.*

Esiste una scelta di $\boldsymbol{\sigma}_i = \hat{\mathbf{q}}_i$ oppure $\boldsymbol{\sigma}_i = \hat{\mathbf{p}}_i$ tale che il minore $n \times n$ della matrice formata dai vettori \mathbf{u}_i rispetto alle componenti σ_i è non nullo.

Dimostro il teorema per la matrice JV ; la tesi è valida perché V è nullo se e solo se lo è JV . Usando il teorema precedente, esiste Π sottospazio nullo coordinato generato da $\hat{\boldsymbol{\sigma}}_i$, $i = 1, \dots, n$, tale che $V \oplus \Pi = \mathbb{R}^{2n}$. Sia ora $\mathbf{z} \in JV$. Esistono $\mathbf{v} \in V$ e $\mathbf{w} \in \Pi$ tale che

$$\mathbf{z} = \mathbf{v} + \mathbf{w}$$

Indico con \mathcal{P}_W il proiettore ortogonale sul sottospazio W . Poiché $\mathbf{z} \in JV$ e $\mathbf{v} \in V$, si ha che $\mathbf{v} = -\mathcal{P}_V \mathbf{w}$, dunque

$$\mathbf{z} = \mathbf{w} - \mathcal{P}_V \mathbf{w}$$

Si noti che l'operatore $\mathbf{I} - \mathcal{P}_V$ definito su Π è iniettivo, altrimenti esisterebbe un \mathbf{w} non nullo tale che

$$\mathbf{w} = \mathcal{P}_V \mathbf{w}$$

ma allora $\mathbf{w} \in V \cap \Pi$. Dunque $\mathbf{w} \rightarrow \mathbf{w} - \mathcal{P}_V \mathbf{w}$ è una biezione tra Π e JV .

Proiettiamo la relazione trovata su Π , con \mathcal{P}_Π :

$$\mathcal{P}_\Pi \mathbf{z} = \mathbf{w} - \mathcal{P}_\Pi \mathcal{P}_V \mathbf{w}$$

Mostriamo che anche questa è iniettiva. Se così non fosse, esisterebbe \mathbf{w} non nullo tale che

$$\mathbf{w} = \mathcal{P}_\Pi \mathcal{P}_V \mathbf{w}$$

ma questo è possibile se e solo se $\mathbf{w} \in V \cap \Pi$, cioè se \mathbf{w} è nullo.

Poiché \mathcal{P}_Π è iniettiva da JV a Π , se \mathbf{w}_i sono vettori linearmente indipendenti che generano JV , allora il minore che corrisponde alle coordinate di Π ha determinante non nullo (lascio questo dettaglio al lettore).

Un risultato più debole ma più semplice da dimostrare è il seguente.

Teorema 7.4. *Siano $\{\mathbf{u}_i\}_{i=1}^n$ indipendenti e antiortogonali. Esiste una trasformazione canonica lineare di nuove variabili \tilde{q}_i, \tilde{p}_i tali che*

$$JV = \text{span}\{J\mathbf{u}_i\}_{i=1}^n = \{\tilde{\mathbf{u}} | \tilde{p}_i = 0 \text{ per } i = 1, \dots, n\} = \{(\tilde{q}_1, \dots, \tilde{q}_n, 0, \dots, 0) | \tilde{q}_i \in \mathbb{R} \text{ per } i = 1, \dots, n\}.$$

Sia $\{\mathbf{x}_i\}_{i=1}^n$ una base ortonormale per $V = \text{span}\{\mathbf{u}_i\}_{i=1}^n$. Ne segue che $\{J\mathbf{x}_i\}_{i=1}^n$ è una base ortonormale per JV , infatti J conserva il prodotto scalare. Inoltre \mathbf{x}_i è ortogonale a $J\mathbf{x}_j$ perché V è ortogonale a JV , infatti $\mathbf{u}_i \cdot J\mathbf{u}_j = 0$. Indico con \mathbf{x}_i^q e \mathbf{x}_i^p le componenti \mathbf{q} e \mathbf{p} del vettore \mathbf{x}_i . Noto che la condizione di ortonormalità è

$$\delta_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j = \mathbf{x}_i^q \cdot \mathbf{x}_j^q + \mathbf{x}_i^p \cdot \mathbf{x}_j^p$$

mentre la condizione di prodotto simplettico nullo è

$$0 = \mathbf{x}_i \cdot J\mathbf{x}_j = \mathbf{x}_i^q \cdot \mathbf{x}_j^p - \mathbf{x}_i^p \cdot \mathbf{x}_j^q$$

Sia ora $\mathbf{z} = (q_1, \dots, q_n, p_1, \dots, p_n)$.

$$\tilde{p}_i = \mathbf{q} \cdot \mathbf{x}_i^q + \mathbf{p} \cdot \mathbf{x}_i^p = \mathbf{z} \cdot \mathbf{x}_i$$

$$\tilde{q}_i = \mathbf{q} \cdot \mathbf{x}_i^p - \mathbf{p} \cdot \mathbf{x}_i^q = \mathbf{z} \cdot J\mathbf{x}_i$$

Questa trasformazione è ortogonale, infatti la matrice che la definisce ha come prime n righe i vettori $J\mathbf{x}_i$ e come ultime n righe i vettori \mathbf{x}_i , e i vettori $\{\mathbf{x}_j, J\mathbf{x}_i\}_{i,j=1\dots n}$ sono una base ortonormale. Calcolo le parentesi di Poisson delle nuove coordinate:

$$\begin{aligned}\{\tilde{p}_i, \tilde{p}_j\} &= \mathbf{x}_i^q \cdot \mathbf{x}_i^p - \mathbf{x}_i^p \cdot \mathbf{x}_j^q = 0 \\ \{\tilde{q}_i, \tilde{q}_j\} &= -\mathbf{x}_i^p \cdot \mathbf{x}_i^q + \mathbf{x}_i^q \cdot \mathbf{x}_j^p = 0 \\ \{\tilde{q}_i, \tilde{p}_j\} &= \mathbf{x}_i^p \cdot \mathbf{x}_j^p - (-\mathbf{x}_i^q) \cdot \mathbf{x}_j^q = \delta_{ij}\end{aligned}$$

La dimostrazione si conclude osservando che $JV = \{\mathbf{z} \mid \forall i \mathbf{z} \cdot \mathbf{x}_i = 0\}$ ma la condizione $\mathbf{z} \cdot \mathbf{x}_i = 0$ è esattamente la condizione $\tilde{p}_i = 0$.

7.3 Integrabilità locale

Teorema 7.5. Liouville *Sia dato un sistema di hamiltoniana H indipendente dal tempo. Supponiamo che in un intorno di un punto $\mathbf{z}_0 = (\mathbf{q}_0, \mathbf{p}_0)$ esistano n funzioni $f_i = f_i(\mathbf{q}, \mathbf{p})$, con $i = 1, \dots, n$, tali che*

- sono integrali primi del moto, cioè $\{f_i, H\} = 0$ per ogni i ;
- sono in involuzione, cioè $\{f_i, f_j\} = 0$, per ogni i, j ;
- sono indipendenti, cioè la matrice $(\partial_{\mathbf{z}} f_1 \dots \partial_{\mathbf{z}} f_n)$ ha rango n in un intorno di \mathbf{z}_0

Indicando con \mathbf{f} il vettore formato da queste n funzioni $\mathbf{f} = (f_1 \dots f_n)$, esiste una funzione $h = h(\mathbf{f})$ tale che in un intorno di \mathbf{z}_0

$$H(\mathbf{q}, \mathbf{p}) = h(\mathbf{f}(\mathbf{q}, \mathbf{p}))$$

e una trasformazione canonica che ha \mathbf{f} come nuovi impulsi. Nelle nuove coordinate $(\tilde{\mathbf{q}}, \mathbf{f})$ il moto è ricondotto alle quadrature:

$$\tilde{\mathbf{q}}(t) = \tilde{\mathbf{q}}(0) + t \partial_{\mathbf{f}} h(\mathbf{f})$$

con \mathbf{f} e $\partial_{\mathbf{f}} h(\mathbf{f})$ vettori costanti.

Assumeremo inizialmente che in un intorno di \mathbf{z}_0

$$\det \partial_{\mathbf{p}} \mathbf{f} \neq 0$$

(si ricordi che per ipotesi c'è un minore $n \times n$ con determinante non nullo della matrice $\partial_{\mathbf{z}} \mathbf{f}$, che però potrebbe non essere $\partial_{\mathbf{p}} \mathbf{f}$). Da questa ipotesi, usando il teorema della funzione implicita, segue che è possibile esprimere le \mathbf{p} in funzione delle \mathbf{q} e delle \mathbf{f} . Esiste, cioè, una funzione $\mathbf{g} = \mathbf{g}(\mathbf{q}, \mathbf{f})$ tale che

$$\mathbf{g}(\mathbf{q}, \mathbf{f}(\mathbf{q}, \mathbf{p})) = \mathbf{p}$$

Differenziando in \mathbf{p} e in \mathbf{q} si ha che

$$\begin{aligned}\partial_{\mathbf{f}} \mathbf{g} \partial_{\mathbf{p}} \mathbf{f} &= \mathbf{I} \\ \partial_{\mathbf{q}} \mathbf{g} + \partial_{\mathbf{f}} \mathbf{g} \partial_{\mathbf{q}} \mathbf{f} &= 0\end{aligned}$$

Usando la prima nella seconda, si ottiene

$$\partial_{\mathbf{q}} \mathbf{g} = -\partial_{\mathbf{p}} \mathbf{f}^{-1} \partial_{\mathbf{q}} \mathbf{f} = 0 \tag{7.1}$$

La condizione di involuzione implica che la forma differenziale

$$\mathbf{g}(\mathbf{f}, \mathbf{q}) \cdot d\mathbf{q}$$

è chiusa. La condizione di chiusura, infatti, equivale a

$$\partial_{\mathbf{q}}\mathbf{g} = \partial_{\mathbf{q}}\mathbf{g}^t \quad (7.2)$$

Usando la (7.1) si ottiene che la (7.2) è equivalente a

$$\partial_{\mathbf{p}}\mathbf{f}^{-1} \partial_{\mathbf{q}}\mathbf{f} = \partial_{\mathbf{q}}\mathbf{f}^t (\partial_{\mathbf{p}}\mathbf{f}^t)^{-1}$$

che, moltiplicando a sinistra per $\partial_{\mathbf{p}}\mathbf{f}$ e a destra per $\partial_{\mathbf{p}}\mathbf{f}^t$, è equivalente a

$$\partial_{\mathbf{q}}\mathbf{f} \partial_{\mathbf{p}}\mathbf{f}^t = \partial_{\mathbf{p}}\mathbf{f} \partial_{\mathbf{q}}\mathbf{f}^t$$

L'elemento di matrice ij del membro di sinistra è $\partial_{\mathbf{q}}\mathbf{f}_i \cdot \partial_{\mathbf{p}}\mathbf{f}_j$, quello del membro di destra è $\partial_{\mathbf{p}}\mathbf{f}_i \cdot \partial_{\mathbf{q}}\mathbf{f}_j$, dunque la condizione è equivalente a $\{f_i, f_j\} = 0$ per ogni i e j .

La chiusura della forma $\mathbf{g} \cdot d\mathbf{q}$ permette di definire, localmente, la sua primitiva

$$S(\mathbf{q}, \mathbf{f}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{g}(\mathbf{q}, \mathbf{f}) \cdot d\mathbf{q}$$

dove l'integrale è esteso a un qualunque cammino intorno a \mathbf{z}_0 che parte da q_0 e arriva in \mathbf{q} . Si noti che stiamo in pratica definendo

$$S(\mathbf{q}, \mathbf{f}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{p} \cdot d\mathbf{q}$$

dove \mathbf{p} diventano funzioni delle \mathbf{q} , fissando le \mathbf{f} . Per definizione,

$$\partial_{\mathbf{q}}S = \mathbf{g} \quad \text{e} \quad \partial_{\mathbf{f},\mathbf{q}}^2 S = \partial_{\mathbf{f}}\mathbf{g} = \partial_{\mathbf{p}}\mathbf{f}^{-1}$$

che sono tutte matrici non singolari per le ipotesi fatte. Dunque S definisce una trasformazione canonica, attraverso

$$\begin{aligned} \mathbf{p} &= \mathbf{g}(\mathbf{q}, \mathbf{f}) \\ \tilde{\mathbf{q}} &= \partial_{\mathbf{f}}S(\mathbf{q}, \mathbf{f}) \end{aligned}$$

dove le nuove coordinate sono le $\tilde{\mathbf{q}}$, mentre i nuovi impulsi sono gli integrali primi \mathbf{f} . Sia $\tilde{H} = \tilde{H}(\tilde{\mathbf{q}}, \mathbf{f})$ l'hamiltoniana nelle nuove variabili. Scriviamo nelle nuove variabili il fatto che f_i sono integrali primi:

$$0 = \{\tilde{H}, f_i\} = \partial_{\tilde{q}_i}\tilde{H}$$

Ma allora \tilde{H} non dipende dalle nuove coordinate $\tilde{\mathbf{q}}$, cioè \tilde{H} è funzione solo dei nuovi impulsi. Ne segue la banalità del moto in $\tilde{\mathbf{q}}$:

$$\dot{\tilde{\mathbf{q}}} = \partial_{\mathbf{f}}\tilde{H}(\mathbf{f})$$

e il membro di destra non dipende dal tempo.

Dimostro ora che posso rimuovere l'ipotesi che il minore non singolare della matrice delle derivate delle f_i sia proprio quello che si ottiene derivando negli impulsi p_j .

Teorema 7.6. *Siano $\{f_i\}_{i=1}^n$ integrali primi indipendenti e in involuzione in un intorno di un punto \mathbf{z}_0 . Intorno a \mathbf{z}_0 esiste una trasformazione canonica tale che $\partial_{\mathbf{p}}\mathbf{f}$ ha determinante non nullo.*

Senza mancare di generalità, assumo $\mathbf{z}_0 = \mathbf{0}$. Indico con

$$M = \{\mathbf{z} \mid f_i(\mathbf{z}) = f_i(\mathbf{0}), i = 1, \dots\}$$

l'insieme che si ottiene fissando i valori delle f_i al valore che assumo in $\mathbf{0}$. Intorno a $\mathbf{0}$, M è una varietà differenziabile. Siano $\mathbf{u}_i = \partial_{\mathbf{z}}f_i(\mathbf{0})$. Per le ipotesi sulle f_i , questi vettori sono linearmente indipendenti e hanno prodotto simplettico nullo. Noto che lo spazio V generato da \mathbf{u}_i è il sottospazio n -dimensionale ortogonale a M in $\mathbf{0}$, mentre lo spazio JV generato da $J\mathbf{u}_i$ è il sottospazio n -dimensionale tangente a M in $\mathbf{0}$.

Per i risultati del paragrafo dedicato alla geometria simplettica, esiste una trasformazione lineare ortogonale e simplettica tale che lo spazio tangente in $\mathbf{0}$ è descritto da $\tilde{p}_i = 0$, per $i = 1 \dots n$. Poiché questo spazio è generato anche dalle $J\partial_{\mathbf{z}}f_i$, ne segue che $\partial_{\mathbf{q}}\mathbf{f}$ è la matrice nulla. Ma allora la condizione di indipendenza dei gradienti delle f_i è esattamente la condizione di non singolarità di $\partial_{\mathbf{p}}\mathbf{f}$.

Questo teorema permette di concludere che il teorema di integrabilità locale vale sotto l'ipotesi generale che $\partial_{\mathbf{z}}\mathbf{f}$ abbia rango n .

Come corollario, vale il seguente teorema generale.

Teorema 7.7. Non esistenza di $n+1$ funzioni indipendenti in involuzione *Se f_i sono n funzioni indipendenti in involuzione e g è una funzione in involuzione con le f_i , allora g è funzione di f_1, \dots, f_n .*

Infatti, il sistema di hamiltoniana g con gli n integrali primi in involuzione f_i è localmente integrabile, e dunque, come dimostrato sopra, g è funzione delle f_i .

7.4 Integrabilità globale

Teorema 7.8. Arnold-Liouville

Nelle ipotesi del teorema precedente, supponiamo che date le costanti c_i , $i = 1, \dots, n$ la varietà n -dimensionale

$$M_{\mathbf{c}} = \{\mathbf{z} \mid f_i(\mathbf{z}) = c_i, i = 1, \dots\}$$

sia connessa e compatta e lo jacobiano $\partial_{\mathbf{z}}\mathbf{f}$ sia di rango n in tutti i punti della varietà $M_{\mathbf{c}}$. Si noti che $M_{\mathbf{c}}$ è invariante per il moto, poiché le f_i sono costanti del moto.

Allora:

- $M_{\mathbf{c}}$ è diffeomorfa a un toro n -dimensionale.
- su $M_{\mathbf{c}}$ si possono scegliere n variabili angolari $\vartheta_i \in [0, 2\pi]$
- in queste variabili il moto è dato da

$$\vartheta(t) = \vartheta(0) + t\boldsymbol{\omega}$$

dove $\boldsymbol{\omega}$ è un vettore costante (vettore delle frequenze)

In questo moto sul toro, ogni variabile ϑ_i si muove con velocità uniforme, e torna al valore iniziale dopo un periodo $T_i = 2\pi/\omega_i$. Per questo fatto prende il nome di **moto quasi periodico sul toro**.

Il punto chiave è considerare i flussi Ψ_i^t di hamiltoniana f_i . Poiché le f_i sono in involuzione, la varietà M_c è invariante anche per questi flussi. Inoltre, sempre per l'involuzione, i flussi commutano. Sia ora $\mathbf{z} \in M_c$, e consideriamo la mappa

$$\Psi^{s_1 s_2 \dots s_n}(\mathbf{z}) = \Psi_1^{s_1} \circ \dots \circ \Psi_n^{s_n}(\mathbf{z})$$

Noto che

$$\partial_{s_i} \Psi^{\mathbf{s}} = J \partial_{\mathbf{z}} f_i(\Psi^{\mathbf{s}})$$

infatti, poiché i flussi commutano, posso far agire $\Psi_i^{s_i}$ per ultimo, e la sua derivata è proprio $J \partial_{\mathbf{z}} f_i$ calcolato sul flusso.

Dunque per valori sufficientemente vicini a $\mathbf{0}$, $\mathbf{s} = (s_1, s_2, \dots, s_n)$ è un sistema di coordinate locali per M_c intorno a \mathbf{x} , infatti i vettori $J \partial_{\mathbf{z}} f_i$ sono una base per lo spazio tangente a M_c (per l'indipendenza delle f_i).

Diamo alcune proprietà per $\Psi^{\mathbf{s}}$.

Rapporto con il flusso Φ^t di hamiltoniana H Vale

$$\Phi^t = \Psi^{t \partial_{\mathbf{f}} H}$$

Infatti, per l'integrabilità locale H è funzione delle sole f , dunque il vettore $\partial_{\mathbf{f}} H$ dipende solo da \mathbf{f} , ed è dunque costante su M_c . Derivo in t . A sinistra ho

$$\partial_t \Phi^t(\mathbf{z}) = J \partial_{\mathbf{z}} H|_{\Phi^t(\mathbf{z})} = \sum_j \partial_{f_i} H J \partial_{\mathbf{z}} f_j \Big|_{\Phi^t(\mathbf{z})}$$

A destra ho

$$\partial_t \Psi^{t \partial_{\mathbf{f}} H}(\mathbf{z}) = \sum_j \partial_{f_i} H J \partial_{\mathbf{z}} f_j \Big|_{\Psi^{t \partial_{\mathbf{f}} H}(\mathbf{z})}$$

Dunque i due flussi verificano la stesso sistema di equazioni differenziali.

Suriettività di $\Psi^{\mathbf{s}}(\mathbf{z}_0)$

Sia $\mathbf{z}(\lambda)$ una curva regolare che unisce \mathbf{z}_0 a un punto $\mathbf{z} \in M_c$. Allora esiste $\mathbf{s}(\lambda)$ tale che $\Psi^{\mathbf{s}(\lambda)} = \mathbf{z}(\lambda)$.

Infatti, siano $a_i(\lambda)$ tali che

$$\sum_i a_i(\lambda) J \partial_{\mathbf{z}} f_i(\mathbf{z}(\lambda)) = \partial_{\lambda} \mathbf{z}(\lambda)$$

(il vettore \mathbf{a} esiste perché $\partial_{\lambda} \mathbf{z}(\lambda)$ è tangente a M_c in $\mathbf{z}(\lambda)$ e i vettori $\{J \partial_{\mathbf{z}} f_i(\mathbf{z}(\lambda))\}_{i=1}^n$ sono una base dello spazio tangente in $\mathbf{z}(\lambda)$). Sia dunque $\mathbf{s}(\lambda)$

$$\mathbf{s}(\lambda) = \int_0^{\lambda} \mathbf{a}(\lambda) d\lambda.$$

Ne segue che

$$\Psi^{\mathbf{s}(\lambda)} = \mathbf{z}(\lambda)$$

(lo si dimostri per esercizio, in modo analogo al punto precedente).

Non iniettività di $\Psi^{\mathbf{s}}(\mathbf{z}_0)$

Poiché $M_{\mathbf{c}}$ è compatta, $\Psi^{\mathbf{s}}(\mathbf{z}_0)$ non può essere iniettiva. Dunque esistono $\mathbf{s}_1 \neq \mathbf{s}_2$ tali che $\Psi^{\mathbf{s}_1}(\mathbf{z}_0) = \Psi^{\mathbf{s}_2}(\mathbf{z}_0)$. Ma allora per la proprietà di flusso, se $\mathbf{s}_0 = \mathbf{s}_1 - \mathbf{s}_2$

$$\Psi^{\mathbf{s}_0}(\mathbf{z}_0) = \mathbf{z}_0$$

cioè esiste \mathbf{s}_0 non nullo per cui \mathbf{z}_0 è un punto fisso di $\Psi^{\mathbf{s}_0}$. Poiché per ogni \mathbf{z} esiste \mathbf{s} tale che $\mathbf{z} = \Psi^{\mathbf{s}}(\mathbf{z}_0)$, ne segue che

$$\mathbf{z} = \Psi^{\mathbf{s}}(\mathbf{z}_0) = \Psi^{\mathbf{s}}(\Psi^{\mathbf{s}_0}(\mathbf{z}_0)) = \Psi^{\mathbf{s}_0}(\Psi^{\mathbf{s}}(\mathbf{z}_0)) = \Psi^{\mathbf{s}}(\mathbf{z})$$

Quindi è non vuoto l'insieme

$$G = \{\mathbf{s} \in \mathbb{R}^n : \Psi^{\mathbf{s}}(\mathbf{x}) = \mathbf{x} \quad \forall \mathbf{x}\}$$

G è un sottogruppo discreto di \mathbb{R}^n

È facile verificare che G è un sottogruppo di \mathbb{R}^n rispetto alla somma vettoriale. Mostriamo che G è un sottogruppo discreto di \mathbb{R}^n , cioè esiste $r > 0$ tale che se $\mathbf{s} \in G$ e $|\mathbf{s}| < r$ allora $\mathbf{s} = \mathbf{0}$ (da cui segue, per la proprietà di gruppo, che due punti di G diversi distano più di r). Infatti, sia $\mathbf{s}^k \rightarrow 0$, con $\mathbf{s}^k \in G$. Poiché

$$\mathbf{z} = \Psi^{\mathbf{s}^k}(\mathbf{z}) = \mathbf{z} + \sum_i s_i^k J \partial_{\mathbf{z}} f_i(\mathbf{z}) + O(|\mathbf{s}^k|^2)$$

sia ha che

$$\sum_i \frac{s_i^k}{|\mathbf{s}^k|} J \partial_{\mathbf{z}} f_i = O(|\mathbf{s}^k|)$$

Per sottosequenze $\mathbf{s}^k/|\mathbf{s}^k|$ converge a un versore \mathbf{v} , e dunque, passando al limite nell'espressione precedente, si ottiene una combinazione lineare nulla a coefficienti non nulli dei vettori $J \partial_{\mathbf{z}} f_i(\mathbf{z})$, in contraddizione con l'indipendenza.

G è generato da n vettori indipendenti di \mathbb{R}^n

Poiché G è un sottogruppo discreto di \mathbb{R}^n , esistono $m \leq n$ vettori linearmente indipendenti $\mathbf{v}_i \in \mathbb{R}^n$, tali che

$$G = \left\{ \sum_{i=1}^m k_i \mathbf{v}_i \mid k_i \in \mathbb{Z} \right\}$$

(non dimostro questo punto, facile ma un po' lungo, vedi Arnold). Siano inoltre $\mathbf{v}_{m+1} \dots \mathbf{v}_n$ una base per l'ortogonale allo spazio generato da $\mathbf{v}_1 \dots \mathbf{v}_m$. Considero ora l'applicazione lineare

$$(\vartheta_1 \dots \vartheta_n) \rightarrow \mathbf{s}(\vartheta) = \frac{1}{2\pi} \sum_{i=1}^n \vartheta_i \mathbf{v}_i$$

Per costruzione

$$\Psi^{\mathbf{s}(\vartheta)}$$

è un diffeomorfismo tra $S^m \times \mathbb{R}^{n-m}$ e $M_{\mathbf{c}}$, dove S è la circonferenza unitaria. Ma, poiché $M_{\mathbf{c}}$ è compatta, $m = n$ e $M_{\mathbf{c}}$ è diffeomorfa al toro n -dimensionale.

Studiamo il moto di hamiltoniana H nelle variabili angolari ϑ definite in termini di \mathbf{s} . Ho già provato che nelle variabili \mathbf{s}_i il moto avviene a velocità costante $\partial_{f_i} H(\mathbf{f})$. Poiché la relazione tra le variabili \mathbf{s} e le variabili $\boldsymbol{\vartheta}$ è lineare, anche nelle variabili $\boldsymbol{\vartheta}$ il moto avviene a velocità costante, cioè esistono $\omega_i = \omega_i(\mathbf{f})$ tali che

$$\vartheta_i(t) = \vartheta_i(0) + t\omega_i$$

7.5 Moti quasi periodici

Consideriamo un moto quasi periodico sul toro $S \times S \times \dots \times S = T^n$, dato da

$$\boldsymbol{\vartheta}(t) = \boldsymbol{\vartheta}(0) + t\boldsymbol{\omega}$$

Le frequenze $\boldsymbol{\omega}$ si dicono **razionalmente indipendenti** se, per ogni $\mathbf{w} \in \mathbb{Z}^n$ non nullo, (ogni componente di \mathbf{w} è un intero) allora

$$\boldsymbol{\omega} \cdot \mathbf{w} \neq 0$$

Equivalentemente, se $\boldsymbol{\omega} \cdot \mathbf{w} = 0$ e $\mathbf{w} \in \mathbb{Z}^n$, allora $\mathbf{w} = \mathbf{0}$. Si noti che queste definizioni non cambiano se si considerano combinazioni lineari a coefficienti razionali invece che interi.

Descriviamo le proprietà di questi moti, considerando prima il caso bidimensionale.

Teorema 7.9. Moti quasi periodici su T^2

Il moto è periodico se e solo se le due "pulsazioni" ω_1 e ω_2 sono razionalmente dipendenti. In caso contrario, il moto è denso su T^2 .

Se c'è almeno un ω_i nullo, il moto è periodico e le frequenze sono banalmente razionalmente dipendenti. Siano dunque entrambe non nulle. Il moto nella variabile ϑ_i è periodico di periodo $T_i = 2\pi/\omega_i$, dunque il moto è periodico se e solo se T_1 e T_2 hanno rapporto razionale. Ma questo è vero se e solo se le due pulsazioni sono razionalmente dipendenti. Se questo non accade, consideriamo due intersezioni successive della traiettoria con l'asse $\vartheta_1 = 0$. Il valore di ϑ_2 cambia di $T_1\omega_2 = 2\pi\omega_2/\omega_1 = \alpha$ che è un numero che ha un rapporto irrazionale con 2π . Il sistema dinamico discreto su S definito da $\Psi^k\vartheta = \vartheta + k\alpha$ prende il nome di **anello di Jacobi**. Se α ha rapporto razionale con 2π , il sistema è periodico. In caso contrario, l'insieme $\{\Psi^k(\vartheta)\}_{k \in \mathbb{N}}$ è denso. Infatti, tra gli m punti (distinti) che si ottengono per $k = 1 \dots m$ ce ne sono almeno due che distano $2\pi/m$, siano essi $\Psi^{k+h}(\vartheta)$ e $\Psi^k(\vartheta)$ con $h > 0$. Ma allora $\Psi^h(\vartheta)$ dista da ϑ meno di $2\pi/m$ e dunque iterando Ψ^h si ricopre S^1 con punti che distano meno di $2\pi/m$. Per l'arbitrarietà di m segue la tesi.

Dalla densità del moto in ϑ_2 segue facilmente la densità del moto in S^2 .

Consideriamo il caso n dimensionale.

Teorema 7.10. Moto quasi periodico su T^n

- Se le pulsazioni sono razionalmente indipendenti, la funzione $t \rightarrow \boldsymbol{\vartheta}(0) + t\boldsymbol{\omega}$ è iniettiva e il moto è denso sul toro.
- Se invece le pulsazioni sono razionalmente dipendenti, sia m il massimo numero di pulsazioni razionalmente indipendenti. Allora il moto risulta denso su una sottovarietà diffeomorfa a un toro di dimensione m .

- In particolare, se $m = 1$, esiste $\nu \in \mathbb{R}$ e esistono n interi k_i tali che $\omega_i = k_i\nu$; in tal caso il moto è periodico e infatti ricopre un toro unidimensionale, cioè una varietà diffeomorfa a una circonferenza.

La terza parte è banale, la seconda non la dimostro. La prima è conseguenza di una proprietà più forte, che enuncio e di cui qui manca la dimostrazione (la aggiungerò...).

Si definisce **media spaziale** di una funzione f il valore

$$\langle f \rangle = \frac{1}{(2\pi)^n} \int_{T^n} f(\vartheta) d\vartheta$$

Si definisce **media temporale** il valore

$$\lim_{T \rightarrow +\infty} \int_0^T f(\vartheta_0 + t\omega) dt$$

Teorema 7.11. Ergodicità nel caso di pulsazioni razionalmente indipendenti *Se le pulsazioni sono razionalmente indipendenti, allora il moto è ergodico, cioè per ogni f integrabile secondo Riemann la sua media spaziale coincide con la media temporale.*

Per il “significato” di questo teorema e la dimostrazione (fatte comunque a lezione) leggete sull’Arnold paragrafo 51, o le dispense di Buttà-Negrini, o le “dispense II” di Caglioti.

Da questo teorema segue facilmente la densità del moto: se così non fosse, esisterebbe in punto del toro con un suo intorno A che ha intersezione nulla con la traiettoria. Poiché A è un insieme misurabile secondo Riemann (infatti $\mathcal{X}\{\vartheta \in A\}$ è una funzione integrabile secondo Riemann) allora $|A|/(2\pi)^n$ deve essere pari alla media temporale. Ma per ipotesi la traiettoria non passa mai da A , dunque la media temporale è nulla, da cui la contraddizione che dimostra il teorema.

7.6 Variabili azione-angolo

Nella dimostrazione del teorema di Arnold abbiamo prima considerato come coordinate locali su M_c i “tempi” s_i di evoluzione per i flussi di hamiltoniane f_i , e poi abbiamo dimostrato che esiste un sistema di coordinate angolari ϑ_i , che si ottiene mediante una trasformazione lineare degli s_i , che è biiettivo da $S^{\times n}$ in M_c . In entrambe queste costruzioni abbiamo considerato fissi i valori degli integrali primi. Ora mostreremo che questi cambiamenti di coordinate si estendono a trasformazioni canoniche.

Teorema 7.12. *Le coordinate coniugate agli integrali primi f_i sono i tempi s_i , ovvero*

$$\mathbf{s} = \partial_{\mathbf{f}} W(\mathbf{q}, \mathbf{f})$$

Do prima una prova locale. Supponiamo che intorno a \mathbf{z}_0 il minore $\partial_{\mathbf{p}} \mathbf{f}$ abbia determinante non nullo. Mostro per prima cosa che le variabili \mathbf{s} sono ben definite, anche al variare di \mathbf{f} . Sia $\mathbf{p}_0(\mathbf{q}_0, \mathbf{f})$ il valore di \mathbf{p} tale che $(\mathbf{q}_0, \mathbf{p}_0) \in M_{\mathbf{f}}$ (che esiste per piccole variazioni di \mathbf{f}). Considero il sistema di coordinate locali \mathbf{s} che va da un intorno di $\mathbf{0} \in \mathbb{R}^n$ a valori in $M_{\mathbf{f}}$ definite da

$$\Psi^{\mathbf{s}}(\mathbf{q}_0, \mathbf{p}_0(\mathbf{q}_0, \mathbf{f}))$$

dove $\Psi^{\mathbf{s}}$ è la composizione dei flussi generati dalle hamiltoniane f_i .

Sia $W(\mathbf{q}, \mathbf{f}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{p}(\mathbf{q}, \mathbf{f}) \cdot d\mathbf{q}$ la funzione generatrice che si utilizza nel teorema di integrabilità locale. Le nuove coordinate sono

$$\partial_{\mathbf{f}} W(\mathbf{q}, \mathbf{f}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \partial_{\mathbf{f}} \mathbf{p}^t d\mathbf{q}$$

Ora cambio variabile nell'integrale da $d\mathbf{q}$ alle variabili $d\mathbf{s}$, con le notazione

$$\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{Q}^s \\ \mathbf{P}^s \end{pmatrix} = \Psi^s(\mathbf{q}_0, \mathbf{p}_0(\mathbf{q}_0, \mathbf{f}))$$

Per definizione di Ψ^s

$$\partial_s \mathbf{Q}^s = \partial_{\mathbf{f}} \mathbf{p}^t$$

e dunque

$$d\mathbf{q} = \partial_{\mathbf{f}} \mathbf{p}^t ds$$

Ma allora, poiché $\partial_{\mathbf{f}} \mathbf{p}$ e $\partial_{\mathbf{p}} \mathbf{f}$ sono una l'inversa dell'altra, si ottiene

$$\partial_{\mathbf{f}} W(\mathbf{q}, \mathbf{f}) = \int_0^{s(\mathbf{q}, \mathbf{f})} ds = s(\mathbf{q}, \mathbf{f})$$

Dunque (\mathbf{s}, \mathbf{f}) sono coordinate canoniche coniugate intorno a un punto per cui $\partial_{\mathbf{p}} \mathbf{f}$ è non singolare. D'altra parte, come notato nel paragrafo precedente, si possono prima trasformare semplicemente \mathbf{q}, \mathbf{p} in modo che questa condizione sia soddisfatta, e poi ripetere la costruzione. Le variabili di arrivo saranno sempre \mathbf{f} e \mathbf{s} (con i tempi \mathbf{s} traslati opportunamente), perché la loro definizione dipende solo da $M_{\mathbf{f}}$ e dal punto di partenza, che non cambia (lascio i dettagli al lettore).

Le coordinate (\mathbf{s}, \mathbf{f}) non sono però globalmente definibili, perché le \mathbf{s} non sono iniettive. D'altra parte le variabili ϑ sono iniettive, se limitate alla loro periodicità. Ci si chiede dunque se esitano delle variabili canoniche con ϑ come nuove coordinate. La risposta è sì ed è sufficiente ridefinire gli integrali primi.

Si faccia variare ϑ_i tra 0 e 2π , tenendo fissi i valori delle altre variabili ϑ_j . In questo modo si descrive una curva γ_i sul $M_{\mathbf{f}}$ che non è omotopa al punto.

Teorema 7.13. Variabili azione-angolo

Siano

$$I_i = \frac{1}{2\pi} \int_{\gamma_i} \mathbf{p} \cdot d\mathbf{q}$$

le azioni.

I valori di I_i dipendono solo da \mathbf{f} , e non dalle ϑ_j . Inoltre $\partial_{\mathbf{f}} \mathbf{I}$ è non singolare. Le coordinate (ϑ, \mathbf{I}) sono ben definite in un intorno di $M_{\mathbf{f}}$ e sono canoniche.

Abbiamo dimostrato che la forma $\mathbf{p} \cdot d\mathbf{q}$ è chiusa su $M_{\mathbf{f}}$ intorno ai punti per cui $\det \partial_{\mathbf{p}} \mathbf{f} \neq 0$. D'altra parte dove questa condizione non è soddisfatta, si cambiano semplicemente variabili in modo che $\tilde{\mathbf{p}} \cdot d\tilde{\mathbf{q}}$ è chiusa. Ma poiché la trasformazione è simplettica e lineare, la differenza tra queste due forme è esatta, dunque $\mathbf{p} \cdot d\mathbf{q}$ è chiusa ovunque su $M_{\mathbf{c}}$ (sull'Arnold c'è una dimostrazione più geometrica e interessante.) Come conseguenza, qualunque deformazione continua della curva γ_i non modifica il valore di I_i . Dunque le I_i sono ben definite, e funzioni delle \mathbf{f} .

Calcoliamo come dipendono dalle \mathbf{f} . Passando come abbiamo già fatto all'integrazione delle derivate in \mathbf{f} nelle variabili \mathbf{s} ,

$$\partial_f I_i = \frac{1}{2\pi} \partial_f \int_{\gamma_i} \mathbf{p} \cdot d\mathbf{q} = \frac{1}{2\pi} \int_{\gamma_i} d\mathbf{s}$$

Poiché

$$\mathbf{s}(\boldsymbol{\vartheta}) = \frac{1}{2\pi} \sum_{k=1}^n \vartheta_k \mathbf{v}_k$$

per definizione delle γ_i :

$$\partial_f I_i = \frac{1}{(2\pi)^2} \sum_k \mathbf{v}_k \int_{\gamma_i} d\vartheta_k = \frac{1}{(2\pi)^2} \sum_k \mathbf{v}_k 2\pi \delta_{ki} = \frac{1}{2\pi} \mathbf{v}_i.$$

Poiché i \mathbf{v}_i sono n vettori linearmente indipendenti, la matrice $\partial_f I$ è non singolare su tutta la varietà. A questo punto, usando come funzione generatrice $W(\mathbf{q}, \mathbf{f}(\mathbf{I}))$ e ricordando che $\partial_f W = \mathbf{s}$, si ottiene facilmente che $\partial_I W = \boldsymbol{\vartheta}$.

Esercizio 4. Variabili azione-angolo per i moti unidimensionali

Sia $H = \frac{1}{2}p^2 + V(q)$, con $V(q)$ strettamente convessa e divergente a $\pm\infty$. Sia E l'energia, sia γ la curva nello spazio delle fasi data dall'equazione $H = E$. Si mostri che

$$\int_{\gamma} p dq = A(E) = \text{Area racchiusa da } \gamma$$

La funzione $A(E)$ è crescente in E , dunque invertibile, e $I(E) = A(E)/(2\pi)$. Dunque

$$H(I) = A^{-1}(2\pi I)$$

Dunque

$$\omega(I) = \partial_I H(I) = 2\pi/A'(E)$$

e quindi se $T(E) = 2\pi/\omega$ è il periodo del moto di energia E ,

$$T(E) = \partial_E A(E)$$

Esercizio 5. Variabili azione-angolo per l'oscillatore armonico

Sia $H = \frac{1}{2m}p^2 + \frac{k}{2}q^2$ l'hamiltoniana di un oscillatore armonico. Si determinino le variabili azione-angolo, ricordando che l'area di una ellisse di semiassi a e b è πab .

8 Il teorema del ritorno di Poincaré

Tra le principali conseguenze della conservazione della misura per il flusso di fase c'è il teorema del ritorno di Poincaré, che afferma che per un flusso a un parametro che conserva la misura, definito su una regione limitata, "quasi ogni" traiettoria ritorna arbitrariamente vicina al suo dato iniziale.

Teorema 8.1. teorema del ritorno I.

Sia Φ una biezione su Ω , insieme di misura limitata, misurabile con inversa misurabile, e che conserva la misura¹, cioè per ogni misurabile $A \subset \Omega$ $|\Phi^{-1}(A)| = |A| = |\Phi(A)|$.

Sia A un sottoinsieme misurabile di Ω , allora, per ogni $k \in \mathbb{N}$ esiste $n \geq k$ tale che

$$|A \cap \Phi^n(A)| > 0$$

e dunque l'intersezione è non vuota. In particolare, esiste $\mathbf{x} \in A$ tale che $\Phi^n(\mathbf{x}) \in A$.

Dimostrazione. Definisco $\Psi = \Phi^k$. Anche Ψ è biettiva e conserva la misura. Considero la sequenza di sottoinsiemi $\{\Psi^h(A)\}_{h \geq 0}$. Se le loro intersezioni avessero misura nulla si avrebbe

$$|\Omega| \geq \left| \bigcup_h \Psi^h(A) \right| = \sum_h |\Psi^h(A)| = \sum_h |A| = +\infty$$

che è assurdo poiché Ω è limitato. Dunque esiste $h \geq 0$ ed esiste $m \geq 1$ tali che

$$\Psi^h(A) \cap \Psi^{h+m}(A) \neq \emptyset$$

e ha misura non nulla. Poiché Ψ è biettiva, ne segue che

$$A \cap \Psi^m(A) \neq \emptyset$$

In termini di Φ , questa relazione si legge

$$A \cap \Phi^{mk} \neq \emptyset$$

Scegliendo $n = mk \geq k$ si ottiene la tesi.

Un corollario di questo teorema è il seguente, che vale se Φ è un omeomorfismo. Sia $\varepsilon > 0$; dirò che \mathbf{x} è un punto che ε -**ritorna** se per ogni $k \in \mathbb{N}$, esiste $n > k$ tale che $|\Phi^n(\mathbf{x}) - \mathbf{x}| < \varepsilon$.

Teorema 8.2. teorema del ritorno II *L'insieme dei punti che non ε -ritornano ha misura nulla.*

Dimostrazione. Sia N_ε l'insieme dei punti che non ε -ritornano, cioè

$$N_\varepsilon = \{\mathbf{x} \mid |\Phi^k(\mathbf{x}) - \mathbf{x}| \geq \varepsilon \text{ definitivamente in } k\} = \bigcup_{k \geq 1} N_\varepsilon^k$$

dove

$$N_\varepsilon^k = \{\mathbf{x} \mid \forall n \geq k \mid \Phi^n(\mathbf{x}) - \mathbf{x} \mid \geq \varepsilon\} = \bigcap_{n \geq k} \tilde{N}_\varepsilon^n$$

con

$$\tilde{N}_\varepsilon^n = \{\mathbf{x} \mid |\Phi^n(\mathbf{x}) - \mathbf{x}| \geq \varepsilon\}$$

\tilde{N}_ε^n è un chiuso, dunque è misurabile e dunque sono misurabili anche N_ε^k e N_ε .

Dimostro che N_ε^k ha misura nulla. Posso ricoprire Ω con una quantità al più numerabile di palle B_1, B_2, \dots di raggio $\varepsilon/2$. Sia $A = B_i \cap N_\varepsilon^k$, che è misurabile. Se per assurdo avesse

¹La condizione di conservazione della misura si esprime in genere sugli osservabili, ed è $\int \alpha(\Phi(\mathbf{x})) \, d\mathbf{x} = \int \alpha(\mathbf{x}) \, d\mathbf{x}$, per ogni funzione test α . È facile vedere che questa condizione è equivalente a $|\Phi^{-1}(A)| = |A|$ per ogni misurabile A

misura positiva, per il teorema di Poincaré esisterebbe $n > k$ tale che $\Phi^n(A) \cap A$ avrebbe misura positiva. Ma allora esisterebbero $\mathbf{x}, \mathbf{y} \in A$ tali che $x = \Psi^n(\mathbf{y})$ e dunque

$$|\mathbf{y} - \Psi^n(\mathbf{y})| = |\mathbf{y} - \mathbf{x}| < \varepsilon$$

perché $\mathbf{x}, \mathbf{y} \in A \subset B_i$. Ma allora $B_i \cap N_\varepsilon^k$ ha misura nulla per ogni i . Usando la σ -subadditività della misura di Lebesgue

$$N_\varepsilon^k = \bigcup_i B_i \cap N_\varepsilon^k \Rightarrow |N_\varepsilon^k| \leq \sum_i |B_i \cap N_\varepsilon^k| = 0$$

Di nuovo usando la σ -subadditività

$$N_\varepsilon = \bigcup_k N_\varepsilon^k \Rightarrow |N_\varepsilon| \leq \sum_k |N_\varepsilon^k| = 0$$

Per esercizio, si verifichi se anche questo teorema vale sotto l'ipotesi di sola misurabilità per Φ e la sua inversa (cioè rimuovendo l'ipotesi di continuità).

Infine, si può provare la seguente affermazione più forte. Dirò che $x \in \Omega$ **ritorna** se

$$\forall \varepsilon > 0, \forall k \in \mathbb{N}, \exists n \geq k \text{ tale che } |\Phi^n(\mathbf{x}) - \mathbf{x}| < \varepsilon$$

Detto in parole, x ritorna se ε -ritorna per ogni ε , ovvero se la traiettoria che parte da x torna infinite volte arbitrariamente vicino a x .

Teorema 8.3. teorema del ritorno II

L'insieme dei punti che non ritornano ha misura nulla.

Se il punto x ritorna, allora $1/n$ -ritorna per ogni $n \in \mathbb{N}$ positivo. Dunque l'insieme dei punti che ritornano è l'intersezione per $n \geq 1$ degli insiemi dei punti che $1/n$ -ritornano. Il complementare è l'insieme dei punti che non ritornano, ed è dunque l'unione per $n \geq 1$ dell'insieme dei punti che non $1/n$ -ritornano. Poiché questi insiemi hanno misura nulla, la loro unione numerabile ha misura nulla.

Per commenti sui paradossi del teorema del ritorno vedi le dispense di Buttà-Negrini. È da notare, comunque, che lo studio delle proprietà a tempi lunghi dei sistemi conservativi (hamiltoniani) sono l'inizio dello studio che porta alla descrizione dei sistemi mediante la Meccanica Statistica (parola chiave **ipotesi ergodica**), ma non dirò nulla in merito.