

Appunti per il corso di Meccanica Analitica

Giancarlo Benettin

Anno accademico 2014–2015

Queste note sono indirizzate agli studenti del corso di Meccanica Analitica, recentemente istituito a Padova per la laurea di primo livello in Fisica. Sono dedicate alla Meccanica di Hamilton e ad alcune sue applicazioni. Lo scopo è quello di mostrare che il formalismo hamiltoniano — facilmente percepito, a livello elementare, come una riscrittura un po' misteriosa e di incerta utilità del formalismo lagrangiano — è invece uno strumento potente e insostituibile per meglio descrivere e comprendere il mondo fisico, e nei limiti del possibile, che è anche una bella costruzione intellettuale.

Il punto di partenza, cioè quello che si dà qui per noto, è la meccanica analitica “elementare” che si studia al secondo anno: ovvero il formalismo lagrangiano e i primissimi elementi di meccanica hamiltoniana (le equazioni di Hamilton e il loro legame con le equazioni di Lagrange, la parentesi di Poisson); di questi ultimi comunque, per completezza, si fa nel primo paragrafo un breve richiamo.

Il punto di arrivo, purtroppo, non arriva a comprendere nessuno dei grandi successi moderni della meccanica hamiltoniana: ma di alcuni (teoria “KAM”, 1954; teoria di Nekhoroshev, 1976) si giunge, per così dire, alla soglia, e gli si dà almeno un'occhiata, anche se un po' di sfuggita. Lo studio potrà eventualmente continuare, per gli studenti interessati, in corsi più avanzati.

Lo spirito con cui le note sono state stese, che si raccomanda agli studenti di far proprio nello studio, è quello classico della Fisica Matematica: mantenersi fedeli al rigore del procedimento e del linguaggio matematico, dando allo stesso tempo enfasi al senso fisico del formalismo introdotto, nella convinzione che i due punti di vista (qui e in ogni campo della fisica matematica, o forse della fisica e della matematica), ben lungi dall'ostacolarsi, si completano a vicenda e sono entrambi necessari per capire.

La stesura, benché già rivista, va considerata ancora incompleta e provvisoria. Raccomando agli studenti di tenerne conto, e conto sulla loro collaborazione per far sì che la prossima versione sia migliore: qualunque commento — dalla segnalazione degli inevitabili errori all'indicazione di quello che non è abbastanza chiaro e si potrebbe dire meglio — è molto gradito.

G. B. — a.a. 2005–06

“Se chiudete la porta a tutti gli errori,
anche la verità ne resterà fuori”

R. Tagore

INDICE

1. Richiami sulle equazioni di Hamilton
 - 1.1 *Sistemi hamiltoniani*
 - 1.2 *Parentesi di Poisson*
2. Le trasformazioni canoniche
 - 2.1 *Nozione ed esempi*
 - 2.2 *Funzioni generatrici di trasformazioni canoniche*
 - 2.3 *Trasformazioni canoniche dipendenti dal tempo*
 - 2.4 *Qualche esercizio*
3. Sistemi hamiltoniani “integrabili”
 - 3.1 *Nozione di sistema integrabile*
 - 3.2 *Le variabili di azione–angolo per il pendolo*
 - 3.3 *Il teorema di Liouville–Arnol’d*
 - 3.4 *Il corpo rigido di Eulero–Poinsot come sistema hamiltoniano integrabile*
 - 3.5 *L’equazione di Hamilton–Jacobi*
4. Introduzione alla teoria hamiltoniana delle perturbazioni
 - 4.1 *Sistemi prossimi a sistemi integrabili*
 - 4.2 *La “stima a priori”*
 - 4.3 *Il “principio della media”*
 - 4.4 *Sistemi isocroni perturbati: un passo perturbativo*
 - 4.5 *Sistemi non isocroni; il modello dei rotatori*
 - 4.6 *I grandi teoremi della teoria hamiltoniana delle perturbazioni*
 - 4.7 *Invarianti adiabatici*

Appendici

- A. *Introduzione al formalismo geometrico per sistemi hamiltoniani*
- B. *Equazioni differenziali che preservano il volume*
- C. *Commutazione di flussi hamiltoniani*
- D. *Alcune dimostrazioni riguardanti le trasformazioni canoniche*
- E. *Risoluzione di alcuni esercizi*
- F. *Il teorema di Liouville–Arnol’d: dimostrazione dei punti i. e ii.*
- G. *Nozioni di base sulla dinamica del corpo rigido*
- H. *L’equazione di Van der Pol*
- I. *Il modello classico per la precessione degli equinozi*
- J. *Oltre i rotatori: un passo perturbativo per generici sistemi convessi*

Letture supplementari

- V.I. Arnol'd, *Metodi matematici della meccanica classica*, Editori Riuniti 1979.
- V.I. Arnol'd, *Metodi geometrici della teoria delle equazioni differenziali ordinarie*, capitolo 4, Editori Riuniti 1989.
- V.I. Arnol'd (editore), *Dynamical Systems III*, collezione *Encyclopaedia of Mathematical Science*, Springer 1985.
- H. Goldstein, *Classical Mechanics*, Addison Wesley 1980 (molto dettagliati i capp. 4 e 5 sul corpo rigido).
- F. Sheck, *From Newton's Laws to Deterministic Chaos*, Springer 2010.
- G. Benettin, *Introduzione ai sistemi dinamici*, Capitolo III, dispense non pubblicate (www.math.unipd.it/~benettin).
- G. Benettin, *The elements of Hamiltonian perturbation theory*, in: *Hamiltonian dynamics and frequency analysis*, volume curato da D. Benest, C. Froeschlé e E. Lega, Cambridge Scientific Publisher 2005.

1 Richiami sulle equazioni di Hamilton

In tutte queste note, e specificamente in questo e nel prossimo capitolo in cui si introducono le basi del formalismo hamiltoniano, il punto di vista adottato è quello cosiddetto locale: l'ambiente è un qualche dominio aperto contenuto in \mathbb{R}^{2n} , e tutto viene discusso con riferimento a opportune coordinate. E' un punto di vista sufficiente, come vedremo, a introdurre le idee di base della meccanica hamiltoniana e anche moderni risultati, belli e fondamentali, come il teorema KAM o il teorema di Nekhoroshev. E' anche l'unico possibile per chi non abbia già familiarità col linguaggio astratto delle varietà. Ma concettualmente la visione locale, confrontata con quella globale geometrica, è un un po' angusta. Una rivisitazione delle nozioni fondamentali da un punto di vista geometrico si trova nell'appendice A.

1.1 Sistemi hamiltoniani

Richiamiamo allora innanzitutto la nozione elementare locale di sistema hamiltoniano. Sia U un aperto di \mathbb{R}^{2n} (la dimensione pari è essenziale) munito di coordinate

$$(p, q) = (p_1, \dots, p_n, q_1, \dots, q_n),$$

e sia H una funzione regolare: $U \rightarrow \mathbb{R}$. Il sistema di $2n$ equazioni differenziali ordinarie

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i}, \quad i = 1, \dots, n, \quad (1.1)$$

è detto *sistema hamiltoniano* di hamiltoniana (o funzione di Hamilton) H ; le (1.1) sono dette *equazioni di Hamilton*. Useremo di regola la scrittura più agile, senza indici,

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = \frac{\partial H}{\partial p}.$$

Le variabili p sono dette *momenti coniugati* alle coordinate q ; n è detto *numero di gradi di libertà* del sistema; U è detto *spazio delle fasi*.

Possiamo utilmente introdurre la notazione compatta

$$x = (x_1, \dots, x_{2n}) = (p_1, \dots, p_n, q_1, \dots, q_n)$$

e scrivere le equazioni di hamilton nella forma

$$\dot{x} = X_H(x), \quad X_H = (X_{H,1}, \dots, X_{H,2n});$$

$X_H : U \rightarrow \mathbb{R}^{2n}$, detto *campo vettoriale hamiltoniano* associato all'hamiltoniana H , è dato evidentemente da

$$X_H = \left(-\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p} \right) = \left(-\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n}, \frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n} \right).$$

Se allora per una generica funzione $f : U \rightarrow \mathbb{R}$ denotiamo con $\partial_x f$ la $2n$ -pla delle sue derivate (i coefficienti del differenziale df):

$$\partial_x f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_{2n}} \right) = \left(\frac{\partial f}{\partial p_1}, \dots, \frac{\partial f}{\partial p_n}, \frac{\partial f}{\partial q_1}, \dots, \frac{\partial f}{\partial q_n} \right),$$

il legame tra X_H e $\partial_x H$ si scrive in modo compatto

$$X_H = \mathbb{E} \partial_x H, \quad (1.2)$$

ove \mathbb{E} è la matrice antisimmetrica

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

in cui \mathbb{O} e \mathbb{I} denotano rispettivamente la matrice nulla e l'identità $n \times n$. Si verifica immediatamente che risulta $\mathbb{E}^T = \mathbb{E}^{-1} = -\mathbb{E}$. Si ha perciò $\mathbb{E}^2 = -\mathbb{I}$, con \mathbb{I} identità $2n \times 2n$; la relazione ricorda $i^2 = -1$, ove i è l'unità immaginaria. \mathbb{E} è detta *identità symplettica*.¹

La soluzione delle equazioni di Hamilton al tempo t , con dato iniziale $(p, q) \in U$, verrà denotata $\Phi_H^t(p, q)$; Φ_H^t è detto *flusso* del sistema hamiltoniano. Supporremo sempre che il flusso sia *completo*, ovvero che $\Phi_H^t : U \rightarrow U$ sia definito per ogni $t \in \mathbb{R}$. In questo caso, grazie al fatto che il sistema è autonomo, ovvero H e di conseguenza le equazioni di Hamilton non dipendono esplicitamente da t , $\{\Phi_H^t, t \in \mathbb{R}\}$ è un gruppo:² l'identità è Φ_H^0 , e poi

$$(\Phi_H^t)^{-1} = \Phi_H^{-t}, \quad \Phi_H^{t+s} = \Phi_H^t \circ \Phi_H^s = \Phi_H^s \circ \Phi_H^t. \quad (1.3)$$

Una proprietà importante del flusso hamiltoniano, nota come *teorema di Liouville*, è che Φ_H^t preserva il volume nello spazio delle fasi, o più in generale preserva la misura di Lebesgue: per ogni $A \subset U$ (misurabile) si ha $\text{Vol}(\Phi_H^t(A)) = \text{Vol}(A)$; si veda l'appendice B per dettagli e ulteriori commenti, nonché per il *teorema del ritorno*, di Poincaré, uno degli aspetti curiosi ma in realtà profondi della meccanica hamiltoniana.

Nei sistemi Hamiltoniani (autonomi) si conserva l'energia, più precisamente l'hamiltoniana stessa è costante lungo le soluzioni:

$$\frac{dH}{dt} = 0, \quad H(\Phi_H^t(p, q)) = H(p, q). \quad (1.4)$$

Corrispondentemente le superfici di energia costante

$$\Sigma_E = \{(p, q) \in U : H(p, q) = E\}$$

sono invarianti: $\Phi_H^t(\Sigma_E) = \Sigma_E$.

Hamiltoniane che differiscono per una costante moltiplicativa $c \neq 0$ hanno gli stessi moti, a meno di un banale riscaldamento del tempo: si vede subito infatti che si ha³

$$\Phi_{cH}^t = \Phi_H^{ct}. \quad (1.5)$$

Una spontanea generalizzazione è quella al *caso non autonomo* in cui H dipende esplicitamente dal tempo: $H : U \times \mathbb{R} \rightarrow \mathbb{R}$, $(p, q, t) \mapsto H(p, q, t)$. Le equazioni di Hamilton (1.1) dipendono allora esse stesse dal tempo e la proprietà grupitale si perde. La (1.4) è sostituita dalla più generale

$$\frac{dH}{dt} = \frac{\partial H}{\partial t},$$

¹“Symplettico” è in effetti una variante di “complesso”, in cui il prefisso greco “συν” sostituisce il latino “cum”.

²Del tutto in generale, se $\dot{x} = X(x)$ è un'equazione differenziale autonoma in \mathbb{R}^n , il suo flusso Φ_X^t è un gruppo, precisamente si ha $\Phi_X^0 = \text{identità}$, $(\Phi_X^t)^{-1} = \Phi_X^{-t}$, $\Phi_X^{t+s} = \Phi_X^t \circ \Phi_X^s = \Phi_X^s \circ \Phi_X^t$. L'osservazione fondamentale è che se $\hat{x}(t)$ è soluzione particolare dell'equazione corrispondente al dato iniziale x , allora anche $\hat{x}'(t) = \hat{x}(t+s)$ è soluzione dell'equazione, con dato iniziale $\hat{x}'(0) = \hat{x}(s)$. Attenzione: se il sistema non è autonomo la proprietà non è più vera.

³Anche al di fuori del caso hamiltoniano, se Φ_X^t denota il flusso associato all'equazione differenziale $\dot{x} = X(x)$, si ha l'identità $\Phi_{cX}^t = \Phi_X^{ct}$. In ambito hamiltoniano vale la proprietà più generale che se $K(p, q) = F(H(p, q))$ con $F : \mathbb{R} \rightarrow \mathbb{R}$, allora K e H hanno ancora gli stessi moti a meno di un riscaldamento del tempo dipendente dall'energia, precisamente dato, sulla superficie di energia costante Σ_E , dal coefficiente $c = F'(E)$.

mentre continua a valere la conservazione del volume nello spazio delle fasi. Un buon punto di vista è quello di pensare un sistema non autonomo a n gradi di libertà come un (particolare) sistema autonomo in uno spazio esteso a $n + 1$ gradi di libertà, munito di coordinate $(p_1, \dots, p_{n+1}, q_1, \dots, q_{n+1}) = (p_1, \dots, p_n, A, q_1, \dots, q_n, \tau)$, di hamiltoniana

$$H^{\text{ext}}(p_1, \dots, p_n, A, q_1, \dots, q_n, \tau) = H(p_1, \dots, p_n, q_1, \dots, q_n, \tau) + A .$$

Con evidenza si ha $\dot{\tau} = 1$ e dunque i moti con dato iniziale $\tau(0) = 0$ danno $\tau(t) = t$; le equazioni di Hamilton relative a H^{ext} , per le variabili (p, q) , vanno allora a coincidere con quelle relative a H , e corrispondentemente le soluzioni del sistema esteso $(p(t), q(t))$ coincidono con le soluzioni del sistema non esteso. Le sorti della variabile ausiliaria A , soggetta all'equazione

$$\dot{A} = -\frac{\partial H^{\text{ext}}}{\partial \tau} = -\frac{\partial H}{\partial t} = -\frac{dH}{dt}$$

non interessano molto, se non per il fatto che la sua variazione nel tempo è opposta a quella di H (corrispondentemente H^{ext} si conserva).

I sistemi hamiltoniani si affacciano naturalmente nel corso dello studio della meccanica lagrangiana, e le equazioni di Hamilton appaiono allora come una possibile riformulazione delle equazioni di Lagrange, ad esse equivalenti. Ma è una visione riduttiva, ed è bene invece prendere il punto di vista, più ampio, in cui i due formalismi sono indipendenti e dotati di vita e interesse proprio, benché sia vero che molti sistemi fisici ammettono l'una e l'altra descrizione. Ricordiamo che si passa dal formalismo lagrangiano a quello hamiltoniano, e viceversa, attraverso la cosiddetta trasformata di Legendre. Telegraficamente: il legame tra le variabili lagrangiane (q, \dot{q}) e le variabili hamiltoniane (p, q) è dato da

$$p(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}} , \quad (1.6)$$

(in dettaglio $p_i = \frac{\partial L}{\partial \dot{q}_i}$, $i = 1, \dots, n$), mentre il legame tra L e H è

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

ove $\dot{q}(p, q)$ è la funzione che inverte la (1.6), mentre il punto denota il consueto prodotto scalare di n -ple.

Il passaggio da un formalismo all'altro si può fare se la (1.6) è invertibile: localmente basta $\det(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}) \neq 0$, mentre se si vuole (come è tipico) che l'inversione si estenda a ogni $\dot{q} \in \mathbb{R}^n$, una buona condizione sufficiente è la convessità di L nelle \dot{q} (assicurata in tutti i casi meccanici, quando l'energia cinetica è una forma quadratica definita positiva nelle \dot{q}).

Ricordiamo qualche esempio elementare di sistema hamiltoniano, scrivendo anche per raffronto la corrispondente lagrangiana.

- i) Un punto materiale di massa m sulla retta, soggetto a potenziale posizionale:

$$L(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - V(q) , \quad H(p, q) = \frac{1}{2m}p^2 + V(q) ;$$

$p = m\dot{q}$ è il momento lineare. Per $V = \frac{1}{2}m\omega^2 q^2$ si ha l'oscillatore armonico lineare di pulsazione ω .

ii) Il pendolo: indicando con ϑ la coordinata, si ha

$$L(\vartheta, \dot{\vartheta}) = \frac{1}{2}ml^2\dot{\vartheta}^2 + mgl \cos \vartheta, \quad H(p, q) = \frac{p^2}{2ml^2} - mgl \cos \vartheta;$$

$p = ml^2\dot{\vartheta}$ è il momento angolare.

iii) Il moto centrale piano. In coordinate polari piane (r, ϑ) si ha

$$L(r, \vartheta, \dot{r}, \dot{\vartheta}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\vartheta}^2) - V(r), \quad H(p_r, p_\vartheta) = \frac{p_r^2}{2m} + \frac{p_\vartheta^2}{mr^2} + V(r);$$

$p_r = m\dot{r}$ è la componente radiale del momento lineare, $p_\vartheta = mr^2\dot{\vartheta}$ è il momento angolare. Per $V(r) = \frac{1}{2}m\omega^2r^2$ si ha l'oscillatore armonico piano, per $V(r) = -Cm/r$ si ha il problema di Keplero.

iv) I tipici sistemi meccanici in cui L è della forma

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

(K energia cinetica, V energia potenziale, a matrice cinetica simmetrica e definita positiva): si ha $p = a\dot{q}$ e poi

$$H(p, q) = \frac{1}{2}p \cdot a^{-1}(q)p + V(q) = K + V.$$

Ma non è sempre così: ad esempio

v) Una carica elettrica e in un campo elettromagnetico assegnato: detti $\varphi(q, t)$ e $A(q, t)$ i potenziali scalare e vettore ($q = (q_1, q_2, q_3)$, $A = (A_1, A_2, A_3)$), si ha

$$L(q, \dot{q}, t) = \frac{m}{2}\dot{q}^2 + e(\dot{q} \cdot A - \varphi), \quad H(p, q, t) = \frac{1}{2m}(p - eA)^2 + e\varphi,$$

e il legame tra p e \dot{q} è $p = m\dot{q} + eA$.

- Vale la pena di osservare che, come risulta delle (1.1), il prodotto $p_i q_i$, per ogni coppia di variabili coniugate, ha la stessa dimensione fisica, precisamente la dimensione dell'hamiltoniana H moltiplicata per un tempo; nel caso tipico della fisica in cui H è un'energia, $p_i q_i$ è un'azione.

1.2 Parentesi di Poisson

In ambito hamiltoniano è naturale introdurre un'operazione binaria tra funzioni, detta *parentesi di Poisson*, definita per ogni coppia di funzioni (regolari) $U \rightarrow \mathbb{R}$ da

$$\{f, g\} = \frac{\partial f}{\partial q} \cdot \frac{\partial g}{\partial p} - \frac{\partial f}{\partial p} \cdot \frac{\partial g}{\partial q}. \quad (1.7)$$

In notazione compatta si ha

$$\{f, g\} = \partial_x f \cdot \mathbb{E} \partial_x g. \quad (1.8)$$

L'operazione è chiusa nello spazio delle funzioni infinitamente differenziabili: $U \rightarrow \mathbb{R}$, e gode di tre proprietà elementari:

$$\begin{aligned} \{f, g\} &= -\{g, f\} && \text{(antisimmetria)} \\ \{c_1 f_1 + c_2 f_2, g\} &= c_1 \{f_1, g\} + c_2 \{f_2, g\} && \text{(linearità)} \\ \{f_1 f_2, g\} &= f_1 \{f_2, g\} + \{f_1, g\} f_2 && \text{(regola di Leibnitz).} \end{aligned}$$

L'antisimmetria implica $\{f, f\} = 0$; più in generale, se $g(p, q) = G(f(p, q))$ con $G: \mathbb{R} \rightarrow \mathbb{R}$, allora $\{f, g\} = 0$. Un'ulteriore importante proprietà della parentesi di Poisson è l'*identità di Jacobi*: per ogni terna di funzioni f, g, h risulta⁴

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

La parentesi di Poisson interviene naturalmente nella meccanica Hamiltoniana grazie alla proprietà, elementare ma fondamentale,

$$\dot{f} = \{f, H\}, \quad (1.9)$$

ove \dot{f} è la derivata di f lungo il flusso:

$$\begin{aligned} \dot{f}(p, q) &= \left. \frac{d}{dt} f(\Phi_H^t(p, q)) \right|_{t=0} = \frac{\partial f}{\partial p} \cdot \dot{p} + \frac{\partial f}{\partial q} \cdot \dot{q} \\ &= -\frac{\partial f}{\partial p} \cdot \frac{\partial H}{\partial q} + \frac{\partial f}{\partial q} \cdot \frac{\partial H}{\partial p}. \end{aligned}$$

Se f dipende esplicitamente anche da t , allora la (1.9) si generalizza in

$$\dot{f} = \{f, H\} + \frac{\partial f}{\partial t}.$$

Per ogni fissata g l'operatore

$$\mathcal{L}_g = \{ \cdot, g \}$$

è un operatore di derivazione i cui coefficienti sono le derivate parziali di g :

$$\mathcal{L}_g = -\frac{\partial g}{\partial q} \cdot \frac{\partial}{\partial p} + \frac{\partial g}{\partial p} \cdot \frac{\partial}{\partial q};$$

\mathcal{L}_g altro non è che la derivata di Lie relativa al campo vettoriale hamiltoniano $X_g = (-\frac{\partial g}{\partial q}, \frac{\partial g}{\partial p})$ associato alla hamiltoniana g . Prese due funzioni f e g , il prodotto $\mathcal{L}_f \mathcal{L}_g$ non è un operatore di derivazione, perché contiene le derivate seconde. Se però consideriamo il commutatore $[\mathcal{L}_f, \mathcal{L}_g] = \mathcal{L}_f \mathcal{L}_g - \mathcal{L}_g \mathcal{L}_f$, i termini di derivata seconda si elidono e si ha ancora un operatore di derivazione: precisamente risulta

$$[\mathcal{L}_f, \mathcal{L}_g] = -\mathcal{L}_{\{f, g\}}; \quad (1.10)$$

questa espressione si vede subito essere l'identità di Jacobi, trascritta in questa notazione.

Come si vede dalla (1.9), se f ha parentesi di Poisson nulla con l'hamiltoniana allora è una costante del moto (e viceversa). Se f e g sono costanti del moto, anche $\{f, g\}$ lo è (identità di Jacobi). La (1.10) mostra che se f e g hanno parentesi di Poisson nulla, le corrispondenti derivate di Lie commutano; si dice anche, un po' impropriamente, che f e g commutano (la stessa parentesi

⁴L'identità di Jacobi è soddisfatta da due altre comuni operazioni binarie: il prodotto vettoriale $u \times v$ in \mathbb{R}^3 e il commutatore di matrici (o di operatori) $[A, B] = AB - BA$.

di Poisson è chiamata talvolta commutatore). In questo caso — ma non è banale, si veda l'appendice C — anche i flussi hamiltoniani di f e g , pensate come funzioni di Hamilton, commutano:

$$\{f, g\} = 0 \iff \Phi_f^s \circ \Phi_g^t = \Phi_g^t \circ \Phi_f^s \quad \forall s, t \in \mathbb{R} .$$

Qualche esempio di parentesi di Poisson:

- i) Le “parentesi di Poisson elementari”: denotando (un po' impropriamente) con p_i, q_i le funzioni coordinate (cioè $p_i(p, q) = p_i, q_i(p, q) = q_i$), per ogni i, j si ha

$$\{p_i, p_j\} = 0, \quad \{q_i, q_j\} = 0, \quad \{q_i, p_j\} = \delta_{ij} .$$

- ii) Per un punto materiale, denotando con x, y, z le coordinate, con $p = (p_x, p_y, p_z)$ i momenti lineari, e con $M = (M_x, M_y, M_z)$ il momento angolare, si ha

$$\{M_x, M_y\} = M_z ,$$

assieme alle analoghe che si ottengono ciclando gli indici, e inoltre

$$\{M_x, M^2\} = 0, \quad \{p_x, M_y\} = p_z, \quad \{p_x, M_z\} = -p_y, \quad \{p_x, M_x\} = 0$$

e analoghe.

ESERCIZIO 1 *Si verifichi l'affermazione fatta sopra che se $g(p, q) = G(f(p, q))$, con $G : \mathbb{R} \rightarrow \mathbb{R}$, allora $\{f, g\} = 0$. Più in generale: se $g(p, q) = G(p, q, f(p, q))$, con $G : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}$, allora nel calcolo di $\{f, g\}$ si può ignorare la dipendenza di g da p e q attraverso f .*

ESERCIZIO 2 *E' facile trovare sistemi di punti materiali per i quali si conservano due componenti della quantità di moto, ma non la terza (ad esempio, un sistema con solo forze interne e la gravità). E' possibile trovare un sistema hamiltoniano in cui si conservano due componenti del momento angolare, senza che si conservi anche la terza?*

2 Le trasformazioni canoniche

2.1 Nozione ed esempi

Uno degli aspetti più interessanti del formalismo lagrangiano è la sua invarianza per trasformazioni di coordinate: preso un qualunque sistema lagrangiano di lagrangiana $L(q, \dot{q}, t)$, se introduciamo la trasformazione di coordinate (eventualmente dipendente dal tempo) $q = q(\tilde{q}, t)$, in dettaglio

$$q_i = q_i(\tilde{q}_1, \dots, \tilde{q}_n, t), \quad i = 1, \dots, n,$$

e la estendiamo alle velocità ponendo

$$\dot{q}_i = \sum_{j=1}^n \frac{\partial q_i}{\partial \tilde{q}_j} \dot{\tilde{q}}_j + \frac{\partial q_i}{\partial t},$$

le equazioni del moto nelle nuove coordinate $\tilde{q}_1, \dots, \tilde{q}_n$ hanno ancora la forma di equazioni di Lagrange, e la nuova lagrangiana \tilde{L} si ottiene da L per semplice sostituzione di variabili:

$$\tilde{L}(\tilde{q}, \dot{\tilde{q}}, t) = L(q(\tilde{q}, t), \dot{q}(\tilde{q}, \dot{\tilde{q}}, t), t) .$$

Vogliamo qui studiare il problema analogo per il formalismo hamiltoniano. Come avremo modo di vedere, il formalismo hamiltoniano da questo punto di vista è sensibilmente più interessante e più ricco del formalismo lagrangiano, perché la classe di trasformazioni di coordinate ammissibili in ambito hamiltoniano, chiamate *trasformazioni canoniche*, è sostanzialmente più vasta. La novità *essenziale* è che oltre alle trasformazioni in cui (similmente all'ambito lagrangiano) q_1, \dots, q_n dipendono dalle sole $\tilde{q}_1, \dots, \tilde{q}_n$ e da t , e i momenti si trasformano di conseguenza (trasformazioni “puntuali” estese naturalmente ai momenti) si prendono in considerazione trasformazioni di coordinate più generali, del tipo

$$p_i = u_i(\tilde{p}, \tilde{q}, t), \quad q_i = v_i(\tilde{p}, \tilde{q}, t), \quad i = 1, \dots, n,$$

o in notazione più agile

$$p = u(\tilde{p}, \tilde{q}, t), \quad q = v(\tilde{p}, \tilde{q}, t),$$

in cui coordinate e momenti sono mescolati insieme dalla trasformazione. Useremo di frequente per la trasformazione di coordinate la notazione più compatta

$$(p, q) = w(\tilde{p}, \tilde{q}, t), \quad w = (u, v). \quad (2.1)$$

Queste trasformazioni di coordinate giocano un ruolo essenziale, in particolare, nella cosiddetta “teoria hamiltoniana delle perturbazioni”, uno dei rami ancor oggi più fecondi della meccanica, ricco di applicazioni fisiche ad esempio alla Meccanica Celeste (e alla fisica degli acceleratori, alla dinamica dei satelliti artificiali, alla Meccanica Statistica...) nel quale continuano a prodursi sempre nuovi rilevanti risultati.

Tuttavia, non tutte le trasformazioni generali del tipo (2.1) conservano la forma delle equazioni di Hamilton, ovvero sono, come si suol dire, “canoniche”. Quali siano, che proprietà abbiano e come si generino tali trasformazioni, è oggetto della *teoria delle trasformazioni canoniche*, che andiamo a esporre.

Ci restringeremo, in un primo momento, alle trasformazioni indipendenti dal tempo,

$$(p, q) = w(\tilde{p}, \tilde{q}), \quad w : \tilde{U} \rightarrow U, \quad (2.2)$$

ove \tilde{U}, U sono aperti di \mathbb{R}^{2n} , che per evitare fastidi supporremo semplicemente connessi; w sarà sempre supposta regolare invertibile (con inversa $w^{-1} : U \rightarrow \tilde{U}$ regolare).⁵ Le trasformazioni dipendenti dal tempo sono certamente importanti (sarebbe spiacevole non saper come passare, in ambito hamiltoniano, a un sistema di riferimento rotante!), ma non altrettanto importanti, e comunque conviene occuparsene in un secondo momento. Quanto al passaggio da una nozione locale, in opportuni aperti di \mathbb{R}^n , a una nozione geometrica globale (su varietà), è una questione importante ma delicata, di cui in questa trattazione elementare non ci occuperemo. Un'introduzione alla visuale geometrica della teoria delle trasformazioni canoniche si trova nell'appendice A.

A. Trasformazioni canoniche “in senso stretto”.

Diamo ora una prima definizione, in un certo senso stretta, di trasformazione canonica.

⁵Si è qui scelto di chiamare trasformazione diretta w quella che esprime le vecchie coordinate in funzione delle nuove, e trasformazione inversa w^{-1} quella che viceversa esprime le nuove coordinate in funzione delle vecchie. La scelta, che può sembrare innaturale, si rivela pratica (semplifica la notazione) al momento di usare il cambio di variabili per trasformare funzioni: a seguito del cambiamento di variabili (2.2) una qualunque $f : U \rightarrow \mathbb{R}$ è mutata in $\tilde{f} : \tilde{U} \rightarrow \mathbb{R}$ definita da $\tilde{f} = f \circ w$, ovvero $\tilde{f}(\tilde{p}, \tilde{q}) = f(w(\tilde{p}, \tilde{q}))$. In linea di principio ovviamente le due possibilità sono perfettamente equivalenti.

DEFINIZIONE 1 *La trasformazione di coordinate indipendente dal tempo (2.2) si dice canonica in senso stretto, se comunque si prenda l'hamiltoniana $H : U \rightarrow \mathbb{R}$ le equazioni di Hamilton*

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = \frac{\partial H}{\partial p}$$

sono equivalenti alle nuove equazioni

$$\dot{\tilde{p}} = -\frac{\partial \tilde{H}}{\partial \tilde{q}}, \quad \dot{\tilde{q}} = \frac{\partial \tilde{H}}{\partial \tilde{p}},$$

ove \tilde{H} denota la funzione trasformata $\tilde{H} = H \circ w$, ovvero $\tilde{H}(\tilde{p}, \tilde{q}) = H(w(\tilde{p}, \tilde{q}))$.

Se ciò accade, allora i moti nei due sistemi di coordinate sono l'uno l'immagine dell'altro:

$$\Phi_H^t \circ w = w \circ \Phi_{\tilde{H}}^t,$$

avendo denotato con $\Phi_H^t : U \rightarrow U$ e $\Phi_{\tilde{H}}^t : \tilde{U} \rightarrow \tilde{U}$ il flusso hamiltoniano nei due sistemi di coordinate. Si osservi la stretta somiglianza con quanto avviene in ambito lagrangiano: invece di trasformare le equazioni differenziali del moto, equivalentemente si trasforma la funzione H , e dalla nuova funzione $\tilde{H} = H \circ w$ si deducono le equazioni del moto nelle nuove variabili. Le due hamiltoniane H e \tilde{H} , e le corrispondenti equazioni di Hamilton, si dicono *canonicamente coniugate* da w .

Diamo qui alcuni esempi molto elementari di trasformazioni di coordinate, la cui stretta canonicità si verifica senza difficoltà sulla base della definizione; la verifica è lasciata come facile *ma importante* esercizio (trattiamo esplicitamente uno degli esempi, perché sia chiaro in che cosa consista concettualmente la verifica).

- (i) La traslazione nello spazio delle fasi:

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

(si osservi che la traslazione nelle p non ha analogo in ambito lagrangiano).

- (ii) La trasformazione lineare

$$p = A^{-T} \tilde{p}, \quad q = A \tilde{q},$$

ove A è una qualunque matrice $n \times n$ invertibile e A^{-T} denota la matrice inversa trasposta, $A^{-T} = (A^T)^{-1} = (A^{-1})^T$.

Verifica: prendiamo $H(p, q)$ qualsiasi, e poniamo $\tilde{H}(\tilde{p}, \tilde{q}) = H(A^{-T} \tilde{p}, A \tilde{q})$.

- Per inversione delle equazioni della trasformazione troviamo per una via le nuove equazioni del moto: precisamente si ha $\tilde{p} = A^T p$, $\tilde{q} = A^{-1} q$, e pertanto

$$\dot{\tilde{p}} = -A^T \frac{\partial H}{\partial q}, \quad \dot{\tilde{q}} = A^{-1} \frac{\partial H}{\partial p}. \quad (2.3)$$

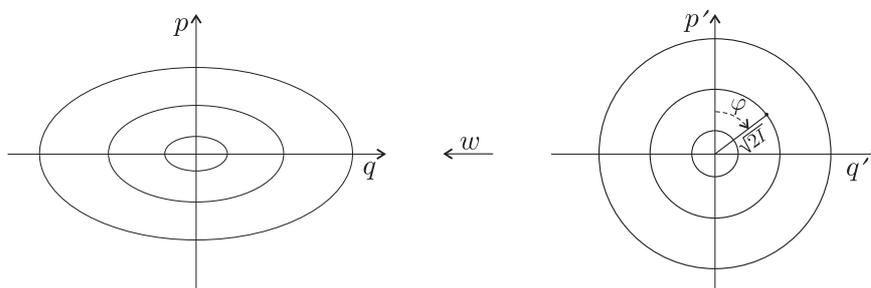


Figura 1: La trasformazione canonica da (p, q) a (p', q') e poi a (I, φ) .

- Scriviamo poi le equazioni di Hamilton relative a \tilde{H} , e verifichiamo che ritroviamo le stesse equazioni. Per \tilde{p} si ha

$$\dot{\tilde{p}}_i = -\frac{\partial \tilde{H}}{\partial \tilde{q}_i} = -\sum_{j=1}^n \frac{\partial H}{\partial q_j} A_{ji} ,$$

e segue subito la prima delle (2.3); in modo analogo si procede per l'altra.

Come caso particolare, corrispondente a A diagonale, abbiamo

- (iii) il riscalamento

$$p_i = \alpha_i \tilde{p}_i , \quad q_i = \alpha_i^{-1} \tilde{q}_i .$$

Questa elementare trasformazione consente di modificare in modo coerente le unità di misura di coordinate e momenti. Ad esempio l'hamiltoniana dell'oscillatore armonico $H(p, q) = \frac{1}{2m}p^2 + \frac{m\omega^2}{2}q^2$ viene mutata da $p = m^{1/2}\tilde{p}$, $q = m^{-1/2}\tilde{q}$ in

$$\tilde{H}(\tilde{p}, \tilde{q}) = \frac{1}{2}(\tilde{p}^2 + \omega^2 \tilde{q}^2)$$

(si porta così la massa a uno, eliminando un parametro); l'ulteriore riscalamento $\tilde{p} = \omega^{1/2}p'$, $\tilde{q} = \omega^{-1/2}q'$ muta poi l'hamiltoniana in

$$H'(p', q') = \frac{\omega}{2}(p'^2 + q'^2) ,$$

con equazioni del moto più simmetriche $\dot{p}' = -\omega q'$, $\dot{q}' = \omega p'$. Si osservi che ora p' e q' hanno la stessa dimensione fisica (nel caso tipico in cui H è un'energia e t è un tempo, sono radici di un'azione) e le curve di livello dell'hamiltoniana da ellissi sono diventate cerchi. La trasformazione complessiva di coordinate è illustrata in figura 1.

- (iv) La trasformazione che per uno o più gradi di libertà scambia le coordinate con i momenti, precisamente

$$p_i = \tilde{q}_i , \quad q_i = -\tilde{p}_i .$$

Questa trasformazione, di modesta utilità concreta, mostra tuttavia come, in ambito hamiltoniano, non vi sia alcuna differenza di principio tra coordinate e momenti coniugati.

Non è difficile vedere che non tutte le trasformazioni di coordinate della forma (2.2) sono canoniche. Consideriamo ad esempio la trasformazione che introduce, nel piano cartesiano p, q , le coordinate polari $(\tilde{p}, \tilde{q}) \equiv (r, \varphi)$, cioè

$$p = r \cos \varphi, \quad q = r \sin \varphi; \quad (2.4)$$

r e φ sono coordinate in qualche modo promettenti per l'oscillatore armonico di Hamiltoniana H' , che rimuovendo gli apici riscriviamo

$$H(p, q) = \frac{\omega}{2}(p^2 + q^2). \quad (2.5)$$

Infatti dalle equazioni del moto $\dot{p} = -\omega q$, $\dot{q} = \omega p$ si deducono subito per r e φ le equazioni semplicissime

$$\dot{r} = 0, \quad \dot{\varphi} = \omega.$$

Ma queste *non* sono le equazioni di Hamilton di $\tilde{H}(r, \varphi) = H(r \cos \varphi, r \sin \varphi) = \frac{\omega}{2}r^2$, come vorremmo: da \tilde{H} segue infatti l'equazione sbagliata (anche dimensionalmente, lo sbaglio è grosso) $\dot{\varphi} = \omega r$. Ciò esclude che la trasformazione (2.4) sia canonica.

Una trasformazione concettualmente simile alla (2.4), ma canonica, si può però costruire: si verifica infatti che

(v) la trasformazione

$$p = \sqrt{2I} \cos \varphi, \quad q = \sqrt{2I} \sin \varphi, \quad (2.6)$$

ove si sono denotati con I e φ rispettivamente il nuovo momento e la nuova coordinata (*variabili di azione angolo* dell'oscillatore; la notazione è tradizionale) è canonica.

La verifica della canonicità sulla base della definizione è un po' noiosa, ma vale la pena di eseguirla per esercizio (più avanti potremo avvalerci di facili *criteri di canonicità*, e la verifica della canonicità di questa trasformazione diventerà immediata).

Verifica: prendiamo H qualsiasi. Le inverse delle (2.6),

$$I = \frac{1}{2}(p^2 + q^2), \quad \varphi = \arctan \frac{q}{p} \quad (2.7)$$

(si sottintende $\varphi = \pi + \arctan \frac{q}{p}$, se $p < 0$) forniscono le equazioni del moto nelle nuove variabili, che si trovano essere

$$\begin{aligned} \dot{I} &= p\dot{p} + q\dot{q} = -\sqrt{2I} \cos \varphi \frac{\partial H}{\partial q} + \sqrt{2I} \sin \varphi \frac{\partial H}{\partial p} \\ \dot{\varphi} &= \frac{p\dot{q} - q\dot{p}}{p^2 + q^2} = \frac{\cos \varphi}{\sqrt{2I}} \frac{\partial H}{\partial p} + \frac{\sin \varphi}{\sqrt{2I}} \frac{\partial H}{\partial q}. \end{aligned}$$

Si verifica ora con facilità che posto $\tilde{H}(I, \varphi) = H(\sqrt{2I} \cos \varphi, \sqrt{2I} \sin \varphi)$ risulta effettivamente, come richiesto, $\dot{I} = -\frac{\partial \tilde{H}}{\partial \varphi}$ e $\dot{\varphi} = \frac{\partial \tilde{H}}{\partial I}$.

Si osservi che l'hamiltoniana (2.5) introdotta sopra per l'oscillatore armonico è mutata da questa trasformazione nell'hamiltoniana

$$\tilde{H}(I, \varphi) = \omega I \quad (2.8)$$

i cui moti sono lineari: precisamente si ha $\dot{I} = 0$, $\dot{\varphi} = \omega$, e dunque

$$I(t) = I(0), \quad \varphi(t) = \varphi(0) + \omega t.$$

Questa hamiltoniana può servire come base per lo studio del comportamento *ancora oggi non compreso a sufficienza* di un sistema di oscillatori armonici debolmente accoppiati.

- Che l'uso di $I = \frac{1}{2}r^2$ al posto di r come coordinata radiale “aggiusti” per l'oscillatore armonico la difficoltà delle (2.4), conducendo per tale sistema all'hamiltoniana lineare (2.8) e dunque all'equazione giusta $\dot{\varphi} = \omega$ per φ , è proprio evidente, e questo ha il valore di suggerimento su come correggere le (2.4) arrivando alle (2.6). Naturalmente la dimostrazione che la trasformazione (2.6) è canonica rimane necessaria (deve funzionare per ogni H , non certo solo per l'oscillatore armonico).
- La (2.4) non è a posto neanche dimensionalmente: la variabile coniugata a un angolo deve essere per forza un'azione, mentre $r = \sqrt{p^2 + q^2}$ è, come p e q separatamente, la radice di un'azione. Anche questa osservazione elementare fa capire che il modo giusto di correggere la (2.4) è la (2.6), a parte il fattore 2 che comunque non è difficile mettere a posto.

ESERCIZIO 3 *Si verifichi che anche l'hamiltoniana della particella libera $H(p, q) = \frac{1}{2}p^2$ costituisce un controesempio alla canonicità delle (2.4). Per questo si scrivano le equazioni della particella libera nelle coordinate r, φ , e si osservi che esse non sono le equazioni del moto di $\tilde{H}(r, \varphi) = \frac{1}{2}r^2 \cos^2 \varphi$. Anzi: si vede facilmente che queste equazioni non sono le equazioni di Hamilton di nessuna altra hamiltoniana $K(r, \varphi)$, anche diversa da $\tilde{H} = H \circ w$. [Suggerimento: se lo fossero, allora le derivate miste $\frac{\partial^2 K}{\partial r \partial \varphi}$ e $\frac{\partial^2 K}{\partial \varphi \partial r} \dots$]*

B. Una condizione necessaria e sufficiente per la canonicità.

Del tutto in generale, consideriamo un sistema di equazioni differenziali (non necessariamente hamiltoniano)

$$\dot{x} = X(x) ,$$

e introduciamo il cambiamento di variabili $x = w(\tilde{x})$. Derivando rispetto al tempo l'equazione della trasformazione si ha $\dot{x} = J(\tilde{x})\dot{\tilde{x}}$, ove J denota la matrice jacobiana

$$J = \begin{pmatrix} \frac{\partial w_i}{\partial \tilde{x}_j} \end{pmatrix} .$$

L'equazione del moto nelle nuove variabili è allora

$$\dot{\tilde{x}} = \tilde{X}(\tilde{x}) ,$$

con X e \tilde{X} legati da

$$X(w(\tilde{x})) = J(\tilde{x})\tilde{X}(\tilde{x}) , \quad \text{ovvero} \quad X \circ w = J\tilde{X} .$$

Nel caso hamiltoniano il campo vettoriale è definito dalla (1.2) mentre J è la matrice a blocchi

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

Dire che w è canonica, nel senso stretto della definizione 1, vuol dire che per ogni H il campo trasformato \tilde{X}_H di X_H coincide con il campo $X_{\tilde{H}}$ dell'hamiltoniana trasformata \tilde{H} , e dunque la condizione di canonicità si scrive

$$X_H \circ w = JX_{\tilde{H}} . \tag{2.9}$$

Vale allora la semplice ma cruciale

PROPOSIZIONE 1 *Condizione necessaria e sufficiente perché la trasformazione di coordinate $x = w(\tilde{x})$ sia canonica in senso stretto è che la matrice jacobiana J della trasformazione soddisfi ovunque la relazione algebrica*

$$J \mathbb{E} J^T = \mathbb{E} . \quad (2.10)$$

DIMOSTRAZIONE. La (2.9), usando la relazione (1.2) tra hamiltoniana e campo vettoriale, si riscrive

$$\mathbb{E} \partial_x H \circ w = J \mathbb{E} \partial_{\tilde{x}} \tilde{H} .$$

Poiché risulta, come è immediato verificare a partire da $\tilde{H}(\tilde{x}) = H(w(\tilde{x}))$,

$$\partial_{\tilde{x}} \tilde{H} = J^T \partial_x H \circ w , \quad (2.11)$$

la condizione di stretta canonicità si riscrive ancora

$$\mathbb{E} \partial_x H \circ w = J \mathbb{E} J^T \partial_x H \circ w ,$$

e l'uguaglianza è soddisfatta *per ogni* H se e solo se è soddisfatta la (2.10). \square

ESERCIZIO 4 *Si verifichi esplicitamente che per tutti gli esempi (i)–(v) sopra introdotti la relazione algebrica (2.10) è soddisfatta.*

Matrici che soddisfano la relazione algebrica (2.10) si dicono *simplettiche*. È immediato verificare che esse formano un gruppo: infatti, la matrice identità è evidentemente simplettica; J^{-1} è simplettica, come segue da

$$J^{-1} \mathbb{E} (J^{-1})^T = J^{-1} (J \mathbb{E} J^T) (J^{-1})^T = \mathbb{E} ;$$

infine, se J e K sono simplettiche, allora JK è simplettica, come segue da

$$(JK) \mathbb{E} (JK)^T = JK \mathbb{E} K^T J^T = J \mathbb{E} J^T = \mathbb{E} .$$

Il gruppo è detto *gruppo simplettico*. La matrice \mathbb{E} è essa stessa simplettica, inoltre se J è simplettica lo è anche J^T : infatti, da $J^{-1} \mathbb{E} (J^{-1})^T = \mathbb{E}$ segue, per inversione, $J^T \mathbb{E}^{-1} J = \mathbb{E}^{-1}$, e dunque (ricordando $\mathbb{E}^{-1} = -\mathbb{E}$) $J^T \mathbb{E} J = \mathbb{E}$. Sulla base della proposizione 1, le trasformazioni canoniche in senso stretto sono dette anche *trasformazioni simplettiche*.

Dalla (2.10) si deduce immediatamente $(\det J)^2 = 1$, e dunque $\det J = \pm 1$; in realtà si può dimostrare, ma non è immediato, che si ha sempre $\det J = 1$. Pertanto

PROPOSIZIONE 2 *Le trasformazioni strettamente canoniche preservano il volume euclideo (orientato) dello spazio delle fasi.*

Si verifica poi immediatamente (esercizio) che

PROPOSIZIONE 3 *Per le matrici 2×2 la proprietà di simpletticità coincide con la proprietà di avere determinante uguale a uno.*

Pertanto per $n = 1$ le trasformazioni strettamente canoniche sono tutte e sole quelle che conservano l'area (orientata) nel piano delle fasi pq . La proprietà *non* si estende a $n > 1$.

C. Trasformazioni canoniche e parentesi di Poisson.

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

DEFINIZIONE 2 *Si dice che la trasformazione di coordinate*

$$(p, q) = w(\tilde{p}, \tilde{q})$$

preserva le parentesi di Poisson se comunque si prendano due funzioni f e g , indicando con $\tilde{f} = f \circ w$, $\tilde{g} = g \circ w$ le funzioni trasformate, risulta

$$\{\tilde{f}, \tilde{g}\} = \{f, g\} \circ w .$$

Si intende (ma non potrebbe essere altrimenti) che nella parentesi di Poisson a sinistra dell'uguale si deriva rispetto a \tilde{p} e \tilde{q} , a destra invece rispetto a p, q . Sottintendendo il cambiamento di variabili, si usa a volte la scrittura breve $\{\tilde{f}, \tilde{g}\} = \{f, g\}$.

Se w preserva le parentesi di Poisson, preserva in particolare le parentesi di Poisson "fondamentali", precisamente si ha, per ogni i, j ,

$$\{u_i, u_j\} = 0, \quad \{v_i, v_j\} = 0, \quad \{v_i, u_j\} = \delta_{ij} .$$

Si vede ora facilmente che *la preservazione delle parentesi di Poisson fondamentali implica la preservazione di tutte le parentesi di Poisson*: un semplice calcolo mostra infatti che

$$\begin{aligned} \{\tilde{f}, \tilde{g}\} &= \sum_{i,j=1}^n \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial p_j} \{u_i, u_j\} + \sum_{i,j=1}^n \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_j} \{u_i, v_j\} \\ &+ \sum_{i,j=1}^n \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_j} \{v_i, u_j\} + \sum_{i,j=1}^n \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial q_j} \{v_i, v_j\}, \end{aligned}$$

e la conclusione è immediata. La verifica di canonicità attraverso le parentesi di Poisson è così finita, nonostante la definizione faccia riferimento a infinite funzioni.

ESERCIZIO 5 *Si verifichi che le trasformazioni (i)-(v) sopra introdotte a titolo di esempio preservano le parentesi di Poisson.*

Si dimostra senza difficoltà la seguente

PROPOSIZIONE 4 *La trasformazione $(p, q) = w(\tilde{p}, \tilde{q})$ preserva le parentesi di Poisson se e solo se è strettamente canonica.*

DIMOSTRAZIONE. La conservazione delle parentesi di Poisson si scrive, con la notazione compatta sopra introdotta,

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

o anche, osservando che $\partial_{\tilde{x}}\tilde{f} = J^T \partial_x f \circ w$, $\partial_{\tilde{x}}\tilde{g} = J^T \partial_x g \circ w$,

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

eliminando l'ormai inessenziale composizione con w , si vede allora che questa uguaglianza è soddisfatta per ogni coppia f e g , se e solo se si ha $J \mathbb{E} J^T = \mathbb{E}$. \square

- La definizione (2.10) di matrice simplettica presenta una certa analogia con la nozione di matrice ortogonale, cui ci si riporta sostituendo all'identità simplettica \mathbb{E} la consueta identità $2n \times 2n$. Come le matrici ortogonali preservano il prodotto scalare euclideo $a \cdot b$ in \mathbb{R}^n , così le matrici simplettiche preservano il cosiddetto *prodotto scalare simplettico* $[a, b] = a \cdot \mathbb{E} b$ (che però non è un prodotto scalare, essendo antisimmetrico). L'ortogonalità tuttavia in un certo senso è una proprietà più stretta: si verifica facilmente infatti che le matrici ortogonali $m \times m$ dipendono da $m(m-1)/2$ parametri reali indipendenti, mentre le matrici simplettiche ($m = 2n$) dipendono da $m(m+1)/2$ parametri indipendenti.⁶ Per $m = 2$ abbiamo da un lato le rotazioni piane (un solo parametro), dall'altro tutte le matrici con determinante 1 (tre parametri).

D. Trasformazioni canoniche e 1-forma di Liouville.

Vi è un'ulteriore importante caratterizzazione delle trasformazioni canoniche, basata sul comportamento, sotto l'effetto della trasformazione, di alcune forme differenziali. La più semplice, di fatto l'unica che prenderemo in considerazione, riguarda la cosiddetta *1-forma di Liouville*

$$p \cdot dq = p_1 dq_1 + \cdots + p_n dq_n . \quad (2.12)$$

DEFINIZIONE 3 Diremo che la trasformazione di coordinate $p = u(\tilde{p}, \tilde{q})$, $q = v(\tilde{p}, \tilde{q})$, preserva debolmente la 1-forma di Liouville, se esiste localmente $f(\tilde{p}, \tilde{q})$ tale che risulti soddisfatta l'identità differenziale

$$u \cdot dv = \tilde{p} \cdot d\tilde{q} + df . \quad (2.13)$$

In altre parole, la forma differenziale $u \cdot dv - \tilde{p} \cdot d\tilde{q}$ è localmente esatta, o chiusa. Con un piccolo abuso di notazione si scrive spesso la condizione (2.13) nella forma più simmetrica e facile da ricordare

$$p \cdot dq = \tilde{p} \cdot d\tilde{q} + df ;$$

da questa scrittura non è immediatamente evidente quali siano le variabili indipendenti (le \tilde{p}, \tilde{q} o le p, q) cui si sta facendo riferimento, ma la cosa non è grave perché la scelta delle variabili indipendenti nelle quali si verifica un'identità differenziale come la (2.13) è irrilevante.

ESERCIZIO 6 Si verifichi che le trasformazioni di coordinate (i)–(v) introdotte sopra al punto B preservano debolmente la 1-forma di Liouville, e per ciascuna si trovi la funzione f . [Risposta: per la (i) si trova $f = a \cdot \tilde{q}$; per la (ii) e (iii) $f = 0$; per la (v) $f = I \sin \varphi \cos \varphi$.]

⁶Una matrice $m \times m$ ha m^2 elementi. La relazione di ortogonalità $JJ^T = \mathbb{I}$ comporta $m(m+1)/2$ equazioni indipendenti (una per ogni elemento indipendente di JJ^T , che è simmetrica), e dunque restano $m^2 - m(m+1)/2 = m(m-1)/2$ parametri indipendenti. Invece $J\mathbb{E}J^T$ è antisimmetrica, pertanto la relazione di simpletticità (2.10) si traduce in sole $m(m-1)/2$ equazioni indipendenti, e il numero di parametri indipendenti che rimangono è $m^2 - m(m-1)/2 = m(m+1)/2$.

La (2.12) è l'espressione più tipica della 1-forma di Liouville, adattata ai sistemi hamiltoniani che provengono da quelli lagrangiani, nei quali le q e le p hanno un ruolo diverso. Ma in meccanica hamiltoniana tutto è simmetrico, e ad esempio la condizione (2.13) si può riscrivere nella forma del tutto equivalente

$$v \cdot du = \tilde{q} \cdot d\tilde{p} + df' ,$$

con $f' = -f + u \cdot v - \tilde{p} \cdot \tilde{q}$ (f esiste se e solo se esiste f'); sottraendo le due si trova la condizione più simmetrica, ancora equivalente alla (2.13),

$$u \cdot dv - v \cdot du = \tilde{p} \cdot d\tilde{q} - \tilde{q} \cdot d\tilde{p} + dg , \quad g = 2f - u \cdot v + \tilde{p} \cdot \tilde{q}$$

(f esiste se e solo se esiste g), interessante perché si presta alla notazione compatta

$$\mathbb{E} w \cdot dw - \mathbb{E} \tilde{x} \cdot d\tilde{x} = dg . \quad (2.14)$$

Si dimostra ora che

PROPOSIZIONE 5 *La trasformazione $p = u(\tilde{p}, \tilde{q})$, $q = v(\tilde{p}, \tilde{q})$ preserva debolmente la 1-forma di Liouville, se e solo se è strettamente canonica.*

DIMOSTRAZIONE. Facciamo riferimento alla (2.14). Ricordando che $dw = Jd\tilde{x}$, la condizione si scrive

$$(J^T \mathbb{E} w - \mathbb{E} \tilde{x}) \cdot d\tilde{x} = dg ;$$

di conseguenza g esiste se e solo se il campo vettoriale

$$X = J^T(\tilde{x}) \mathbb{E} w(\tilde{x}) - \mathbb{E} \tilde{x}$$

è tale che

$$X_i = \frac{\partial g}{\partial \tilde{x}_i} \quad \forall i ,$$

e dunque se e solo se

$$\frac{\partial X_i}{\partial \tilde{x}_j} - \frac{\partial X_j}{\partial \tilde{x}_i} = 0 \quad \forall i, j .$$

Ora non è difficile verificare che si ha

$$\frac{\partial X_i}{\partial \tilde{x}_j} - \frac{\partial X_j}{\partial \tilde{x}_i} = 2(J^T \mathbb{E} J - \mathbb{E})_{ij} ;$$

la deduzione è un po' complicata per il fatto che anche J^T dipende da \tilde{x} , ma si vede facilmente che i termini contenenti le derivate di J^T (che sono le derivate seconde di w) si elidono e non contribuiscono. Segue immediatamente che g esiste, e la (2.14) è soddisfatta, se e solo se la matrice J è simplettica. \square

- La 1-forma di Liouville non è esatta e corrispondentemente il suo integrale lungo un cammino chiuso non è nullo. Per $n = 1$ l'integrale della 1-forma di Liouville su un cammino chiuso nel piano pq è, a meno del segno, l'area racchiusa dal cammino, si veda la figura 2; per $n > 1$ gradi di libertà l'integrale è la somma delle aree delle proiezioni del cammino sugli n piani coordinati $p_i q_i$. Se si aggiunge a una 1-forma il differenziale df di una funzione, la 1-forma cambia, ma

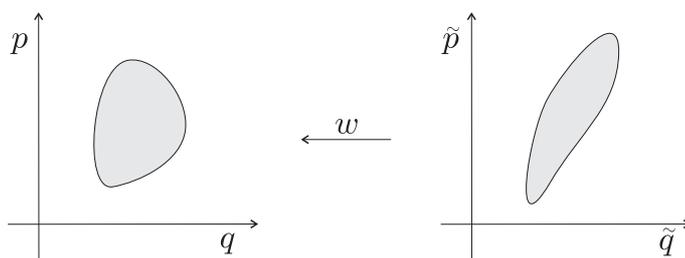


Figura 2: Ad illustrazione della condizione (2.13).

il suo integrale su un ciclo (supponendo qui la semplice connessione del dominio) non cambia. Se una trasformazione di coordinate $(p, q) = w(\tilde{p}, \tilde{q})$ soddisfa la (2.13), e si considerano due cicli che siano uno l'immagine dell'altro nelle coordinate p, q e \tilde{p}, \tilde{q} , si ottengono ($n = 1$) cicli di forma diversa, ma della stessa area; si veda la figura 2. La preservazione debole della 1-forma di Liouville, grazie alla proposizione 5, fornisce così una caratterizzazione geometrica della canonicità di una trasformazione, che affianca la caratterizzazione algebrica basata sulla simpletticità della matrice jacobiana.

ESERCIZIO 7 Si consideri (per n qualsiasi) la semplice trasformazione di coordinate $p = \tilde{p} + a(\tilde{q})$, $q = \tilde{q}$; di trovi la condizione su a perché la trasformazione sia canonica. Si faccia lo stesso per $p = \tilde{p}$, $q = \tilde{q} + b(\tilde{p})$.

Alcuni degli esempi dati sopra, in particolare la trasformazione (2.6), mostrano come l'insieme delle trasformazioni canoniche contenga cambiamenti di variabili che non hanno analogo in ambito lagrangiano. Viceversa,

PROPOSIZIONE 6 Data una qualunque trasformazione di coordinate puntuale $q = v(\tilde{q})$ la si può sempre estendere canonicamente ai momenti ponendo

$$p = A^{-T}(\tilde{q}) \tilde{p}, \quad (2.15)$$

ove A denota la matrice jacobiana di v , e la (2.13) è soddisfatta con $f = 0$.

DIMOSTRAZIONE. Si ha $dv = A(\tilde{q})d\tilde{q}$, pertanto $u \cdot dv = A^{-T}\tilde{p} \cdot A d\tilde{q} = \tilde{p} \cdot d\tilde{q}$. \square

Si generalizza in questo modo l'esempio (ii), in cui la matrice A era costante. Queste trasformazioni, dette "puntuali estese" o brevemente puntuali, hanno così la proprietà di preservare esattamente, anziché debolmente, la 1-forma di Liouville. L'affermazione non si inverte: vi sono trasformazioni che preservano esattamente la 1-forma di Liouville, senza essere puntuali estese.

ESERCIZIO 8 Si studi la condizione generale perché una trasformazione canonica preservi esattamente la 1-forma di Liouville. [Risposta: $v(\tilde{p}, \tilde{q})$ deve essere omogenea di grado zero nei nuovi momenti \tilde{p}]

- E' bene osservare che la (2.15) è coerente con la definizione dei momenti $p_i = \frac{\partial L}{\partial \dot{q}_i}$ data in ambito lagrangiano. (Per verificarlo si prenda L qualsiasi, si costruisca \tilde{L} e si confrontino $p = \frac{\partial L}{\partial \dot{q}}$ e $\tilde{p} = \frac{\partial \tilde{L}}{\partial \dot{\tilde{q}}}$; si osservi che la relazione tra p e \tilde{p} è indipendente da L .)

E. Trasformazioni canoniche generali (indipendenti dal tempo).

Nella definizione 1 data sopra di trasformazione canonica in senso stretto abbiamo chiesto, in sostanza, che w muti ogni sistema hamiltoniano in un sistema hamiltoniano, e *inoltre* che la nuova hamiltoniana si ottenga dalla precedente per sostituzione di variabili. Possiamo utilmente indebolire la nozione di trasformazione canonica rinunciando a quest'ultima richiesta, e dare la seguente più generale

DEFINIZIONE 4 *La trasformazione di variabili $(p, q) = w(\tilde{p}, \tilde{q})$, $w : \tilde{U} \rightarrow U$, si dice canonica in senso generale se per ogni $H : U \rightarrow \mathbb{R}$ esiste $K : \tilde{U} \rightarrow \mathbb{R}$ tale che le equazioni di Hamilton relative a H e a K siano equivalenti.*

La nozione è effettivamente più debole, come mostra il “riscaldamento omogeneo” w definito da

$$p_i = \alpha_i \tilde{p}_i, \quad q_i = \beta_i \tilde{q}_i \quad \text{con} \quad \alpha_i \beta_i = c^{-1} \text{ costante} \quad \forall i,$$

che generalizza il precedente esempio (iii) nel quale era $c = 1$ (il riscaldamento è detto omogeneo perché la costante c è la stessa per tutti i gradi di libertà). Si vede subito infatti che le nuove equazioni del moto sono le equazioni di Hamilton relative a $K = cH \circ w$, pertanto per $c \neq 1$ la trasformazione w è canonica ma non in senso stretto.

L'aggiunta dei riscaldamenti con $c \neq 1$ allarga il gruppo simplettico al gruppo simplettico generalizzato definito da

$$cJ \mathbb{E} J^T = \mathbb{E}, \quad c \neq 0, \quad (2.16)$$

mentre la proprietà di w di preservare le parentesi di Poisson si generalizza in

$$c\{\tilde{f}, \tilde{g}\} = \{f, g\} \circ w, \quad (2.17)$$

e infine la condizione (2.13) si generalizza in

$$cp \cdot dq = \tilde{p} \cdot d\tilde{q} + df. \quad (2.18)$$

Tale allargamento è utile in diversi casi (si veda ad esempio il paragrafo 4.1), ma si capisce che non muta in modo sostanziale la nozione di trasformazione canonica; tra l'altro, come mostra la (1.5), la moltiplicazione dell'hamiltoniana per una costante si può sempre compensare con un riscaldamento temporale. *La generalizzazione tuttavia è tutta qui:* vale infatti la seguente

PROPOSIZIONE 7 *Se la trasformazione di coordinate indipendente dal tempo $(p, q) = w(\tilde{p}, \tilde{q})$ è canonica nel senso generale della definizione 4, allora esiste $c \neq 0$ tale che comunque si prenda H , la nuova hamiltoniana è (a meno di una costante additiva) $K = cH \circ w$.*

La costante c (che, si osservi, dipende solo dalla trasformazione e non da H) è detta *valenza* della trasformazione; le trasformazioni canoniche in senso stretto sono dette a volte *univalenti*. La dimostrazione della proposizione 7 è riportata nell'appendice D.

- La distinzione tra trasformazioni canoniche in senso stretto e in senso generale, come si è visto, è modesta. Libri e articoli scientifici spesso la ignorano, senza danno, usando il termine “canoniche” indifferentemente per le une e per le altre. Anche in queste note useremo spesso la sola espressione “canoniche”, senza la pedanteria di specificare ogni volta se in senso stretto o in senso generale.

2.2 Funzioni generatrici di trasformazioni canoniche

L'interesse per le trasformazioni canoniche deriva principalmente dalla possibilità di utilizzarle per introdurre coordinate che descrivano in modo più semplice, e anche più profondo, la dinamica di un sistema, restando nell'ambito del formalismo hamiltoniano. Esse sono usate in modo intenso in particolare nella teoria delle perturbazioni, ove si cerca per quanto possibile di riassorbire l'effetto di una perturbazione in un opportuno cambiamento di coordinate. Per questo però è importante saper produrre famiglie sufficientemente ampie di cambiamenti di coordinate la cui canonicità sia, per così dire, garantita a priori. Vi sono fondamentalmente due metodi adatti allo scopo: il metodo delle "funzioni generatrici" in variabili miste e il "metodo di Lie". Ci restringeremo qui, per semplicità di linguaggio, alle sole trasformazioni univalenti; l'estensione di quanto diremo al caso $c \neq 1$ è del tutto ovvia.

A. Funzioni generatrici.

Consideriamo una funzione di $2n$ variabili $F(q, \tilde{q})$ definita in un aperto di \mathbb{R}^{2n} , e supponiamo che nel dominio di definizione risulti

$$\det \left(\frac{\partial^2 F}{\partial q_i \partial \tilde{q}_j} \right) \neq 0 . \quad (2.19)$$

Allora le equazioni

$$p = \frac{\partial F}{\partial q}(q, \tilde{q}) , \quad \tilde{p} = -\frac{\partial F}{\partial \tilde{q}}(q, \tilde{q}) \quad (2.20)$$

definiscono implicitamente una trasformazione locale di coordinate $(p, q) = w(\tilde{p}, \tilde{q})$, da un qualche aperto $\tilde{U} \subset \mathbb{R}^{2n}$ alla sua immagine $U \subset \mathbb{R}^{2n}$. Infatti, la condizione (2.19) garantisce l'invertibilità locale della seconda delle (2.20) rispetto a q , fornendo così $q = v(\tilde{p}, \tilde{q})$, e poi per sostituzione nella prima si trova $p = u(\tilde{p}, \tilde{q})$; invertendo invece \tilde{q} nella prima e sostituendo nella seconda si trova la trasformazione inversa, a prova dell'invertibilità locale di w .

PROPOSIZIONE 8 *La trasformazione w così generata è canonica.*

DIMOSTRAZIONE. La condizione (2.13) è soddisfatta con $f(\tilde{p}, \tilde{q}) = F(v(\tilde{p}, \tilde{q}), \tilde{q})$. Infatti si ha

$$df = \frac{\partial F}{\partial q}(v(\tilde{p}, \tilde{q}), \tilde{q}) \cdot dv + \frac{\partial F}{\partial \tilde{q}}(v(\tilde{p}, \tilde{q}), \tilde{q}) \cdot d\tilde{q} = u \cdot dv - \tilde{p} \cdot d\tilde{q} ,$$

come richiesto (non c'è bisogno, si osservi, di sviluppare il differenziale dv). □

Questo modo di generare trasformazioni canoniche scrivendole dapprima in variabili miste e poi eseguendo un'inversione può apparire bizzarro. In realtà è molto naturale se si pensa di verificare l'identità (2.13), come è spontaneo, pensando che siano variabili indipendenti q e \tilde{q} anziché \tilde{p} e \tilde{q} , così da non dover sviluppare dv .

Le trasformazioni generate in questo modo hanno $v(\tilde{p}, \tilde{q})$ necessariamente dipendente da \tilde{p} (infatti v è ottenuta dalla seconda delle (2.20) per inversione di \tilde{p} , cosicché $\det \frac{\partial v}{\partial \tilde{p}} \neq 0$), perciò non comprendono la classe importante delle trasformazioni puntuali estese canonicamente ai momenti,

$$p = u(\tilde{p}, \tilde{q}) , \quad q = v(\tilde{q})$$

(non comprendono nemmeno l'identità!). Si rimedia con una variante banale: precisamente si riscrive la (2.13) nella forma, ovviamente equivalente,

$$p \cdot dq = -\tilde{q} \cdot d\tilde{p} + dg, \quad g = f + \tilde{p} \cdot \tilde{q}, \quad (2.21)$$

e in modo assolutamente identico si ottiene la seguente proposizione:

PROPOSIZIONE 9 *Sia data $S(\tilde{p}, q)$ tale che nel suo dominio di definizione risulti*

$$\det \left(\frac{\partial^2 S}{\partial \tilde{p}_i \partial q_j} \right) \neq 0. \quad (2.22)$$

Allora le equazioni

$$p = \frac{\partial S}{\partial q}(\tilde{p}, q), \quad \tilde{q} = \frac{\partial S}{\partial \tilde{p}}(\tilde{p}, q) \quad (2.23)$$

definiscono implicitamente una trasformazione di coordinate

$$p = u(\tilde{p}, \tilde{q}), \quad q = v(\tilde{p}, \tilde{q})$$

che laddove è definita è canonica.

L'identità è generata da $S(\tilde{p}, q) = \tilde{p} \cdot q$, e si vede subito che le trasformazioni puntuali sono generate da funzioni S del tipo

$$S(\tilde{p}, q) = \tilde{p} \cdot \tilde{v}(q).$$

ESERCIZIO 9 *Quale classe di trasformazioni canoniche resta esclusa questa volta?*

ESERCIZIO 10 *Si determini la funzione generatrice S per le trasformazioni canoniche (i), (ii) e (v) introdotte nel paragrafo 2.1. [Risposte: (i) $S(\tilde{p}, q) = (\tilde{p} + a) \cdot (q - b)$; (ii) $S(\tilde{p}, q) = \tilde{p} \cdot A^{-1}q$; (v) $S(I, q) = \int_0^q \sqrt{2I - x^2} dx$.] Si determini la generatrice S dell'inversa della (v): $I = \frac{1}{2}(p^2 + q^2)$, $\varphi = \arctan \frac{q}{p}$. [Risposta: $S(p, \varphi) = \frac{1}{2}p^2 \tan \varphi$.]*

ESERCIZIO 11 *Si introducano funzioni generatrici simili a F e a S , che però adoperino come variabili indipendenti le coppie p, \tilde{q} oppure p, \tilde{p} . Si introducano funzioni generatrici più generali, che operano diversamente, al modo di F o di S , per i diversi gradi di libertà.⁷*

Nella teoria delle perturbazioni giocano un ruolo importante le cosiddette trasformazioni canoniche prossime all'identità, precisamente trasformazioni canoniche dipendenti (in modo regolare) da un parametro ε definito in un intorno dell'origine, che si riducono all'identità per $\varepsilon = 0$:

$$(p, q) = w_\varepsilon(\tilde{p}, \tilde{q}), \quad w_0(\tilde{p}, \tilde{q}) = (\tilde{p}, \tilde{q}). \quad (2.24)$$

Un modo semplice di generare trasformazioni di questo genere è di prendere una funzione generatrice della forma

$$S(\tilde{p}, q) = \tilde{p} \cdot q + \varepsilon \mathcal{S}(\tilde{p}, q, \varepsilon);$$

la condizione di invertibilità (2.22), si osservi, è certamente soddisfatta per ε piccolo.

⁷Con quest'ultima generalizzazione è possibile dimostrare che tutte le trasformazioni canoniche locali ammettono una funzione generatrice.

PROPOSIZIONE 10 *Tutte le trasformazioni canoniche prossime all'identità si possono scrivere, per ε piccolo, in questo modo.*

La proposizione interviene in alcuni teoremi di non esistenza, nei quali bisogna essere certi di aver preso in considerazione tutte le trasformazioni a priori possibili. La dimostrazione (facile, ma di interesse secondario per noi) è riportata nell'appendice D.

B. Il flusso hamiltoniano come trasformazione canonica; il metodo di Lie

Dato un sistema di equazioni differenziali $\dot{x} = X(x)$ in \mathbb{R}^n , il suo flusso Φ_X^t fornisce per ogni t_0 fissato una trasformazione regolare e invertibile: $\mathbb{R}^n \rightarrow \mathbb{R}^n$, che associa al dato iniziale \tilde{x} il dato x all'istante $t = t_0$. Consideriamo ora il caso di un sistema hamiltoniano, denotiamo

$$\Phi_H^t(\tilde{p}, \tilde{q}) = (u(\tilde{p}, \tilde{q}, t), v(\tilde{p}, \tilde{q}, t))$$

(restano così definite u e v), e per ogni t_0 fissato consideriamo la trasformazione di variabili

$$p = u(\tilde{p}, \tilde{q}, t_0), \quad q = v(\tilde{p}, \tilde{q}, t_0). \quad (2.25)$$

PROPOSIZIONE 11 *Per ogni t_0 la trasformazione 2.25 è strettamente canonica, precisamente soddisfa la condizione (2.13) con f data da*

$$f(\tilde{p}, \tilde{q}, t_0) = \int_0^{t_0} \left[u(\tilde{p}, \tilde{q}, t) \cdot \frac{\partial v}{\partial t}(\tilde{p}, \tilde{q}, t) - H(u(\tilde{p}, \tilde{q}, t), v(\tilde{p}, \tilde{q}, t)) \right] dt.$$

Dall'espressione di f , tenendo conto del significato di u e v , si vede che la funzione integranda è $p \cdot \dot{q} - H = L$, pertanto f coincide con l'azione hamiltoniana calcolata per il moto vero, espressa in funzione del dato iniziale (\tilde{p}, \tilde{q}) e del tempo finale t_0 .

DIMOSTRAZIONE. Sviluppando i differenziali dv e df nella (2.13) si vede subito che affinché la 1-forma di Liouville sia debolmente preservata bisogna e basta che per ogni $i = 1, \dots, n$ risulti

$$\frac{\partial f}{\partial \tilde{p}_i}(\tilde{p}, \tilde{q}, t_0) = u(\tilde{p}, \tilde{q}, t_0) \cdot \frac{\partial v}{\partial \tilde{p}_i}(\tilde{p}, \tilde{q}, t_0) \quad (2.26)$$

$$\frac{\partial f}{\partial \tilde{q}_i}(\tilde{p}, \tilde{q}, t_0) = u(\tilde{p}, \tilde{q}, t_0) \cdot \frac{\partial v}{\partial \tilde{q}_i}(\tilde{p}, \tilde{q}, t_0) - \tilde{p}_i. \quad (2.27)$$

Verifichiamo la prima di esse. Dalla definizione di f segue

$$\frac{\partial f}{\partial \tilde{p}_i} = \int_0^{t_0} \left[\frac{\partial u}{\partial \tilde{p}_i} \cdot \frac{\partial v}{\partial t} + u \cdot \frac{\partial^2 v}{\partial t \partial \tilde{p}_i} - \frac{\partial H}{\partial p} \cdot \frac{\partial u}{\partial \tilde{p}_i} - \frac{\partial H}{\partial q} \cdot \frac{\partial v}{\partial \tilde{p}_i} \right] dt.$$

Ma per ipotesi u e v nelle (2.25) sono soluzioni delle equazioni di Hamilton relative all'hamiltoniana H , e dunque $\frac{\partial H}{\partial p} = \frac{\partial v}{\partial t}$, $-\frac{\partial H}{\partial q} = \frac{\partial u}{\partial t}$; segue allora

$$\frac{\partial f}{\partial \tilde{p}_i} = \int_0^{t_0} \left[u \cdot \frac{\partial^2 v}{\partial t \partial \tilde{p}_i} + \frac{\partial u}{\partial t} \cdot \frac{\partial v}{\partial \tilde{p}_i} \right] dt = \int_0^{t_0} \frac{\partial}{\partial t} \left(u \cdot \frac{\partial v}{\partial \tilde{p}_i} \right) dt = u \cdot \frac{\partial v}{\partial \tilde{p}_i} \Big|_0^{t_0}.$$

Osservando infine che $v(\tilde{p}, \tilde{q}, 0) = \tilde{q}$, cosicché $\frac{\partial v}{\partial \tilde{p}_i}(\tilde{p}, \tilde{q}, 0) = 0$, la (2.26) è immediata. In modo analogo si ottiene facilmente anche la (2.27). \square

Dalle proposizioni 11 e 2 si deduce immediatamente che

COROLLARIO 12 *Il flusso hamiltoniano preserva il volume euclideo $dp_1 \cdots dp_n dq_1 \cdots dq_n$ nello spazio delle fasi.*

- Abbiamo così ritrovato per questa via il teorema di Liouville sulla conservazione del volume nello spazio delle fasi (paragrafo 1.1). Dimostrare il teorema in questo modo tuttavia è cosa sproporzionata: il complesso apparato delle trasformazioni canoniche ha poco a che fare col risultato, che è assai più elementare e si deduce facilmente (appendice B) dal semplice fatto che il campo vettoriale hamiltoniano ha divergenza nulla.

Per $t_0 = \varepsilon$ piccolo, la trasformazione $w_\varepsilon = \Phi_H^\varepsilon$ è prossima all'identità; l'hamiltoniana H che con il suo flusso la genera è detta *hamiltoniana generatrice*. Nel seguito la denoteremo di regola con χ , per riservare il simbolo H all'hamiltoniana sulla quale vogliamo lavorare. Questo metodo per produrre trasformazioni canoniche prossime all'identità, detto *metodo di Lie*, è di largo uso in teoria delle perturbazioni. Infatti, benché la trasformazione Φ_χ^ε non sia scritta esplicitamente, è molto chiaro come viene trasformata una generica funzione F (ad esempio l'hamiltoniana H del problema allo studio) tramite la composizione $F \circ \Phi_\chi^\varepsilon$: da $\dot{F} = \mathcal{L}_\chi F$, ove $\mathcal{L}_\chi = \{., \chi\}$ denota la derivata di Lie associata al campo vettoriale di χ , segue subito lo sviluppo in serie

$$F \circ \Phi_\chi^\varepsilon = \sum_{r=0}^{\infty} \frac{\varepsilon^r}{r!} \mathcal{L}_\chi^r F = e^{\varepsilon \mathcal{L}_\chi} F . \quad (2.28)$$

Non ci addentriamo qui nei problemi, un po' delicati anche se non difficili, della convergenza di questa serie, detta *serie di Lie* della funzione F . Nell'uso concreto (si veda ad esempio il paragrafo 4.4) è spesso sufficiente scrivere sviluppi del primo o al massimo del secondo ordine, con il resto in forma opportuna, per ottenere risultati di rilievo.

Nel trattare le trasformazioni canoniche, fino ad ora abbiamo preso in considerazione, per semplicità più che altro di notazione, hamiltoniane $H(p, q)$ non dipendenti esplicitamente dal tempo. Ma è facile verificare (basta scorrere i passaggi delle dimostrazioni) che tutto si estende senza variazione alcuna al caso più generale $H(p, q, t)$. Come vedremo nel prossimo paragrafo, c'è invece qualche differenza (marginale, a ben guardare) nel caso in cui la trasformazione w dipenda da t .

2.3 Trasformazioni canoniche dipendenti dal tempo

Consideriamo la trasformazione di Galileo in una dimensione,

$$q = \tilde{q} + Vt . \quad (2.29)$$

In ambito Lagrangiano la trasformazione si completa con $\dot{q} = \dot{\tilde{q}} + V$, e poi per ogni assegnata $L(q, \dot{q}, t)$ si pone al solito, per sostituzione, $\tilde{L}(\tilde{q}, \dot{\tilde{q}}, t) = L(\tilde{q} + Vt, \dot{\tilde{q}} + V, t)$. Il nuovo momento \tilde{p} risulta allora essere

$$\tilde{p} = \frac{\partial \tilde{L}}{\partial \dot{\tilde{q}}} = \frac{\partial L}{\partial \dot{q}} = p ,$$

e ciò suggerisce di completare il cambiamento puntuale di coordinate (2.29) con $p = \tilde{p}$. La risultante trasformazione nello spazio delle fasi

$$(p, q) = w(\tilde{p}, \tilde{q}, t) = (\tilde{p}, \tilde{q} + Vt)$$

è dipendente dal tempo. Presa ora una qualsiasi hamiltoniana $H(p, q, t)$, la trasformazione muta le sue equazioni di Hamilton in

$$\dot{\tilde{p}} = \dot{p} = -\frac{\partial H}{\partial q}(\tilde{p}, \tilde{q} + Vt, t), \quad \dot{\tilde{q}} = \dot{q} - V = \frac{\partial H}{\partial p}(\tilde{p}, \tilde{q} + Vt, t) - V,$$

e in queste espressioni si riconoscono le equazioni di Hamilton corrispondenti all'hamiltoniana

$$K(\tilde{p}, \tilde{q}, t) = H(\tilde{p}, \tilde{q} + Vt, t) - V\tilde{p}.$$

ESERCIZIO 12 *Si verifichi che K è la trasformata di Legendre di \tilde{L} .*

L'esempio si generalizza in modo ovvio in \mathbb{R}^n : la trasformazione si scrive ancora

$$p = \tilde{p}, \quad q = \tilde{q} + Vt, \quad (2.30)$$

e la nuova hamiltoniana ha la forma

$$K(\tilde{p}, \tilde{q}, t) = H(\tilde{p}, \tilde{q} + Vt, t) + K_0(\tilde{p}, \tilde{q}, t), \quad (2.31)$$

con $K_0 = -V \cdot \tilde{p}$. L'esempio mostra che per le trasformazioni dipendenti dal tempo il legame tra vecchia e nuova hamiltoniana non si può ridurre a una semplice sostituzione di variabili (con eventuale introduzione di una valenza $c \neq 1$), e qualche novità è inevitabile.

Per comprendere la novità, e in particolare trovare naturale la presenza di un termine additivo K_0 come nella (2.31), la via più semplice e soprattutto più chiara è quella di riformulare il problema nello spazio esteso a $2n + 2$ variabili già introdotto nel paragrafo 2.1 a proposito delle hamiltoniane non autonome $H(p, q, t)$. Tali hamiltoniane, ricordiamo, erano state estese ad hamiltoniane autonome $H^{\text{ext}}(p, A, q, \tau) = H(p, q, \tau) + A$: in questo modo la variabile τ coincide di fatto (supponendo $\tau(0) = 0$) con t , mentre A è una variabile ausiliaria che serve a muovere τ con $\dot{\tau} = 1$.

Estendiamo similmente la generica trasformazione di coordinate dipendente dal tempo

$$(p, q) = w(\tilde{p}, \tilde{q}, t) = (u(\tilde{p}, \tilde{q}, t), v(\tilde{p}, \tilde{q}, t))$$

alla trasformazione autonoma nello spazio esteso

$$(p, A, q, \tau) = w^{\text{ext}}(\tilde{p}, \tilde{A}, \tilde{q}, \tilde{\tau}), \quad (2.32)$$

della forma

$$\begin{aligned} p &= u(\tilde{p}, \tilde{q}, \tilde{\tau}) \\ q &= v(\tilde{p}, \tilde{q}, \tilde{\tau}) \\ A &= a(\tilde{p}, \tilde{A}, \tilde{q}, \tilde{\tau}) \\ \tau &= \tilde{\tau}, \end{aligned}$$

con $a(\tilde{p}, \tilde{A}, \tilde{q}, \tilde{\tau})$ per il momento da determinarsi; si è inserita la prescrizione $\tau = \tilde{\tau}$ in modo che, di fatto, il tempo non venga toccato. A questo punto possiamo dare una definizione:

DEFINIZIONE 5 *La trasformazione di coordinate*

$$(p, q) = w(\tilde{p}, \tilde{q}, t), \quad w : \mathbb{R}^{2n} \times \mathbb{R} \rightarrow \mathbb{R}^{2n},$$

regolare e invertibile per ogni fissato $t \in \mathbb{R}$, si dice (strettamente) canonica dipendente dal tempo se esiste a tale che la sua estensione (2.32) sia (strettamente) canonica.

Ci limiteremo per semplicità alla canonicità stretta; la generalizzazione al caso di valenza $c \neq 1$ è del tutto ovvia.

All'estensione w^{ext} si applicano evidentemente le note condizioni necessarie e sufficienti per la canonicità, in particolare la preservazione delle parentesi di Poisson; imponendo $\{\tau, A\} = 1$ si vede immediatamente che a deve avere la forma

$$a(\tilde{p}, \tilde{A}, \tilde{q}, \tilde{\tau}) = \tilde{A} + K_0(\tilde{p}, \tilde{q}, \tilde{\tau}) ,$$

con un qualche K_0 indipendente da \tilde{A} . Preso allora $H(p, q, t)$ qualsiasi, sia $H^{\text{ext}}(p, A, q, \tau) = H(p, q, \tau) + A$ la sua estensione; se w^{ext} è strettamente canonica, la nuova hamiltoniana, che denotiamo K^{ext} , deve essere

$$\begin{aligned} K^{\text{ext}}(\tilde{p}, \tilde{A}, \tilde{q}, \tilde{\tau}) &= H^{\text{ext}}(w^{\text{ext}}(\tilde{p}, \tilde{A}, \tilde{q}, \tilde{\tau})) \\ &= H(w(\tilde{p}, \tilde{q}, \tilde{\tau})) + \tilde{A} + K_0(\tilde{p}, \tilde{q}, \tilde{\tau}) , \end{aligned}$$

e in questa espressione si riconosce l'estensione dell'hamiltoniana non autonoma

$$K(\tilde{p}, \tilde{q}, t) = H(w(\tilde{p}, \tilde{q}, t)) + K_0(\tilde{p}, \tilde{q}, t) . \quad (2.33)$$

Per la supposta canonicità di w^{ext} , i moti di K^{ext} sono mappati nei moti di H^{ext} ; proiettando nel dominio delle variabili p, q si ottiene corrispondentemente che i moti di K sono mappati da w nei moti di H . L'hamiltoniana canonicamente coniugata a H è dunque, necessariamente, della forma (2.33).

Veniamo allora alle condizioni perché w^{ext} sia (strettamente) canonica; per via troveremo anche l'espressione di K_0 . Si dimostra senza troppa difficoltà che

PROPOSIZIONE 13 *Condizione necessaria e sufficiente perché w^{ext} come descritta sopra sia strettamente canonica è che w sia strettamente canonica per ogni t fissato. Inoltre, se $f(\tilde{p}, \tilde{q}, t)$ è tale che per ogni t fissato*

$$u \cdot dv = \tilde{p} \cdot d\tilde{q} + df ,$$

allora si ha $K_0 = \frac{\partial f}{\partial t} - u \cdot \frac{\partial v}{\partial t}$.

DIMOSTRAZIONE. Denotiamo con d^{ext} il differenziale nello spazio esteso. Supponiamo che w^{ext} sia canonica; allora esiste $\tilde{f}(\tilde{p}, \tilde{A}, \tilde{q}, \tilde{\tau})$ tale che

$$u \cdot d^{\text{ext}}v + A d\tau - \tilde{p} \cdot d\tilde{q} - \tilde{A} d\tilde{\tau} = d^{\text{ext}}\tilde{f} . \quad (2.34)$$

Ma per quanto detto sopra la forma di w^{ext} è necessariamente

$$\begin{aligned} p &= u(\tilde{p}, \tilde{q}, \tilde{\tau}) \\ q &= v(\tilde{p}, \tilde{q}, \tilde{\tau}) \\ A &= \tilde{A} + K_0(\tilde{p}, \tilde{q}, \tilde{\tau}) \\ \tau &= \tilde{\tau} , \end{aligned}$$

e dunque la (2.34) si riscrive

$$u \cdot dv + K_0 d\tilde{\tau} - \tilde{p} \cdot d\tilde{q} = d\tilde{f} + \frac{\partial \tilde{f}}{\partial \tilde{A}} d\tilde{A} + \frac{\partial \tilde{f}}{\partial \tilde{\tau}} d\tilde{\tau} . \quad (2.35)$$

Ma \tilde{A} a sinistra dell'uguale scompare, pertanto neanche \tilde{f} può dipendere da \tilde{A} . Dall'identità differenziale (2.35) si deduce allora

$$u \cdot dv = \tilde{p} \cdot \tilde{q} + df, \quad u \cdot \frac{\partial v}{\partial t} + K_0 = \frac{\partial f}{\partial t}, \quad (2.36)$$

con $f = \tilde{f}$ dipendente da \tilde{p}, \tilde{q}, t . Pertanto w a ogni t fissato è canonica, e K_0 ha la forma voluta. Percorrendo la dimostrazione al contrario, precisamente assumendo le (2.36), ove la seconda va intesa come definizione di K_0 , si ottiene subito la canonicità di w^{ext} . \square

Se ne deduce immediatamente che valgono anche le altre due condizioni necessarie e sufficienti: precisamente

COROLLARIO 14 *Condizioni necessarie e sufficienti per la canonicità di w^{ext} sono che w preservi per ogni t le parentesi di Poisson, e che la jacobiana di w sia simplettica per ogni t .*

Tutto si estende in modo ovvio al caso sopra trascurato di valenza $c \neq 1$, con la nuova hamiltoniana, che generalizza di poco la (2.33),

$$K(\tilde{p}, \tilde{q}, t) = c H(w(\tilde{p}, \tilde{q}, t)) + K_0(\tilde{p}, \tilde{q}, t); \quad (2.37)$$

si vede anche che c deve essere indipendente da t (c nella trasformazione estesa è una costante indipendente da tutte le variabili, in particolare da τ).

Tornando all'esempio (2.30) da cui siamo partiti, si ha $f = 0$ (la trasformazione è infatti puntuale), perciò $K_0 = -u \cdot \frac{\partial v}{\partial t} = -\tilde{p} \cdot V$, come trovato in modo diretto.

Un secondo esempio un po' curioso di trasformazione canonica dipendente da t è la trasformazione (priva di analogo lagrangiano)

$$p = \tilde{p}, \quad q = \tilde{q} + \tilde{p}t.$$

Questa trasformazione (esercizio) è canonica e muta una generica $H(p, q, t)$ nella nuova hamiltoniana

$$K(\tilde{p}, \tilde{q}, t) = H(\tilde{p}, \tilde{q} + \tilde{p}t, t) + K_0(\tilde{p}, \tilde{q}, t), \quad K_0(\tilde{p}, \tilde{q}, t) = -\frac{1}{2}\tilde{p}^2.$$

L'aspetto curioso della trasformazione è che essa muta l'hamiltoniana della particella libera di massa unitaria, $H = \frac{1}{2}p^2$, nell'hamiltoniana costante $K(p, q) = 0$. Le nuove coordinate \tilde{p} e \tilde{q} , nella dinamica di K , conservano evidentemente il valore iniziale, e in sostanza la dinamica è finita nelle equazioni della trasformazione, che per il problema della particella libera forniscono l'integrale generale (\tilde{p} e \tilde{q} sono i dati iniziali).

Come ultimo importante esempio consideriamo la rotazione uniforme in \mathbb{R}^2 :

$$\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = R_{\omega t} \begin{pmatrix} \tilde{q}_1 \\ \tilde{q}_2 \end{pmatrix}, \quad \text{ove } R_{\vartheta} = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}.$$

La trasformazione, puntuale, si estende ai momenti ponendo

$$\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = R_{\omega t} \begin{pmatrix} \tilde{p}_1 \\ \tilde{p}_2 \end{pmatrix}$$

(ci andrebbe $R_{\omega t}^{-T}$ che però coincide con $R_{\omega t}$). Si ha $f = 0$, $K_0 = -u \cdot \frac{\partial v}{\partial t}$, e si trova subito

$$K_0(\tilde{p}, \tilde{q}) = -\omega(\tilde{q}_1\tilde{p}_2 - \tilde{q}_2\tilde{p}_1) = -\omega\tilde{M}_3;$$

una riscrittura utile è anche

$$K_0(\tilde{p}, \tilde{q}) = -\tilde{p} \cdot \Omega \tilde{q}, \quad \Omega = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}. \quad (2.38)$$

Consideriamo, come naturale esempio, il caso di un moto centrale, con un qualunque potenziale $V(r)$; l'espressione (2.38) di K_0 dà

$$K(\tilde{p}, \tilde{q}) = \frac{\tilde{p}^2}{2m} + V(\tilde{r}) - \tilde{p} \cdot \Omega \tilde{q},$$

che aggiungendo e togliendo il termine $-\frac{1}{2}m\omega^2\tilde{r}^2$ si riscrive anche

$$K(\tilde{p}, \tilde{q}) = \frac{1}{2m}(\tilde{p} - m\Omega\tilde{q})^2 - \frac{1}{2}m\omega^2\tilde{r}^2 + V(\tilde{r}); \quad (2.39)$$

questa è una espressione migliore perché fa comparire il potenziale centrifugo $-\frac{1}{2}m\omega^2\tilde{r}^2$ altrimenti nascosto, e inoltre sottolinea la forte analogia tra cariche in campo magnetico e masse in sistemi rotanti ($\Omega\tilde{q}$ gioca il ruolo del potenziale vettore; m sta al posto della carica).⁸

Dal fatto che le trasformazioni canoniche dipendenti da t altro non sono che trasformazioni canoniche per ogni t fissato segue che le trasformazioni canoniche dipendenti dal tempo si possono generare attraverso una funzione generatrice in variabili miste, ad esempio una funzione di tipo S dipendente dal tempo, $S(\tilde{p}, q, t)$, scrivendo, a generalizzazione delle (2.23),

$$p = \frac{\partial S}{\partial q}(\tilde{p}, q, t), \quad \tilde{q} = \frac{\partial S}{\partial \tilde{p}}(\tilde{p}, q, t). \quad (2.40)$$

Precisamente:

PROPOSIZIONE 15 *Le (2.40) definiscono implicitamente una trasformazione canonica dipendente dal tempo. La nuova hamiltoniana risulta essere $K = H \circ w + K_0$, con*

$$K_0(\tilde{p}, \tilde{q}, t) = \frac{\partial S}{\partial t}(\tilde{p}, q(\tilde{p}, \tilde{q}, t), t), \quad (2.41)$$

ove $q(\tilde{p}, \tilde{q}, t)$ è ottenuta per inversione a partire dalla seconda delle (2.40).

La verifica della (2.41) è lasciata come esercizio.

- Si potrebbe pensare di generalizzare la definizione 5 di trasformazione canonica dipendente dal tempo, chiedendo semplicemente, a somiglianza della definizione 4 di trasformazione canonica, che per ogni $H(p, q, t)$ esista $K(\tilde{p}, \tilde{q}, t)$ tale che le equazioni di Hamilton relative a H e a K siano equivalenti. Concettualmente la definizione appare più ampia, di fatto non lo è: si dimostra infatti che tutte le trasformazioni canoniche in questo senso apparentemente più ampio rientrano nella definizione 5 e danno in particolare K della forma (2.37).

⁸Vale la pena di osservare che se si parte dalla lagrangiana del moto centrale, si passa alle coordinate rotanti, e infine con la trasformata di Legendre si passa al formalismo Hamiltoniano, K appare spontaneamente nella forma 2.39. E' un buon esercizio.

2.4 Qualche esercizio

ESERCIZIO 13 Si consideri la trasformazione puntuale $q = v(\tilde{q})$ e si denoti con A la matrice jacobiana $\frac{\partial v}{\partial \tilde{q}}$. Si determini una condizione sulla funzione $a(\tilde{q})$ tale che

$$p = A^{-T}(\tilde{q})[\tilde{p} + a(\tilde{q})]$$

sia un completamento canonico di v . Esistono altri completamenti canonici di v ? [Risposta: la condizione è $\frac{\partial a_i}{\partial \tilde{q}_j} = \frac{\partial a_j}{\partial \tilde{q}_i}$; non esistono altri modi di completare canonicamente v .]

ESERCIZIO 14 Si determini il completamento canonico del passaggio alle coordinate polari del piano:

$$q_1 = r \cos \vartheta, \quad q_2 = r \sin \vartheta.$$

Si scriva l'hamiltoniana di un generico moto centrale. [Risposta: il completamento è

$$p_1 = p_r \cos \vartheta - (p_\vartheta/r) \sin \vartheta, \quad p_2 = p_r \sin \vartheta + (p_\vartheta/r) \cos \vartheta,$$

e viceversa si ha

$$p_r = (p \cdot q)/r, \quad p_\vartheta = q_1 p_2 - q_2 p_1$$

(componente radiale del momento; terza componente del momento angolare). L'hamiltoniana del moto centrale $H = \frac{p^2}{2m} + V(r)$ è mutata in $\tilde{H} = \frac{p_r^2}{2m} + \frac{p_\vartheta^2}{2mr^2} + V(r)$.]

ESERCIZIO 15 Si consideri la trasformazione ($n = 2$)

$$\begin{aligned} p_1 &= \tilde{p}_1 + \tilde{q}_1 & p_2 &= \tilde{p}_2 + 2\tilde{q}_2 \\ q_1 &= \tilde{p}_1 + 2\tilde{q}_1 + \tilde{p}_2 + a\tilde{q}_2 & q_2 &= \tilde{p}_2 + b\tilde{q}_2 + \tilde{p}_1 + \tilde{q}_1 \end{aligned}$$

e si determinino a e b che la rendono canonica. [Risposta: $a = 2$, $b = 3$.]

ESERCIZIO 16 Si consideri la trasformazione ($n = 1$)

$$p = a I^\alpha \cos \varphi, \quad q = b I^\beta \sin \varphi.$$

Si determinino a , b , α , β che la rendono strettamente canonica. [Risposta: l'unica possibilità è $ab = 2$, $\alpha = \beta = \frac{1}{2}$, cioè la trasformazione (2.6) a meno di un riscalamento.]

ESERCIZIO 17 Si consideri la trasformazione ($n = 1$)

$$p = c\tilde{p}e^{-\beta\tilde{q}}, \quad q = ce^{\beta\tilde{q}}$$

e si determinino c , β che la rendono canonica. [Risposta: è canonica per ogni c , β ; strettamente canonica per $c^2\beta = 1$.]

ESERCIZIO 18 Si consideri la trasformazione ($n = 1$)

$$p = c\tilde{p}^\alpha e^{\beta\tilde{q}}, \quad q = d\tilde{p}^\alpha e^{\gamma\tilde{q}}.$$

Si determinino c , d , α , β , γ che la rendono strettamente canonica. [Risposta: $\alpha = \frac{1}{2}$, $\gamma = -\beta$, $cd\beta = -1$. Il caso complesso $\beta = i$ e ad esempio $c = 1$, $d = i$, è di uso concreto in teoria delle perturbazioni.]

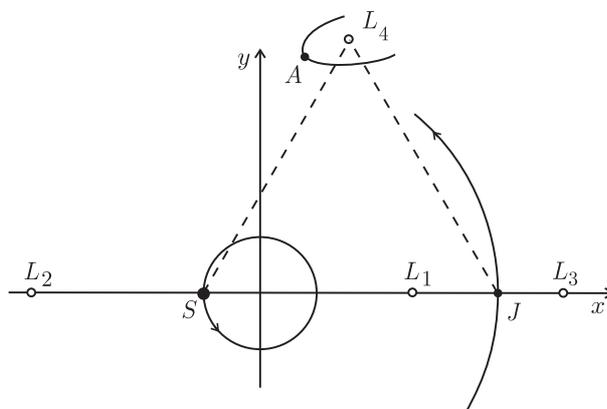


Figura 3: I punti di equilibrio lagrangiani nel problema a tre corpi circolare ristretto.

ESERCIZIO 19 Si consideri un qualunque moto centrale con energia potenziale $V(r)$. Si passi a un sistema rotante con velocità angolare ω , determinando ω in modo che nel sistema rotante ci sia un punto di equilibrio in $\tilde{q}^* = (a, 0)$. Si studi in particolare il caso kepleriano $V(r) = -km/r$. [Risposta: in generale la condizione è $m\omega^2 = V'(a)$; nel caso kepleriano si ha $\omega^2 = k/a^3$ (terza legge di Keplero).]

ESERCIZIO 20 (problema a due corpi nel sistema corotante) Si consideri il problema piano a due corpi con potenziale Kepleriano:

$$H = \frac{P^2}{2M} + \frac{p^2}{2m} - \frac{GMm}{r},$$

ove

$$P^2 = P_1^2 + P_2^2, \quad p^2 = p_1^2 + p_2^2, \quad r^2 = (Q_1 - q_1)^2 + (Q_2 - q_2)^2.$$

Ci si metta in un sistema baricentrico rotante con velocità angolare ω e si determini ω in modo che nel sistema rotante vi sia la configurazione di equilibrio

$$\tilde{Q}_1^* = -\frac{m}{M+m}a, \quad \tilde{q}_1^* = \frac{M}{M+m}a, \quad Q_2^* = q_2^* = 0.$$

[Risposta: $\omega^2 = G(M+m)/a^3$ (terza legge di Keplero per m non trascurabile).]

ESERCIZIO 21 (comincia a essere difficile) Si consideri il problema precedente. Ci si metta nel sistema corotante in cui M e m sono ferme, e si aggiunga un terzo corpo di massa trascurabile, cioè non in grado di alterare il moto di M e m . (Problema a tre corpi “ristretto”: il Sole, un pianeta, un asteroide. Problema “circolare” avendo messo il Sole e il pianeta in moto circolare anziché ellittico; “piano” se ci si restringe a moti dell’asteroide nel piano del moto di M, m .) (1) Si scriva l’hamiltoniana del terzo corpo; (2) si studino le sue configurazioni di equilibrio (moti circolari corotanti). [Risposta: in unità di misura tali che $a = 1$ e $\omega = 1$, l’hamiltoniana è

$$H(p, q) = \frac{1}{2}p^2 - (q_1p_2 - q_2p_1) - \frac{1-\eta}{\|q - (-\eta, 0, 0)\|} - \frac{\eta}{\|q - (1-\eta, 0, 0)\|},$$

ove $\eta = m/(M+m)$. Si trovano cinque configurazioni di equilibrio (“punti lagrangiani” L_1, \dots, L_5). Tre sono collineari a M, m , due invece sono tali da formare con M e m un triangolo equilatero (“punti lagrangiani triangolari” L_4, L_5). Si veda la figura 3.]

ESERCIZIO 22 (difficile) *Si studi la stabilità lineare di L_4 e L_5 .* [Risposta: sono linearmente stabili se il rapporto di massa $\eta = m/(M + m)$ è sufficientemente piccolo, precisamente se si ha $\eta < \eta_R = \frac{1}{2}(1 - \sqrt{\frac{23}{27}}) \simeq 0.039$ (“limite di Routh”). Per il sistema Sole–Giove si ha $\eta \simeq 10^{-3}$; per il sistema Terra–Luna $\eta \simeq 0.012$. Più precisamente, prendendo come origine delle coordinate il punto di equilibrio, e introducendo opportune coordinate normali che denotiamo Q_i, P_i , se $\eta < \eta_R$ l’hamiltoniana troncata al secondo ordine ha la forma

$$H_2(P, Q) = \sum_{i=1}^3 \frac{\omega_i}{2} (P_i^2 + Q_i^2) .$$

Le ω_i tuttavia non hanno segno concorde (si veda per i dettagli l’appendice E): ciò non pregiudica la stabilità lineare, ma rende problematica la stabilità non lineare.]

ESERCIZIO 23 (molto, molto, molto difficile; un vero abuso del termine esercizio) *Si studi la stabilità (senza approssimazione lineare) di L_4, L_5 .* [Risposta: nel caso piano Arnol’d (1962) ha dimostrato la stabilità non lineare, come applicazione del cosiddetto “Teorema KAM” (paragrafo 4.5). Nel problema non piano la questione è ancora parzialmente aperta.]

3 Sistemi hamiltoniani “integrabili”

3.1 Nozione di sistema integrabile

Consideriamo un sistema hamiltoniano autonomo a n gradi di libertà, di hamiltoniana $H(p, q)$; con evidenza se H non dipende da una o più coordinate, ad esempio q_{m+1}, \dots, q_n , $0 \leq m < n$, allora i corrispondenti momenti coniugati p_{m+1}, \dots, p_n sono costanti del moto del sistema, e in pratica il sistema hamiltoniano a n gradi di libertà si riduce a un sistema a soli m gradi di libertà. Se poi H è indipendente da tutte le coordinate q_1, \dots, q_n , cioè si ha $H(p, q) = h(p)$, allora l’integrazione delle equazioni del moto è del tutto banale: posto infatti $\omega_i(p) = \frac{\partial h}{\partial p_i}$, $i = 1, \dots, n$, le equazioni del moto divengono $\dot{p}_i = 0$, $\dot{q}_i = \omega_i(p)$, e quindi l’integrale generale si scrive

$$p_i(t) = p_i^o, \quad q_i(t) = q_i^o + \omega_i(p^o)t, \quad i = 1, \dots, n, \quad (3.1)$$

dove p_i^o, q_i^o , $i = 1, \dots, n$, indicano i dati iniziali. Le coordinate assenti dall’hamiltoniana sono chiamate *variabili cicliche*; la ragione di questa terminologia sta nel fatto che tipicamente (più precisamente nel caso in cui il moto sia limitato, di gran lunga il più interessante) queste variabili sono angoli, che come mostrano le (3.1) avanzano (ruotano) con velocità angolare costante. In questo caso ciascuna delle q_i sta su un cerchio, e dunque l’ n -pla $q = (q_1, \dots, q_n)$ sta sul toro n -dimensionale $\mathbb{T}^n = \mathbb{R}^n / (2\pi\mathbb{Z})^n$. Se il dominio di p è l’aperto $B \subset \mathbb{R}^n$, allora lo spazio delle fasi è il prodotto $B \times \mathbb{T}^n$; possiamo pensare che in ogni punto della *base* B stia attaccato come *fibra* un toro \mathbb{T}^n , e allora il moto avviene sulla fibra ed è lineare.

Una struttura di questo genere può essere nascosta dal fatto che non si sta descrivendo il sistema nel modo appropriato. Precisamente, data una hamiltoniana $H(p, q)$ dipendente in generale da tutte le variabili, $H : U \rightarrow \mathbb{R}$, è possibile che esista una trasformazione canonica a nuove variabili che denoteremo per comodità $(I, \varphi) = (I_1, \dots, I_n, \varphi_1, \dots, \varphi_n) \in B \times \mathbb{T}^n$,

$$(p, q) = w(I, \varphi), \quad w : B \times \mathbb{T}^n \rightarrow w(B \times \mathbb{T}^n) \subset U,$$

tale che la nuova hamiltoniana $K = H \circ w$ dipenda solo da I ,

$$K(I, \varphi) = h(I) .$$

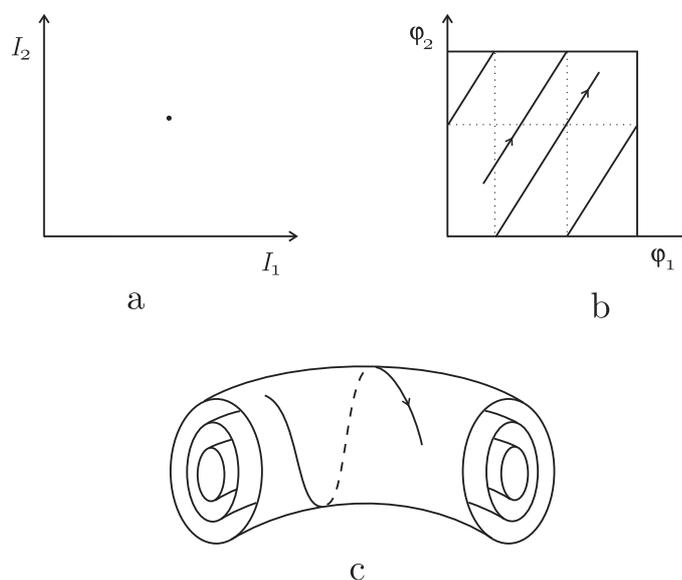


Figura 4: il moto quasi periodico su \mathbb{T}^n , in due diverse rappresentazioni

Il moto nelle nuove variabili si svolge allora su un toro \mathbb{T}^n ed è lineare:

$$I(t) = I^o, \quad \varphi(t) = \varphi^o + \omega(I^o)t \quad \text{con} \quad \omega = \frac{\partial h}{\partial I},$$

e nelle nuove variabili lo spazio delle fasi appare naturalmente decomposto nel prodotto $B \times \mathbb{T}^n$.⁹ La figura 4a,b illustra il comportamento delle variabili I e φ .

La corrispondente struttura esiste evidentemente anche nelle vecchie variabili, semplicemente trasformata da w . In particolare, ogni moto di H (nel dominio $D = w(B \times \mathbb{T}^n) \subset U$, che risulta invariante) ha la forma

$$(p(t), q(t)) = w(I^o, \varphi^o + \omega(I^o)t),$$

ovvero è quasi periodico.¹⁰ La figura 4c dà un'immagine pittorica di una famiglia di questi tori nello spazio delle fasi.

Prendiamo ad esempio un sistema linearizzato attorno a un punto di equilibrio stabile, con hamiltoniana della forma

$$H(p, q) = \frac{1}{2}p \cdot A^{-1}p + \frac{1}{2}q \cdot Bq, \quad (3.2)$$

ove A e B sono matrici simmetriche definite positive. Le coordinate p, q non rivelano immediatamente la struttura geometrica del sistema. Ma come insegna la teoria delle piccole oscillazioni, si può innanzitutto passare alle coordinate normali, che qui denotiamo \tilde{p}, \tilde{q} , ottenendo un sistema di n oscillatori armonici disaccoppiati:

$$\tilde{H}(\tilde{p}, \tilde{q}) = \frac{1}{2} \sum_{i=1}^n (\tilde{p}_i^2 + \omega_i^2 \tilde{q}_i^2); \quad (3.3)$$

⁹E' completamente equivalente pensare agli angoli φ sul toro \mathbb{T}^n oppure pensare, più semplicemente ma meno geometricamente, che gli angoli corrono su \mathbb{R}^n , ma ovunque ci si restringe funzioni periodiche di periodo 2π negli angoli.

¹⁰Una funzione $f(t)$ si dice *quasi periodica con n frequenze* $\omega_1, \dots, \omega_n$ se esiste $F(\varphi_1, \dots, \varphi_n)$, periodica di periodo 2π in ciascun argomento, tale che $f(t) = F(\omega_1 t, \dots, \omega_n t)$.

poi, procedendo come nell'esempio (v) del paragrafo 2.1, si possono introdurre le variabili di azione–angolo di ciascun oscillatore:

$$\tilde{p}_i = \sqrt{2\omega_i I_i} \cos \varphi_i, \quad \tilde{q}_i = \sqrt{2I_i/\omega_i} \sin \varphi_i,$$

ottenendo una hamiltoniana dipendente dalle sole azioni, precisamente

$$K(I, \varphi) = \omega \cdot I.$$

Sistemi con questa proprietà sono detti *integrabili* nel senso di Liouville. Una possibile definizione formale è questa:

DEFINIZIONE 6 *L'hamiltoniana $H : U \rightarrow \mathbb{R}$ si dice integrabile secondo Liouville nel dominio $D \subset U$ se esiste una trasformazione canonica $(p, q) = w(I, \varphi)$, definita per $I \in B \subset \mathbb{R}^n$ e $\varphi \in \mathbb{T}^n$, tale che (i) $D = w(B \times \mathbb{T}^n)$, e (ii) la nuova hamiltoniana $K = H \circ w$ è funzione delle sole azioni:*

$$K(I, \varphi) = h(I), \quad h : B \rightarrow \mathbb{R}.$$

Con evidenza non tutti i sistemi sono integrabili: ad esempio, non appena nel dominio c'è un punto di equilibrio instabile (p^*, q^*) con una separatrice e un moto $(p(t), q(t))$ che vi converge per $t \rightarrow \infty$, questo moto non è quasi periodico e dunque il sistema non è integrabile. L'integrabilità, in un senso che si potrebbe definire, è un fatto eccezionale.¹¹ Ma è un fatto di grande importanza, perché riguarda alcuni tra i più significativi sistemi della fisica, come i sistemi armonici (masse legate da forze elastiche), il corpo rigido di Eulero (corpo rigido con un punto fisso, in assenza di forze attive) o di Lagrange (corpo rigido simmetrico con un punto fisso, nella gravità: la comune trottola), e soprattutto il sistema di Keplero e ogni altro moto centrale. Nello studio di questi e (pochi) altri simili problemi, l'apparato hamiltoniano non va però pensato come una via per risolvere le equazioni del moto (tutti questi problemi si sapevano già risolvere prima della meccanica hamiltoniana); l'apparato hamiltoniano ne fornisce piuttosto una descrizione più profonda, utile sia a capirli meglio da un punto di vista geometrico, sia soprattutto a studiare il problema *molto interessante ma molto difficile* dei sistemi prossimi a questi: ad esempio un sistema di masse legate da forze leggermente non lineari; un corpo rigido in un debole campo di forze (equivalentemente, in rapida rotazione); due o più problemi di Keplero debolmente accoppiati (modelli di Sistema Solare, e numerosi altri problemi di Meccanica Celeste). Sono questioni difficili, nelle quali non ci sarà possibile addentrarci veramente. Ma ci arriveremo alla soglia, riuscendo a dare almeno un'occhiata all'interno.

3.2 Le variabili di azione–angolo per il pendolo

Per prendere familiarità con la nozione di sistema integrabile studiamo qui il caso facile (ma non banale: c'è da imparare) dei sistemi a un grado di libertà. Per essere definiti ci riferiremo al pendolo:

$$H(p, q) = \frac{p^2}{2} + V(q), \quad V(q) = -\omega^2 \cos q, \quad (p, q) \in \mathbb{R}^2, \quad (3.4)$$

ma il modo di procedere è generale e riguarda sostanzialmente tutti i sistemi con un grado di libertà che abbiano moti limitati.

¹¹Questa cosa, pur chiara negli scritti di Poincaré, si è in realtà compresa bene relativamente di recente (si pensi all'età che hanno questi problemi!), diciamo nel corso degli anni '60 del '900.

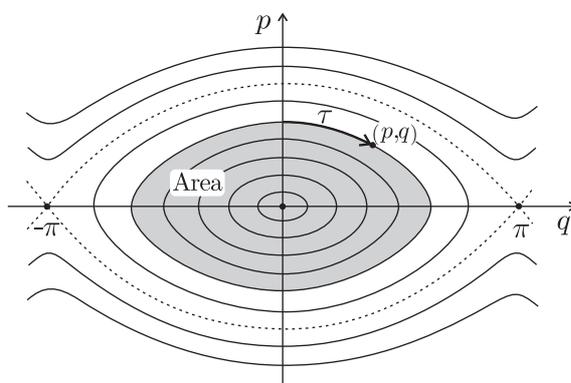


Figura 5: La costruzione delle variabili di azione–angolo per il pendolo.

Il ritratto in fase del sistema è riportato in figura 5; le traiettorie coincidono con le superfici di livello (curve di livello) dell'energia, che denotiamo

$$\Sigma_E = \{(p, q) \in \mathbb{R}^2 : H(p, q) = E\} . \quad (3.5)$$

Ci restringiamo alle librazioni (il caso delle rotazioni è semmai più semplice), escludendo anche il punto di equilibrio (l'origine), ovvero ci restringiamo all'intervallo di energia $-\omega^2 < E < \omega^2$ e al corrispondente dominio $D \subset \mathbb{R}^2$. Allora lo spazio delle fasi topologicamente è il prodotto di un intervallo per un cerchio, ovvero un cilindro (o se si preferisce una corona). Vogliamo costruire coordinate canoniche (\tilde{p}, \tilde{q}) sul cilindro, tali che \tilde{p} resti costante e \tilde{q} avanzi uniformemente. Una prima scelta spontanea per \tilde{p} è l'energia stessa

$$E = H(p, q) ,$$

mentre una quantità che avanza uniformemente nel tempo, buona candidata (anche dimensionalmente) ad essere la nuova coordinata \tilde{q} , è il tempo stesso: precisamente, fissata una sezione \mathcal{S} che tagli tutte le traiettorie, come la semiretta $q = 0, p > 0$, poniamo

$$\tau = \mathcal{T}(p, q) = \text{tempo per arrivare da } \mathcal{S} \text{ a } (p, q) ,$$

ovvero (ricordando che $dt = dq/v$, con in questo caso $v = p = \pm \sqrt{2(H(p, q) - V(q))}$)

$$\begin{aligned} \tau &= \int_0^q \frac{dx}{\sqrt{2(H(p, q) - V(x))}} && (p \geq 0) \\ &= \frac{1}{2} T(H(p, q)) - \int_0^q \frac{dx}{\sqrt{2(H(p, q) - V(x))}} && (p < 0) , \end{aligned} \quad (3.6)$$

ove $T(E)$ è il periodo:

$$T(E) = 2 \int_{q_-(E)}^{q_+(E)} \frac{dx}{\sqrt{2(E - V(x))}} , \quad (3.7)$$

avendo denotato con $q_{\pm}(E)$ i due punti di arresto all'energia E . In questo modo restano definite E e τ come funzioni di p e q , ovvero resta definita la trasformazione inversa

$$\begin{aligned} (E, \tau) &= w^{-1}(p, q) \\ &= (H(p, q), \mathcal{T}(p, q)) . \end{aligned}$$

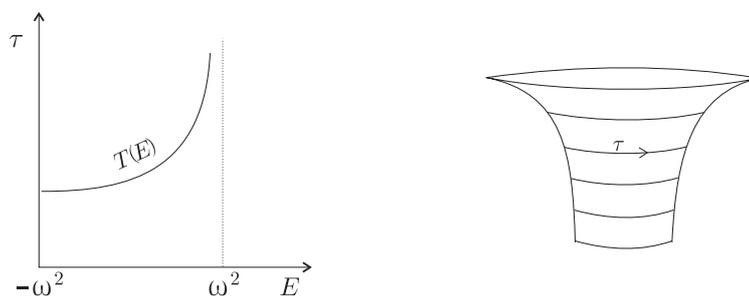


Figura 6: La curva $T(E)$ e il corrispondente dominio delle variabili E, τ .

Il tempo τ è in realtà funzione multivoca di (p, q) , definita a meno di multipli del periodo $T(E)$; corrispondentemente, se scriviamo

$$(p, q) = w(E, \tau),$$

la funzione w è periodica in τ di periodo $T(E)$.

Si verifica senza difficoltà che w^{-1} è canonica, e dunque lo è w . Per questo è sufficiente verificare che risulta $\{\tau, E\} = 1$, più precisamente $\{\mathcal{J}, H\} = 1$ (è questa l'unica parentesi di Poisson elementare non banale). La cosa è del tutto ovvia: $\{\mathcal{J}, H\}$ altro non è che $\dot{\mathcal{J}}$ nella dinamica di H , ed evidentemente $\dot{\mathcal{J}} = 1$ (se denotiamo $\hat{p}(t), \hat{q}(t)$ il movimento, allora $\mathcal{J}(\hat{p}(t), \hat{q}(t)) = t$).

ESERCIZIO 24 Si calcoli esplicitamente la parentesi di Poisson $\{\mathcal{J}, H\}$, usando la (3.6) e l'espressione (3.4) di H . Convieni osservare preliminarmente che la dipendenza di \mathcal{J} da p e q attraverso $H(p, q)$, a denominatore della funzione integranda, non contribuisce a $\{\mathcal{J}, H\}$.

ESERCIZIO 25 Si determini la funzione generatrice $S(E, q)$ della trasformazione w . [Risposta: $S(E, q) = \int_0^q \sqrt{2(E - V(x))} dx$.]

Nelle nuove variabili l'hamiltoniana e le equazioni del moto sono evidentemente

$$\tilde{H}(E, \tau) = E, \quad \dot{E} = 0, \quad \dot{\tau} = 1.$$

Questa scelta della coppia energia–tempo come variabili che integrano il sistema con moto lineare ha in realtà una complicazione nascosta: il periodo infatti dipende da E , pertanto il dominio della coppia (E, τ) non ha forma semplice (non è il prodotto di un dominio per ciascuna variabile); si veda la figura 6. Per rimediare, sostituendo a τ un vero angolo con periodo fisso 2π , possiamo cercare nuove coordinate canoniche (I, φ) , con

$$\varphi = 2\pi \frac{\tau}{T(E)}$$

e $I = I(E)$ che completi canonicamente la trasformazione (è indispensabile che I sia funzione della sola E). La funzione $I(E)$ si trova subito: infatti per la canonicità deve risultare

$$1 = \{\varphi, I\} = \frac{\partial \varphi}{\partial \tau} \frac{\partial I}{\partial E},$$

ovvero

$$\frac{\partial I}{\partial E} = \frac{1}{2\pi} T(E). \quad (3.8)$$

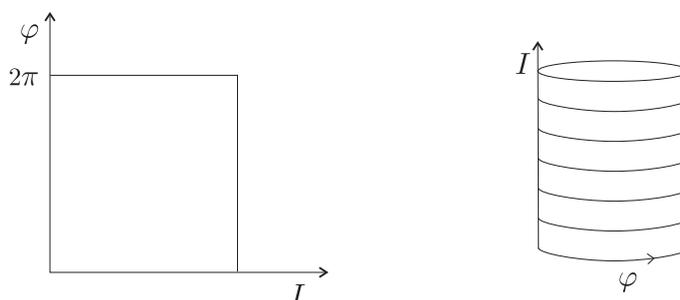


Figura 7: Il dominio rettangolare (cilindrico) nelle variabili I, φ .

Si vede facilmente che la funzione cercata è l'area sottesa dalla curva Σ_E , divisa per 2π :

$$I(E) = \frac{1}{2\pi} \oint_{\Sigma_E} p \, dq = \frac{1}{\pi} \int_{q_-(E)}^{q_+(E)} \sqrt{2(E - V(x))} \, dx . \quad (3.9)$$

Derivando rispetto a E si ottiene infatti la (3.8) con $T(E)$ espresso dalla (3.7). (Vale la pena di osservare che la dipendenza da E attraverso gli estremi non conta nella derivata, grazie al fatto che la funzione integranda agli estremi si annulla.)

Il dominio D delle nuove variabili ora è rettangolare, $D = (0, I_{\max}) \times S^1$; si veda la figura 7. La nuova hamiltoniana ha la forma

$$K(I, \varphi) = h(I) ,$$

avendo indicato con $E = h(I)$ la funzione che inverte la (3.9), e le equazioni del moto sono

$$\dot{I} = 0 , \quad \dot{\varphi} = \omega(I) ,$$

con $\omega = \frac{\partial h}{\partial I}$. Abbiamo costruito in questo modo le variabili di azione–angolo per il pendolo, nel dominio ristretto alle librazioni.

ESERCIZIO 26 *In modo analogo si introducano le variabili di azione–angolo per le rotazioni del pendolo ($E > \omega^2$).*

ESERCIZIO 27 *Si verifichi che per l'oscillatore armonico l'azione $I(E)$, definita come $\text{area}/2\pi$ secondo la (3.9), coincide con E/ω .*

- È bene sottolineare il fatto che nella nostra trattazione abbiamo dovuto escludere dallo spazio delle fasi del sistema l'origine, ovvero il punto di equilibrio stabile, e le separatrici, ovvero la superficie di livello dell'energia contenente il punto di equilibrio instabile. Nell'origine si ha un'evidente singolarità geometrica: una famiglia di curve invarianti collassa lì in un punto, e in un intorno del punto singolare non è più possibile definire né τ né φ . È la consueta singolarità delle coordinate polari. Sulle separatrici invece la singolarità è di altra natura: lì il periodo $T(E)$ diverge, i moti non sono più periodici, sarebbe ancora possibile introdurre la coppia di variabili (E, τ) ma non più la coppia (I, φ) . È interessante osservare (perché la cosa si generalizzerà) che i valori di E da escludere sono tutti e soli quelli per cui Σ_E definito dalla (3.5) non è una curva (una sottovarietà dello spazio delle fasi), poiché $\partial_{p,q}H$ non è ovunque diverso da zero su Σ_E .

- Quanto abbiamo fatto può sembrare tutto implicito e difficilmente utilizzabile. In realtà abbiamo fatto una cosa grossa: la soluzione di un’equazione differenziale è stata riportata alle operazioni assai più elementari di integrazione ordinaria e inversione di funzioni. In questo modo ad esempio è più facile fare approssimazioni numeriche delle soluzioni, e costruire grafici, *senza il deterioramento al crescere di t che ci sarebbe tentando di risolvere numericamente le equazioni differenziali*. Questo era, in buona sostanza, il programma di Liouville: riportare ove possibile la soluzione delle equazioni di Hamilton alle operazioni di quadratura e inversione. In effetti, i sistemi integrabili secondo Liouville, che stiamo trattando, si dicono anche *integrabili per quadrature*. La posta in gioco in realtà, come si vedrà nei prossimi paragrafi, è assai più alta del fare approssimazioni: si tratta di capire la struttura dello spazio delle fasi dei sistemi integrabili e su queste basi comprendere la vera cosa che interessa, ovvero il comportamento dei sistemi prossimi a sistemi integrabili, di cui è disseminata la fisica.

3.3 Il teorema di Liouville–Arnol’d

Il procedimento introdotto nel paragrafo precedente si applica a tutti i sistemi hamiltoniani a un grado di libertà, attorno a superfici di livello Σ_E sulle quali sia, in ogni punto, $\partial_{p,q}H \neq 0$ (assenza di punti di equilibrio). Il problema *fondamentale* che ci poniamo è di generalizzare queste considerazioni a sistemi con n gradi di libertà. Una generalizzazione banale è quella di n sistemi a un grado di libertà disaccoppiati:

$$H(p, q) = \sum_{i=1}^n F_i(p_i, q_i) .$$

In questo caso se, per una certa scelta di $f = (f_1, \dots, f_n)$, si ha che $\partial_{p_i, q_i} F_i$ non si annulla mai sulla curva di livello definita da $F_i(p_i, q_i) = f_i$ (ciò assicura che si tratta di una vera curva), e se questa è chiusa, allora la varietà su cui si svolge il moto, che denotiamo Σ_f , è un toro n -dimensionale (il prodotto delle n curve) e il moto su di esso è quasi periodico con n frequenze. In questa facile situazione, procedendo separatamente per ciascun sottosistema, si possono chiaramente introdurre in un intorno di Σ_f variabili di energia–tempo (E, τ) e poi variabili di azione–angolo (I, φ) , che danno a ciascun sottosistema la forma $K_i(I_i, \varphi_i) = h_i(I_i)$ e al sistema complessivo la forma integrabile

$$K(I, \varphi) = h(I) , \quad h(I) = h_1(I_1) + \dots + h_n(I_n) .$$

L’idea fondamentale di Liouville (metà ’800), completata (inizio anni ’60 del ’900) da Arnol’d, è che questa situazione persiste anche nel caso interagente, purché esistano n costanti del moto F_1, \dots, F_n indipendenti e inoltre “in involuzione”, precisamente con mutua parentesi di Poisson nulla:¹²

$$\{F_i, F_j\} = 0 , \quad i, j = 1, n .$$

Vale infatti il fondamentale *Teorema di Liouville–Arnol’d*:

¹²Quest’ultima condizione, come si è accennato nel paragrafo 1.2, è necessaria e sufficiente perché gli n flussi hamiltoniani associati alle n hamiltoniane F_1, \dots, F_n commutino:

$$\Phi_{F_i}^{\tau_i} \circ \Phi_{F_j}^{\tau_j} = \Phi_{F_j}^{\tau_j} \circ \Phi_{F_i}^{\tau_i} \quad \forall i, j ;$$

si veda l’appendice C.

PROPOSIZIONE 16 Sia U un aperto di \mathbb{R}^{2n} munito di coordinate canoniche $(p, q) = (p_1, \dots, p_n, q_1, \dots, q_n)$, e siano $F_1, \dots, F_n, F_i : U \rightarrow \mathbb{R}$, funzioni regolari a due a due in involuzione:

$$\{F_i, F_j\} = 0 \quad \forall i, j . \quad (3.10)$$

Per $f \in \mathbb{R}^n$ denotiamo

$$\Sigma_f = \{(p, q) \in U : F_i(p, q) = f_i, i = 1, \dots, n\} ,$$

e assumiamo che F_1, \dots, F_n siano indipendenti su $\Sigma_{f^*} \neq \emptyset$ per un opportuno $f^* \in \mathbb{R}^n$, precisamente

$$\text{rango} \frac{\partial(F_1, \dots, F_n)}{\partial(p_1, \dots, p_n, q_1, \dots, q_n)} = n \quad (3.11)$$

in ogni punto di Σ_{f^*} . Allora

- i. Σ_{f^*} è una varietà n -dimensionale, invariante per ciascuno degli n flussi hamiltoniani $\Phi_{F_i}^t$.
- ii. Σ_{f^*} , se compatta e connessa,¹³ è diffeomorfa al “toro standard” $\mathbb{T}^n = \mathbb{R}^n / (2\pi\mathbb{Z})^n$;
- iii. esiste un intorno $\mathcal{F} \subset \mathbb{R}^n$ di f^* , tale che l’insieme

$$\Sigma_{\mathcal{F}} = \bigcup_{f \in \mathcal{F}} \Sigma_f$$

(“intorno tubolare di Σ_{f^*} ”) è diffeomorfo a $\mathcal{F} \times \mathbb{T}^n$;

- iv. in $\Sigma_{\mathcal{F}}$ esistono coordinate canoniche di azione–angolo, precisamente esiste un cambiamento di coordinate $(p, q) = w(I, \varphi)$, definito per

$$(I, \varphi) \in B \times \mathbb{T}^n , \quad B \subset \mathbb{R}^n ,$$

tale che $\Sigma_{\mathcal{F}} = w(B \times \mathbb{T}^n)$, e ciascuna funzione trasformata $\tilde{F}_i = F_i \circ w$ è funzione delle sole azioni,

$$\tilde{F}_i(I, \varphi) = \tilde{f}_i(I) ;$$

viceversa le azioni I sono funzioni invertibili delle sole F , precisamente esiste $\mathcal{J} : \mathcal{F} \rightarrow B$ tale che $I(p, q) = \mathcal{J}(F(p, q))$.

Si osservi che l’enunciato è simmetrico in F_1, \dots, F_n : una qualsiasi F_i può essere l’hamiltoniana che ci interessa studiare, e le altre sono allora le sue costanti del moto. Se quella che ci interessa è denotata H , allora il teorema dice, in particolare, che ogni moto di H (nel dominio in cui il teorema si applica) si svolge su un toro n -dimensionale, è quasi periodico con n frequenze, e si possono introdurre coordinate di azione angolo (I, φ) che mutano H in una nuova hamiltoniana K dipendente solo dalle azioni. L’hamiltoniana H è allora integrabile secondo Liouville, nel dominio

¹³Ci si può comunque sempre restringere a una sua componente connessa. Se manca la compattezza (ma ugualmente i flussi sono prolungabili a $t \in \mathbb{R}$) allora Σ_f è diffeomorfa al cilindro $\mathcal{C}^{n,k}$ prodotto di k cerchi e $n - k$ rette, $\mathcal{C}^{n,k} = \mathbb{R}^n / (2\pi\mathbb{Z})^k$; nel caso compatto, l’unico veramente interessante, si ha $k = n$ e il cilindro diventa un toro. Nel caso non compatto si possono comunque introdurre k coppie di coordinate di azione–angolo I_i, φ_i , mentre per gli altri gradi di libertà è possibile introdurre coordinate di energia–tempo E_i, τ_i (con le τ_i non periodiche).

$\Sigma_{\mathcal{F}}$, secondo la definizione 6.¹⁴ Parimenti integrabile è ogni hamiltoniana che sia funzionalmente dipendente da F_1, \dots, F_n , cioè $H(p, q) = \mathcal{H}(F_1(p, q), \dots, F_n(p, q))$ con $\mathcal{H} : \mathbb{R}^n \rightarrow \mathbb{R}$ qualsiasi. (Come lo si vede?)

La dimostrazione del teorema, complessa e non facile, sicuramente va oltre gli scopi di queste note. Nell’appendice F tuttavia è riportata la dimostrazione dei punti i. e ii., che sono la parte più interessante.

- La scelta delle variabili di azione–angolo non è mai unica. In particolare se (I, φ) è una scelta possibile, e A è una qualunque matrice intera con $|\det A| = 1$ (equivalentemente: A è intera con inversa intera), allora la trasformazione canonica

$$I = A^{-T} J, \quad \varphi = A\psi \quad (3.12)$$

fornisce un’altra scelta possibile di variabili di azione–angolo. La condizione su A garantisce che alla trasformazione lineare $\varphi = A\psi$ in \mathbb{R}^2 corrisponda una mappa regolare e biunivoca: $\mathbb{T}^2 \rightarrow \mathbb{T}^2$; si veda l’esercizio qui sotto. La libertà offerta dalla (3.12) in qualche caso è molto utile; si veda ad esempio l’esercizio 31.

ESERCIZIO 28 *Si dimostri che alla mappa lineare $\varphi = A\psi$ su \mathbb{R}^2 corrisponde una mappa regolare biunivoca: $\mathbb{T}^2 \rightarrow \mathbb{T}^2$ se e solo se A e A^{-1} sono intere. [Bisogna e basta che, nei due versi, punti equivalenti di \mathbb{R}^2 , cioè punti le cui coordinate differiscono per multipli interi di \mathbb{R}^2 , siano inviati in punti equivalenti.]*

Vediamo ora alcuni esempi di sistemi integrabili, presi tra quelli fisicamente più interessanti.

(i) *I sistemi isocroni*, precisamente sistemi con l’hamiltoniana della forma (3.2), di cui si è già detto. Sono chiamati anche sistemi lineari in quanto le equazioni del moto sono lineari. Dopo il passaggio ai modi normali l’hamiltoniana ha la forma (3.3); sono costanti del moto in involuzione le n energie degli oscillatori, cioè (rimuovendo la tilde) $E_i = \frac{1}{2}(p_i^2 + \omega_i^2 q_i^2)$, $i = 1, \dots, n$. La matrice di cui va verificato che abbia rango n è, ordinando qui per comodità le variabili $(p_1, q_1, \dots, p_n, q_n)$,

$$\frac{\partial(E_1, \dots, E_n)}{\partial(p_1, q_1, \dots, p_n, q_n)} = \begin{pmatrix} p_1 & \omega_1^2 q_1 & & & \\ & & p_2 & \omega_2^2 q_2 & \\ & & & & \dots \\ & & & & & p_n & \omega_n^2 q_n \end{pmatrix}.$$

Si ha allora una caduta di rango se uno o più oscillatori sono all’equilibrio: $p_i, q_i = 0$ per almeno un indice i . Al di fuori di questo caso il rango è n e corrispondentemente il moto avviene sul prodotto di n cerchi, cioè su un toro \mathbb{T}^n , ed è quasi periodico con n frequenze:

$$I_i(t) = I_i^o, \quad \varphi_i(t) = \varphi_i^o(t) + \omega_i t, \quad i = 1, \dots, n. \quad (3.13)$$

Invece in corrispondenza all’equilibrio di uno o più oscillatori una famiglia di tori n dimensionali collassa su un toro di dimensione inferiore (il moto è ancora quasi periodico, con un numero inferiore di frequenze).

¹⁴Il teorema di Liouville–Arnol’d è a volte usato, in letteratura, per definire i sistemi integrabili: i sistemi integrabili secondo Liouville sono in questo caso direttamente definiti come quei sistemi per cui H , assieme a $n - 1$ opportune funzioni F_i , soddisfa le ipotesi del teorema di Liouville–Arnol’d.

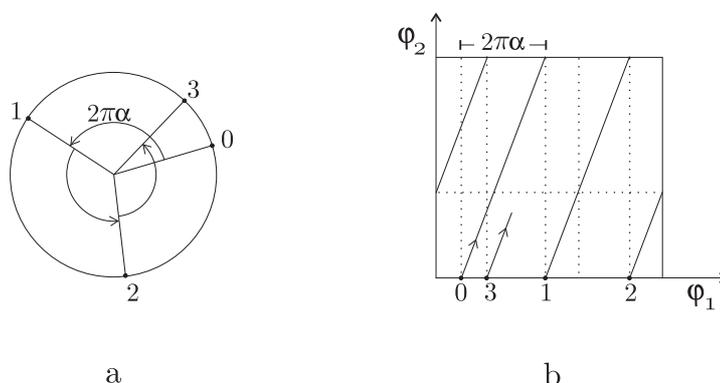


Figura 8: (a) la traslazione sul circolo; (b) il legame col moto lineare su \mathbb{T}^2 .

ESERCIZIO 29 Per $n = 2$ si consideri il moto quasi periodico di φ_1, φ_2 su \mathbb{T}^2 definito dalle (3.13). Si dimostri che (a) se $\alpha = \omega_1/\omega_2$ è razionale, allora tutti i moti sono periodici; (b) se invece α è irrazionale, la traiettoria è densa sul toro (Jacobi, 1835). La seconda parte non è immediata. Convieni prima risolvere l'esercizio seguente: si consideri la traslazione sul circolo $(0, 2\pi)$ definita da $\Phi(\varphi) = \varphi + 2\pi\alpha$ (figura 8a), e per ogni φ_0 si consideri la successione $\varphi_0, \varphi_1, \dots, \varphi_t, \dots$ definita da $\varphi_{t+1} = \Phi(\varphi_t)$, $t \geq 0$. Si dimostri che la successione è densa sul circolo se e solo se α è irrazionale. Anche questo però richiede una certa ingegnosità; si veda l'appendice E. Per comprendere il legame tra i due esercizi si osservi la figura 8b.

ESERCIZIO 30 Con le notazioni dell'esercizio precedente si consideri il caso di α razionale, $\alpha = k/l$. Si determini una ulteriore costante del moto del sistema, in aggiunta a I_1 e I_2 .

ESERCIZIO 31 Si consideri l'hamiltoniana lineare integrabile $H(I_1, I_2, \varphi_1, \varphi_2) = \omega_1 I_1 + \omega_2 I_2$. (i) Per $\omega_1 = \omega_2$ si trovi una trasformazione canonica $(I, \varphi) = w(J, \psi)$ tale che la nuova hamiltoniana dipenda solo da J_1 . (ii) Si estenda il risultato a ω_1/ω_2 razionale qualsiasi; si cerchi per questo w della forma (3.12). La seconda parte richiede un po' di ingegnosità (o qualche conoscenza di aritmetica); si veda l'appendice E.

ESERCIZIO 32 Si consideri un cambiamento di variabili del tipo

$$I = u(J), \quad \varphi = v(J, \psi). \quad (3.14)$$

Qual è la forma più generale possibile di u e v tale che la trasformazione sia canonica, con $v : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ biunivoca? [Risposta: è una marginale generalizzazione della (3.12), precisamente

$$I = A^{-T} J, \quad \varphi = A[\psi + a(J)],$$

con A intera con $|\det A| = 1$ e a tale che $\frac{\partial a_i}{\partial J_j} = \frac{\partial a_j}{\partial J_i}$.]

(ii) Il moto centrale piano, in particolare il problema di Keplero. Per questi sistemi l'hamiltoniana (supponendo la massa unitaria) ha la forma

$$H(p_r, p_\vartheta, r, \vartheta) = \frac{p_r^2}{2} + \frac{p_\vartheta^2}{2r^2} + V(r); \quad (3.15)$$

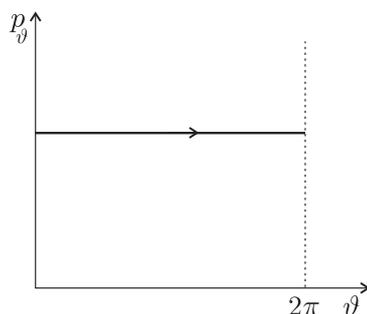


Figura 9: Il punto (ϑ, p_ϑ) corre su un circolo.

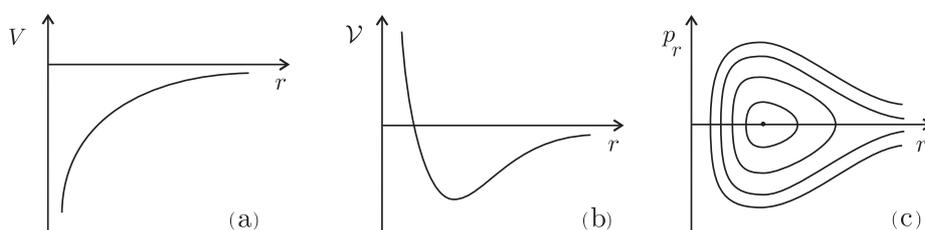


Figura 10: (a) Il potenziale attrattivo V ; (b) il corrispondente potenziale efficace $\mathcal{V}_{p_\vartheta}$; (c) il ritratto in fase di $\mathcal{H}_{p_\vartheta}$.

con evidenza si conservano e sono in involuzione l'energia H e il momento angolare p_ϑ . Il moto nel piano ϑp_ϑ è allora come in figura 9: p_ϑ resta costante, mentre ϑ , supponendo $p_\vartheta \neq 0$, avanza monotonamente ($\dot{\vartheta} = p_\vartheta/r^2 \neq 0$). Per quanto riguarda la coppia (p_r, r) , per ogni fissato p_ϑ essa è regolata dall'hamiltoniana efficace

$$\mathcal{H}_{p_\vartheta}(p_r, r) = \frac{p_r^2}{2} + \mathcal{V}_{p_\vartheta}(r), \quad \mathcal{V}_{p_\vartheta}(r) = \frac{p_\vartheta^2}{2r^2} + V(r).$$

Supponiamo V attrattivo della forma qualitativa riportata in figura 10a, con (eventuale) divergenza nell'origine più debole di $-1/r^2$ in modo che $\mathcal{V}_{p_\vartheta}$ abbia la forma qualitativa riportata in figura 10b e corrispondentemente il ritratto in fase di $\mathcal{H}_{p_\vartheta}$ sia come in figura 10c. Verifichiamo l'ipotesi che il rango della matrice $\frac{\partial(H, p_\vartheta)}{\partial(p_r, p_\vartheta, r, \vartheta)}$ sia due. Si ha

$$\frac{\partial(H, p_\vartheta)}{\partial(p_r, p_\vartheta, r, \vartheta)} = \begin{pmatrix} p_r & \frac{p_\vartheta}{r^2} & \frac{\partial \mathcal{V}_{p_\vartheta}}{\partial r} & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

pertanto si ha una caduta di rango se simultaneamente risulta $p_r = 0$ e $\frac{\partial \mathcal{V}_{p_\vartheta}}{\partial r} = 0$. Ciò si verifica sui punti di equilibrio del sistema ridotto $\mathcal{H}_{p_\vartheta}$, ovvero in corrispondenza ai moti circolari di H . Al di fuori di questo caso invece ogni moto di $\mathcal{H}_{p_\vartheta}$ (con energia negativa) avviene su una traiettoria chiusa, cioè su un circolo. Corrispondentemente, ogni moto di H avviene sul prodotto di questo circolo e del circolo considerato sopra nel piano ϑp_ϑ , in cui ϑ avanza monotonamente a p_ϑ costante. Il moto si svolge così su un toro bidimensionale.¹⁵ In corrispondenza ai moti circolari invece una

¹⁵Non è difficile vedere che per $E \geq 0$, venendo a mancare la compattezza delle curve in figura 10c, il moto avviene su un cilindro; si veda la nota 13, a proposito del teorema di Liouville–Arnol'd.

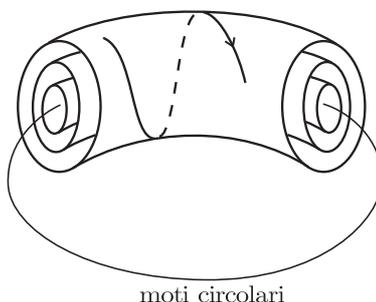


Figura 11: in corrispondenza ai moti circolari una famiglia di tori \mathbb{T}^2 collassa in un cerchio.

famiglia di tori bidimensionali collassa in un cerchio, e corrispondentemente un angolo non è più definito; si veda la figura 11.

I moti di (p_ϑ, ϑ) e (p_r, r) nei due cerchi hanno periodi T_ϑ e T_r dipendenti dalle costanti del moto E e p_ϑ , in generale diversi e incommensurabili: i moti non circolari allora in generale non sono chiusi e anzi (si veda sopra l’esercizio 29) riempiono densamente il toro cui appartengono. Proiettati nello spazio delle configurazioni essi danno il noto moto “a rosetta”, si veda la figura 12, in cui durante ogni periodo T_r l’angolo ϑ , misurato ad esempio al perielio, avanza di una quantità fissa Θ in generale non commensurabile a 2π . Fa eccezione il caso kepleriano $V = -k/r$, per il quale come è noto T_r e T_ϑ sono sempre uguali tra loro, perciò $\Theta = 0$ e tutte le traiettorie anche non circolari sono chiuse.

La descrizione dettagliata del passaggio alle variabili di azione–angolo per il moto centrale piano sarà fatta più avanti nel paragrafo 3.5C, tramite l’equazione di Hamilton–Jacobi; qui ci limitiamo a esporre il risultato. Una scelta possibile di queste variabili, di fatto la più comune e la più conveniente nelle applicazioni, è

$$\begin{aligned} I_1 &= \frac{1}{2\pi} \mathcal{A}(E, p_\vartheta) + |p_\vartheta| \\ I_2 &= p_\vartheta \\ \varphi_1 &= 2\pi \frac{\tau(r, p_r, E, p_\vartheta)}{T_r(E, p_\vartheta)} \\ \varphi_2 &= \psi + \Theta \frac{\tau(r, p_r, E, p_\vartheta)}{T_r(E, p_\vartheta)} \end{aligned}$$

ove $E = H(p_r, p_\vartheta, r, \vartheta)$; \mathcal{A} è l’area sottesa dalla curva passante per (r, p_r) nel piano rp_r (figura 13), a p_ϑ costante (\mathcal{A} dipende solo da p_ϑ e da E); τ è il tempo necessario a raggiungere (r, p_r) a partire dal perielio (figure 13 o anche 12); infine ψ è l’angolo che individua l’ultimo perielio a partire da una direzione assegnata (figura 12). La nuova hamiltoniana ha una forma che non si scrive in termini di funzioni elementari, comunque dipende solo da I_1 e I_2 ,

$$K(I_1, I_2, \varphi_1, \varphi_2) = h(I_1, I_2) .$$

L’angolo φ_1 , detto *anomalia media*, ha origine nell’ultimo perielio e avanza uniformemente fino a 2π al momento del passaggio al perielio successivo.¹⁶ L’angolo φ_2 , detto *argomento del perielio*, avanza invece uniformemente di Θ nel tempo intercorrente tra due passaggi successivi al perielio (resta definita, in questo modo, una posizione del perielio anche tra un passaggio e l’altro).

¹⁶Per la legge delle aree, φ_1 è dunque proporzionale all’area spazzata dal punto che esegue il moto centrale a partire dal perielio.

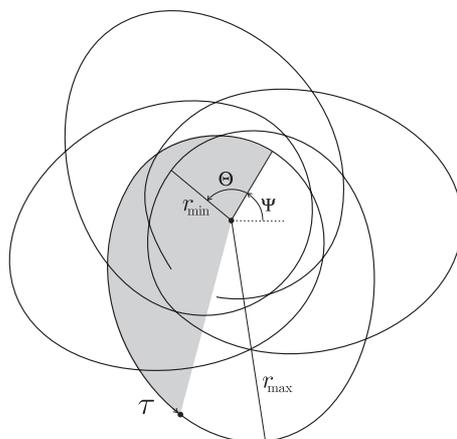


Figura 12: Una tipica traiettoria per un moto centrale

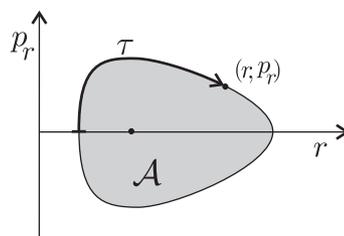


Figura 13: La variabili A e τ per il moto centrale.

Nel caso kepleriano, cioè per $V(r) = -k/r$, come è noto, si ha una costante del moto in più, che qui si traduce nel fatto che φ_2 resta costante: il perielio non avanza e tutti i moti sono periodici (anziché quasi periodici con due frequenze). L'hamiltoniana h nel caso di Keplero si scrive esplicitamente; essa risulta dipendere dalla sola I_1 , e a conti fatti si trova

$$h(I_1) = -\frac{k^2}{2I_1^2} . \tag{3.16}$$

Come poi è noto dalla meccanica newtoniana, il semiasse maggiore a delle ellissi di Keplero è funzione della sola energia h , precisamente si ha $a = -k/(2h)$; sostituendo h nella (3.16) si trova allora il legame tra l'azione I_1 e il semiasse,

$$I_1 = \sqrt{ka} .$$

Il risultato è coerente con la terza legge di Keplero: infatti dall'espressione (3.16) di h si trova

$$\omega_1(I_1) = \frac{\partial h}{\partial I_1} = \frac{k^2}{I_1^3} = \frac{k^{1/2}}{a^{3/2}} ,$$

da cui segue

$$T^2 = \frac{(2\pi)^2}{k} a^3 .$$

ESERCIZIO 33 Si verifichi che nel caso kepleriano $V = -k/r$ il vettore

$$A = p \times l - kq/r , \quad l = q \times p$$

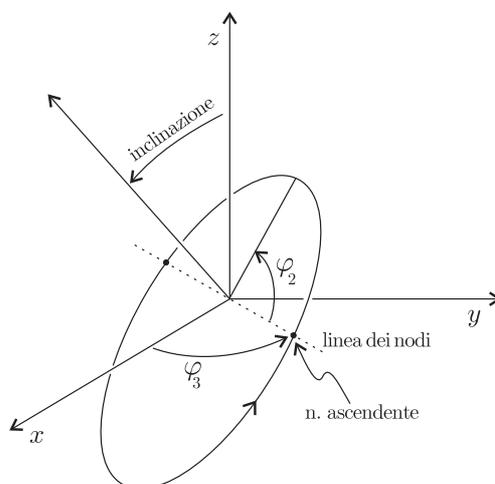


Figura 14: Il problema di Keplero visto in tre dimensioni.

(“vettore di Runge–Lenz”) è costante nel tempo. (Si può calcolare \dot{A} lungo il flusso hamiltoniano, o equivalentemente calcolare algebricamente le tre parentesi di Poisson $\{A_i, H, \}$). Si osservi che la direzione del vettore è parallela al semiasse maggiore delle ellissi. (Infatti, $A \cdot p = 0$ al perielio, cioè quando $q \cdot p = 0$.)

Alla luce del caso kepleriano diamo la seguente definizione:

DEFINIZIONE 7 *Sistemi integrabili in cui una o più azioni non compaiono nell’hamiltoniana, mentre corrispondentemente gli angoli ad esse coniugati restano fermi, sono detti degeneri, e i gradi di libertà assenti da h sono detti essi stessi degeneri.*

Il problema di Keplero piano dunque è degenere. I sistemi degeneri hanno così un numero n_0 di frequenze (non nulle) inferiore a n e corrispondentemente tutti¹⁷ i moti sono ristretti a tori di dimensione n_0 (dimensione uno, e dunque orbite chiuse con moti tutti periodici, per Keplero). Il numero di costanti del moto viceversa è più alto, $2n - n_0 > n$, e per questo i sistemi degeneri sono anche detti *superintegrabili*.

ESERCIZIO 34 *Si studi l’oscillatore armonico piano, con hamiltoniana (3.15) e $V(r) = \frac{1}{2}\omega^2 r^2$. (Al pari del problema di Keplero, il sistema è degenere e ha tutti i moti periodici. Il teorema di Bertrand assicura che non vi sono altri casi di moti centrali con questa proprietà.)*

ESERCIZIO 35 *Per $n = 2$ si consideri il sistema integrabile lineare di hamiltoniana $h(I_1, I_2) = \omega_1 I_1 + \omega_2 I_2$. Si mostri che se ω_2/ω_1 è razionale, allora il sistema con una trasformazione canonica del tipo (3.12) prende la forma degenere. [Suggerimento: ci si rifaccia all’esercizio 31. Per $\omega_1 = \omega_2$ si ha l’oscillatore armonico piano.]*

(iii) *Il moto centrale come problema tridimensionale* (cenno). Possiamo guardare a un generico moto centrale non più nel piano ma nello spazio. Il sistema è integrabile: precisamente, detto

¹⁷Non ci si confonda col fatto che per particolari dati iniziali, ad esempio in corrispondenza ai moti circolari nel problema del moto centrale, un toro \mathbb{T}^n può degenerare in un toro di dimensione inferiore.

M il momento angolare, si conservano e sono indipendenti e in involuzione H , M^2 (o il modulo $\|M\|$) e una qualunque componente M_z di M , con scelta arbitraria dell'asse z .¹⁸ Il sistema è superintegrabile, o degenerare, qualunque sia il potenziale: ha infatti una costante del moto in più, ad esempio $M_{z'}$ con $z' \neq z$ qualsiasi; corrispondentemente ha due anziché tre frequenze. Il caso kepleriano a sua volta si presenta doppiamente degenerare, con una sola frequenza.

Il moto di fatto è piano, e mettendoci nel piano dell'orbita possiamo ripetere la trattazione precedente (si osservi che $p_\vartheta = \|M\|$), costruendo in modo identico $I_1, I_2, \varphi_1, \varphi_2$. Come terza azione si può prendere $I_3 = M_z$; I_3 determina così l'inclinazione del piano dell'orbita (evidentemente costante) rispetto al piano xy . L'angolo φ_3 coniugato a I_3 risulta essere l'angolo nel piano xy da una direzione prefissata, ad esempio l'asse x , alla "linea dei nodi" in cui il piano dell'orbita interseca il piano xy . Il piano dell'orbita è fisso, e corrispondentemente φ_3 non avanza; I_3 non compare in H — l'inclinazione dell'orbita certo non cambia l'energia — coerentemente con il fatto che il sistema è degenerare. La figura 14 illustra il caso kepleriano.

3.4 Il corpo rigido di Eulero come sistema hamiltoniano integrabile

Il corpo rigido di Eulero è il corpo rigido con un punto fisso in assenza di forze attive. In questo paragrafo ci proponiamo di descriverlo come sistema hamiltoniano integrabile, e in particolare di introdurre le sue coordinate di azione–angolo; indispensabile premessa è la trattazione classica di questo sistema, richiamata nell'appendice G.

Denotiamo con (O, e_x, e_y, e_z) un sistema di riferimento fisso nello spazio, con origine nel punto fisso del corpo, e con (O, e_1, e_2, e_3) un sistema di riferimento solidale al corpo e principale (cioè con gli assi cartesiani paralleli agli assi principali di inerzia relativi a O). La configurazione del corpo è univocamente individuata da una matrice $R \in SO(3)$, ad esempio la matrice tale che $e_1 = Re_x$ e così via; la varietà delle configurazioni del sistema dunque è il gruppo $SO(3)$. Lo stato del sistema è poi univocamente determinato assegnando, oltre a R , il momento angolare m . Quest'ultimo si può pensare come vettore di \mathbb{R}^3 facendo riferimento, equivalentemente, alla terna di componenti $m^c = (m_1, m_2, m_3)$ nella base propria oppure a alla terna $m^s = (m_x, m_y, m_z)$ nella base fissa nello spazio; la relazione tra le due è $m^s = Rm^c$. Per essere definiti, e come risulta di fatto più conveniente, faremo riferimento a m^c , pertanto identificheremo gli stati del sistema con le coppie

$$(R, m^c) \in SO(3) \times \mathbb{R}^3 .$$

Lo spazio delle fasi risulta così essere la varietà prodotto¹⁹

$$\mathcal{M} = \mathbb{R}^3 \times SO(3) .$$

Nel sistema si conserva l'energia (cinetica)

$$K(m^c) = \frac{m_1^2}{2A_1} + \frac{m_2^2}{2A_2} + \frac{m_3^2}{2A_3} , \quad (3.17)$$

¹⁸Si conservano anche, ma non sono in involuzione, le tre componenti del momento angolare M_x, M_y, M_z .

¹⁹Che lo spazio delle fasi abbia globalmente la struttura di prodotto è un fatto eccezionale. In genere, se \mathcal{Q} è la varietà delle configurazioni di un sistema olonomo a n gradi di libertà, la meccanica lagrangiana e la meccanica hamiltoniana sono definite su varietà che solo localmente hanno la struttura di prodotto $\mathcal{U} \times \mathbb{R}^n$, ove \mathcal{U} è un aperto di \mathcal{Q} ; globalmente si ottengono invece varietà più complesse, dette *fibrato tangente* e *fibrato cotangente* rispettivamente per la meccanica Lagrangiana e per quella Hamiltoniana. Per il corpo rigido si ottiene la struttura globale di prodotto grazie al fatto eccezionale che lo spazio delle configurazioni $SO(3)$ è un gruppo.

ove si sono denotati con A_1, A_2, A_3 i momenti di inerzia relativi alla base propria. Oltre a K si conservano le tre componenti di m nello spazio m_x, m_y e m_z , dunque in tutto le quantità conservate sono quattro. Una terna di costanti del moto in involuzione è (K, m^2, m_z) , ma evidentemente al posto di m_z si può prendere una qualunque altra componente $m_{z'}$ di m in una direzione fissa z' dello spazio. Il sistema dunque è integrabile, ma ha una costante del moto in più ovvero, secondo la terminologia sopra introdotta, è superintegrabile (si osservi la forte somiglianza con il caso del moto centrale in tre dimensioni).

A. Il caso simmetrico e le coordinate di Andoyer.

Ci restringiamo ora per semplicità al caso di corpo rigido simmetrico: $A_1 = A_2 \neq A_3$. Risulta allora

$$K(m^c) = \frac{1}{2A_1}(m_1^2 + m_2^2) + \frac{1}{2A_3} m_3^2 = \frac{1}{2A_1}(m^2 + \eta m_3^2), \quad (3.18)$$

ove η è il “fattore di forma”

$$\eta = \frac{A_1 - A_3}{A_3}.$$

Poichè si ha sempre $0 < A_3 \leq A_1 + A_2 = 2A_1$, risulta

$$-\frac{1}{2} \leq \eta < \infty;$$

$\eta = -1/2$ corrisponde a $A_3 = 2A_1$, cioè a un corpo piatto (una situazione limite che non presenta difficoltà), mentre per $\eta = \infty$ si avrebbe un corpo filiforme, $A_3 = 0$: una situazione singolare che escludiamo dalle nostre considerazioni. Come si vede dalla (3.18), al posto di K si può equivalentemente usare la quantità conservata m_3 , e far così riferimento alle costanti del moto

$$\|m\|, \quad m_3, \quad m_z, \quad m_{z'};$$

le prime tre, ma anche le prime due e l'ultima, sono in involuzione (invece $\{m_z, m_{z'}\} \neq 0$).

La descrizione buona del moto, adatta a mettere in evidenza la struttura integrabile del sistema, è quella suggerita da Poinsot, che qui ricordiamo brevemente (si veda la figura 15a).

- Il momento angolare m è costante nello spazio.
- La componente m_3 , e dunque l'angolo tra m e e_3 , sono costanti.
- I vettori m, Ω e e_3 , ove Ω denota la velocità angolare del corpo (appendice G) sono complanari, dunque

$$\Omega = \Omega' e_m + \Omega'' e_3 \quad \text{ove} \quad e_m = m/\|m\|; \quad (3.19)$$

con semplici considerazioni geometriche si trova

$$\Omega' = \frac{\|m\|}{A_1}, \quad \Omega'' = \eta \frac{m_3}{A_1}.$$

- Il piano di m, Ω, e_3 precede attorno a m con velocità angolare Ω' costante; simultaneamente il corpo ruota attorno a e_3 con velocità angolare Ω'' costante. Per descrivere questi due moti è utile tracciare due coni \mathcal{C}' e \mathcal{C}'' (coni di Poinsot) con vertice comune nell'origine, il primo con asse m , fisso nello spazio, e il secondo con asse e_3 , solidale al corpo; i coni si toccano sulla direttrice comune Ω . I punti del corpo su questa retta, che è asse istantaneo di rotazione, hanno velocità nulla, perciò il cono solidale \mathcal{C}'' rotola senza strisciare su \mathcal{C}' fisso, come se ci fosse un vincolo di puro rotolamento tra due oggetti materiali.

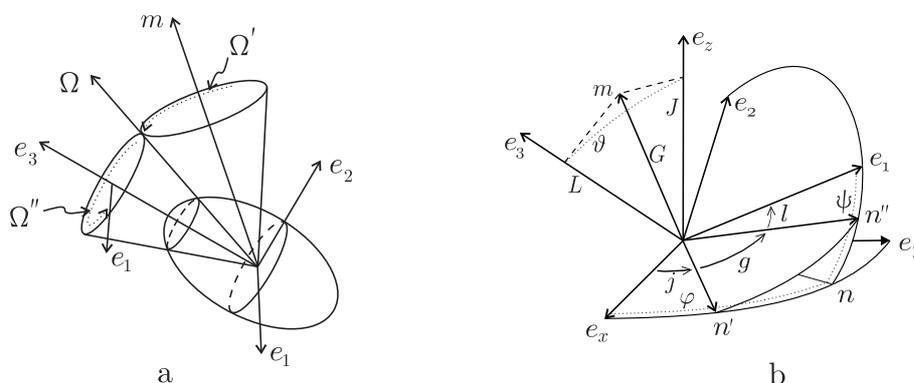


Figura 15: (a) La descrizione di Poinsot; (b) le coordinate di Andoyer (per confronto sono stati indicati, con tratto punteggiato, i consueti angoli di Eulero ϑ, φ, ψ e la linea dei nodi $n = e_z \times e_3$ che entra nella loro costruzione).

La descrizione ha senso solo se m e e_3 non sono paralleli, ovvero al di fuori delle rotazioni proprie sulle quali tutto diviene singolare. Qualche dettaglio in più si trova nell'appendice G.

Non è difficile introdurre sei coordinate adattate alla descrizione del moto di Eulero–Poinsot, che descrivano completamente lo stato di moto del corpo (si veda la figura 15b).

- Poniamo innanzitutto $G = \|m\|$, $L = m_3$. Queste quantità sono costanti e determinano l'apertura dei coni e le frequenze delle precessioni su di essi. Sono quantità intrinseche, cioè definite indipendentemente dalla scelta (arbitraria) della terna fissa (e_x, e_y, e_z) .
- Con una coppia di coordinate fissiamo l'orientamento di e_m nello spazio. Una scelta conveniente è (J, j) , con $J = m_z$ e $j =$ angolo nel piano xy da e_x alla linea dei nodi $n' = e_z \times m$ (coordinate cilindriche di e_m nella base fissa). Le coordinate (J, j) dipendono evidentemente dalla scelta della terna fissa, e sono ben definite se e_z e m non sono paralleli, $e_z \times m \neq 0$.
- Infine, con due angoli determiniamo l'avanzamento del sistema lungo i coni di Poinsot. Un primo angolo, che denoteremo g , descrive la precessione di Ω e e_3 attorno a m , e precisamente è l'angolo, nel piano ortogonale a m , dalla linea dei nodi n' a una seconda linea dei nodi $n'' = m \times e_3$; l'angolo g dipende (solo) per l'origine n' dalla scelta della terna fissa. Il secondo angolo, che denoteremo l , descrive la rotazione propria, ed è definito come l'angolo nel piano equatoriale e_1e_2 da n'' a e_1 ; la coordinata l così definita è intrinseca. Entrambi gli angoli sono ben definiti se m e e_3 non sono paralleli, $m \times e_3 \neq 0$. Il confronto con la figura 15a mostra che g è l'angolo tratteggiato sul cono \mathcal{C}' , solo ne è stata precisata l'origine; l differisce di $\pi/2$ dall'angolo tratteggiato sul secondo cono.

Si raccomanda di rifare da sé una o due volte la figura 15b. Si osservi che g, l , e j sono angoli nei piani ortogonali, rispettivamente, a m, e_3 e e_z ; corrispondentemente G, L e J sono le componenti del momento angolare lungo le tre direzioni.

E' laborioso, ma non veramente difficile, verificare che (G, L, J, g, l, j) sono effettivamente delle buone coordinate (una carta di $SO(3) \times \mathbb{R}^3$), se si escludono: (i) le rotazioni proprie $m \times e_3 = 0$ (ovvero $L = \pm G$), ove i coni di Poinsot sono singolari; (ii) inoltre, per evitare la singolarità delle coordinate cilindriche J, j , il caso $e_z \times m = 0$. Ma quest'ultima non è una vera limitazione, data

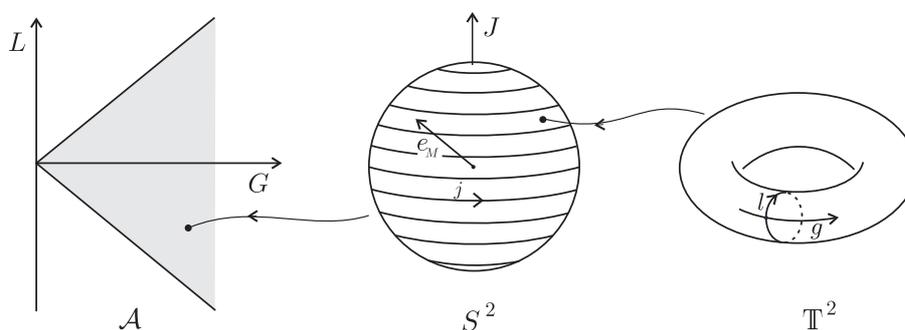


Figura 16: Lo spazio delle fasi del corpo rigido.

l'arbitrarietà della scelta dell'asse z . Quanto al dominio delle diverse coordinate, per (G, L) il dominio è

$$\mathcal{A} = \{(G, L) \in \mathbb{R}^2 : G > 0, -G < L < G\};$$

il dominio di e_m è la sfera S^2 , mentre il dominio delle coordinate (J, j) su S^2 è $(-G, G) \times S^1$ (si escludono i due “poli” ove l'asse z buca la sfera); infine $(g, l) \in \mathbb{T}^2$. Si veda la figura 16: in ogni punto di \mathcal{A} è attaccata una sfera S , con coordinate (J, j) ; in ogni punto della sfera è attaccato un toro bidimensionale. Questo è lo spazio delle fasi del sistema, se si escludono le rotazioni proprie.²⁰

L'interesse principale per le coordinate (G, L, J, g, l, j) è che, prese in quest'ordine, esse costituiscono coordinate canoniche, nelle quali le lettere minuscole denotano le coordinate e le corrispondenti lettere maiuscole denotano i rispettivi momenti coniugati. Vale infatti la seguente importante proposizione, dovuta a Andoyer (1923), in seguito “riscoperta” da Deprit (1961) e chiamata di solito teorema di Andoyer–Deprit:

PROPOSIZIONE 17 *Le coordinate (G, L, J, g, l, j) , ordinate come indicato sopra e nel dominio indicato sopra, sono coordinate canoniche analitiche su $SO(3) \times \mathbb{R}^3$. Due carte di queste coordinate, relative a diverse scelte di e_z , coprono la varietà*

$$\mathcal{M}_0 = \{(R, m) \in SO(3) \times \mathbb{R}^3 : m \times e_3 \neq 0\}.$$

L'hamiltoniana del corpo rigido simmetrico in queste coordinate è $H(G, L, J, g, l, j) = h(G, L)$, con

$$h(G, L) = \frac{1}{2A_1}(G^2 + \eta L^2). \quad (3.20)$$

Omettiamo la dimostrazione, ancora una volta non difficile ma faticosa. Ammessa tuttavia la canonicità, la forma (3.20) di H nel caso simmetrico segue immediatamente dalla (3.18): H è infatti l'energia, qui soltanto cinetica, scritta in coordinate canoniche. Si vede così che *le coordinate di Andoyer sono le coordinate di azione–angolo del corpo rigido simmetrico*. Il sistema è con evidenza degenere, e il grado di libertà degenere è (J, j) . Il moto si svolge sul toro bidimensionale \mathbb{T}^2 , con frequenza $\omega = (\omega_g, \omega_l)$ data da

$$\omega = \left(\frac{\partial K}{\partial G}, \frac{\partial K}{\partial L} \right) = \frac{1}{A_1}(G, \eta L);$$

²⁰E' interessante osservare che due delle coordinate di Andoyer, precisamente L e J , coincidono con due dei momenti coniugati delle coordinate di Eulero: si ha infatti $L = p_\psi$, $J = p_\varphi$; si veda l'appendice J.

si osservi che ω_g, ω_l coincidono con le velocità Ω', Ω'' delle precessioni sui coni di Poinot riportate nella (3.19).

B. Il caso triassiale (cenno).

Nel caso triassiale le coordinate di Andoyer (G, L, J, g, l, j) restano, evidentemente, coordinate canoniche su $SO(3)$, dal momento che la canonicità delle coordinate è una proprietà a priori, indipendente dall'hamiltoniana cui si è interessati. Ma non sono più coordinate di azione–angolo: infatti, questa volta (figura 15b) si ha

$$m_1 = \sqrt{m^2 - m_3^2} \sin l, \quad m_2 = \sqrt{m^2 - m_3^2} \cos l,$$

e dalla (3.17) si deduce allora

$$K = (G^2 - L^2) \left(\frac{\sin^2 l}{2A_1} + \frac{\cos^2 l}{2A_2} \right) + \frac{L^2}{2A_3}.$$

Ma questa espressione si può interpretare come hamiltoniana di un sistema a un solo grado di libertà, per la coppia (L, l) — una variante del pendolo — in cui G è un parametro. Procedendo allora come per il pendolo (o meglio come per un qualunque sistema a un grado di libertà) non è difficile comprendere che si può introdurre, al posto di (L, l) , una coppia di coordinate di azione–angolo (I, φ) , nell'intorno di ogni curva di livello di K non contenente punti singolari. Qualche dettaglio in più è demandato al prossimo esercizio.

ESERCIZIO 36 *Si tracci il ritratto in fase relativo all'hamiltoniana K (pensata come funzione di L, l). In quali regioni del piano Ll si possono introdurre le coordinate di azione–angolo I, φ ? Chi è I ?*

3.5 L'equazione di Hamilton–Jacobi

La ricerca di una trasformazione di coordinate $(p, q) = w(\tilde{p}, \tilde{q})$ che porti un sistema hamiltoniano in forma integrabile è naturalmente associata a una importante equazione alle derivate parziali, detta *equazione di Hamilton–Jacobi*. Si tratta di un'equazione che interviene in vario modo nella Fisica Matematica. Hamilton la introdusse originariamente in connessione alla propagazione dei fronti d'onda in ottica, mentre fu Jacobi a sottolinearne l'importanza in relazione all'integrabilità dei sistemi hamiltoniani. Ci limiteremo qui alla cosiddetta *equazione ridotta* di Hamilton–Jacobi, adatta a problemi indipendenti dal tempo, accennando solo brevemente alla fine al caso generale.

A. L'equazione ridotta di Hamilton–Jacobi

Cominciamo con una definizione:

DEFINIZIONE 8 *Data la funzione di Hamilton $H(p, q)$, $H : \mathbb{R}^{2n} \rightarrow \mathbb{R}$, l'equazione alle derivate parziali*

$$H\left(\frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_n}, q_1, \dots, q_n\right) = h, \quad (3.21)$$

nella quale sia la funzione $S(q_1, \dots, q_n)$ sia la costante h sono incognite, è detta “equazione ridotta di Hamilton–Jacobi” associata ad H .

Si osservi che ciascuna derivata parziale $\frac{\partial S}{\partial q_i}$ sta in H al posto di p_i . Ad esempio, all'hamiltoniana dell'oscillatore armonico $H(p, q) = \frac{1}{2}(p^2 + \omega^2 q^2)$ è associata l'equazione di Hamilton–Jacobi

$$\frac{1}{2} \left(\frac{\partial S}{\partial q} \right)^2 + \frac{1}{2} \omega^2 q^2 = h . \quad (3.22)$$

Non è difficile verificare che una famiglia di soluzioni di questa particolare equazione è data da

$$S(\alpha, q) = \pm \int_0^q \sqrt{2\alpha\omega - \omega^2 x^2} dx , \quad h(\alpha) = \alpha\omega , \quad (3.23)$$

ove α è una arbitraria costante positiva.

DEFINIZIONE 9 *Si dice integrale completo dell'equazione ridotta di Hamilton–Jacobi ogni famiglia di soluzioni*

$$S(\alpha_1, \dots, \alpha_n, q_1, \dots, q_n) , \quad h(\alpha_1, \dots, \alpha_n)$$

dipendente da n parametri reali $\alpha_1, \dots, \alpha_n$, tale che risulti

$$\det \left(\frac{\partial^2 S}{\partial \alpha_i \partial q_j} \right) \neq 0 . \quad (3.24)$$

Attenzione: l'integrale *completo* è una classe “sufficientemente ampia” di soluzioni e non va confuso con l'integrale *generale*, cioè con l'insieme di tutte le soluzioni (che per le equazioni alle derivate parziali contiene sempre funzioni arbitrarie anziché un numero finito di costanti arbitrarie; si pensi all'equazione delle onde). La (3.23) fornisce un integrale completo della (3.22).

Il legame tra l'equazione ridotta di Hamilton–Jacobi e l'integrabilità di H è espresso dalla seguente facile proposizione:

PROPOSIZIONE 18 *Sia $S(\alpha_1, \dots, \alpha_n, q_1, \dots, q_n)$, $h(\alpha_1, \dots, \alpha_n)$ un integrale completo dell'equazione ridotta di Hamilton–Jacobi (3.21). Allora la funzione generatrice $S(\tilde{p}_1, \dots, \tilde{p}_n, q_1, \dots, q_n)$, nella quale le costanti $\alpha_1, \dots, \alpha_n$ dell'integrale completo sono reinterpretate come momenti, genera implicitamente, tramite le (2.23), una trasformazione canonica $(p, q) = w(\tilde{p}, \tilde{q})$ che muta H in*

$$\tilde{H}(\tilde{p}_1, \dots, \tilde{p}_n, \tilde{q}_1, \dots, \tilde{q}_n) = h(\tilde{p}_1, \dots, \tilde{p}_n) . \quad (3.25)$$

DIMOSTRAZIONE. Dalle (2.23) segue

$$\tilde{H}(\tilde{p}, \tilde{q}) = H \left(\frac{\partial S}{\partial q}(\tilde{p}, q), q \right) ,$$

ove a secondo membro, in linea di principio, è sottintesa la sostituzione $q = v(\tilde{p}, \tilde{q})$, con v definita implicitamente dalle (2.23). Ma per l'equazione di Hamilton–Jacobi, il secondo membro non dipende da q e vale $h(\tilde{p})$, ovvero la sostituzione è inutile e si ha la (3.25). La condizione di invertibilità (2.22) è garantita dalla (3.24). \square

Restando nell'esempio della soluzione (3.23) alla (3.22), prendendo il segno più e denotando con I anziché con α il parametro (il nuovo momento), si ha la funzione generatrice

$$S(I, q) = \int_0^q \sqrt{2\omega I - \omega^2 x^2} dx ,$$

e questa (esercizio (10-v)) si è già vista essere la generatrice del passaggio alle coordinate di azione–angolo: la forma implicita della trasformazione è infatti

$$p = \frac{\partial S}{\partial q} = \sqrt{2\omega I - \omega^2 q^2}, \quad \varphi = \frac{\partial S}{\partial I} = \int_0^q \frac{dx}{\sqrt{2I/\omega - x^2}} = \arcsin \sqrt{\omega/2I} q,$$

e per inversione si ottengono facilmente le equazioni esplicite

$$q = \sqrt{2I/\omega} \sin \varphi, \quad p = \sqrt{2\omega I} \cos \varphi, \quad (3.26)$$

con determinazione dell'angolo (funzione arcoseno) tra $-\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2}$. Prendendo il segno meno si ottiene similmente

$$q = -\sqrt{2I/\omega} \sin \varphi', \quad p = -\sqrt{2\omega I} \cos \varphi'$$

con $-\frac{\pi}{2} \leq \varphi' \leq \frac{\pi}{2}$, e con la traslazione $\varphi' = \pi + \varphi$ si recuperano le (3.26), questa volta per $\frac{\pi}{2} \leq \varphi \leq \frac{3\pi}{2}$.

In conclusione, il dominio della trasformazione è $I > 0$, $\varphi \in S^1$, e l'hamiltoniana è mutata nell'hamiltoniana integrabile $\tilde{H}(I, \varphi) = h(I) = \omega I$.

All'integrale completo (3.23) della (3.22) si arriva facilmente scrivendo

$$\frac{\partial S}{\partial q} = \pm \sqrt{2h - \omega^2 q^2},$$

prendendo poi $h(\alpha) = \alpha\omega$; una quadratura dà allora banalmente la (3.23). La scelta $h(\alpha) = \alpha\omega$, si osservi, dà al parametro α le dimensioni di un'azione. Altre scelte di $h(\alpha)$ sono possibili, in particolare si può prendere h stessa come costante libera; in questo caso al posto delle variabili di azione–angolo sopra costruite si ottengono le variabili energia–tempo (h, τ) . Vale la pena di vederlo in dettaglio nel caso più generale di un sistema eventualmente anisocrono, come il pendolo.

B. Esempio: le variabili di energia–tempo per il pendolo.

Riprendiamo il pendolo (o un qualunque altro sistema con un grado di libertà che esegua moti periodici). L'equazione di Hamilton–Jacobi si scrive

$$\frac{1}{2} \left(\frac{\partial S}{\partial q} \right)^2 + V(q) = h,$$

da cui segue

$$\frac{\partial S}{\partial q} = \pm \sqrt{2(h - V(q))}.$$

Ci limitiamo qui al segno più, ma tutto si estende come indicato sopra anche al segno meno. Come parametro α prendiamo h stessa, e vediamo che allora l'equazione di Hamilton–Jacobi fornisce proprio le variabili energia–tempo (h, τ) del pendolo introdotte nel paragrafo 3.2. La generatrice $S(h, q)$, con questa scelta, prende la forma

$$S(h, q) = \int_0^q \sqrt{2(h - V(x))} dx;$$

la condizione di invertibilità (2.22) è soddisfatta,²¹

$$\frac{\partial^2 S}{\partial h \partial q} = \frac{1}{\sqrt{2(h - V(q))}} \neq 0,$$

²¹Sui punti di inversione (i punti q tali che $V(q) = h$, cosicché $\dot{q} = p = 0$) c'è un'evidente singolarità, ma i punti si recuperano poi per continuità.

e la trasformazione (in forma implicita) si scrive

$$p = \frac{\partial S}{\partial q} = \sqrt{2(h - V(q))} , \quad \tau = \frac{\partial S}{\partial h} = \int_0^q \frac{dx}{\sqrt{2(h - V(x))}} .$$

Per inversione nella prima e sostituzione nella seconda si ottengono le espressioni esplicite di h e τ in funzione di p e q :

$$h = \frac{p^2}{2} + V(q) = H(p, q) , \quad \tau = \int_0^q \frac{dx}{\sqrt{2(H(p, q) - V(x))}} ,$$

e queste sono proprio le coordinate energia–tempo precedentemente introdotte. Da queste, come si è visto sopra, si passa poi all'occorrenza alle coordinate di azione–angolo.

C. Esempio: il moto centrale piano.

Consideriamo il moto centrale piano già brevemente trattato nel paragrafo 3.3; l'hamiltoniana $H(p_r, p_\vartheta, r, \vartheta)$, di fatto indipendente da ϑ , è la (3.15), con V e \mathcal{V} come in figura 10, e l'equazione ridotta di Hamilton–Jacobi si scrive

$$\frac{1}{2} \left(\frac{\partial S}{\partial r} \right)^2 + \frac{1}{2r^2} \left(\frac{\partial S}{\partial \vartheta} \right)^2 + V(r) = h .$$

Ci proponiamo di risolvere questa equazione per costruire coordinate (h, l, τ, ψ) , in cui h e τ sono energia e tempo, l è il momento angolare p_ϑ , mentre ψ è una opportuna variabile angolare coniugata a l . In un secondo momento passeremo da queste variabili alle variabili di azione–angolo che stiamo cercando.

Cerchiamo allora un integrale completo dell'equazione, che sia lineare in ϑ , precisamente sia della forma²²

$$S(h, l, r, \vartheta) = \mathcal{S}(h, l, r) + l\vartheta , \quad (3.27)$$

ove h e l sono le costanti da cui l'integrale completo dipende, ovvero i nuovi momenti; questa forma di S assicura che sia, come vogliamo, $p_\vartheta = \frac{\partial S}{\partial \vartheta} = l$. L'equazione di Hamilton–Jacobi si riduce allora a

$$\frac{1}{2} \left(\frac{\partial \mathcal{S}}{\partial r} \right)^2 + \frac{l^2}{2r^2} + V(r) = h ,$$

e in questa equazione si riconosce l'equazione di Hamilton–Jacobi del problema a un grado di libertà

$$\mathcal{H}(p_r, r, l) = \frac{p_r^2}{2} + \mathcal{V}(r, l) , \quad \mathcal{V}(r, l) = \frac{l^2}{2r^2} + V(r) ,$$

che si ottiene dal problema di partenza di Hamiltoniana (3.15) fissando il valore l del momento angolare p_ϑ . Possiamo allora procedere come abbiamo fatto sopra per il pendolo: ricaviamo

$$\frac{\partial \mathcal{S}}{\partial r} = \pm \sqrt{2(h - \mathcal{V}(r, l))} ,$$

²²Questa forma di S si presenta spontanea se si segue il metodo, tipico in diverse equazioni alle derivate parziali, della *separazione delle variabili*. Precisamente si cerca S della forma $S = \mathcal{S}(r) + S'(\vartheta)$ (si separa la dipendenza da r e da ϑ), cosicché l'equazione di Hamilton–Jacobi si scrive

$$\frac{1}{2} \left(\frac{\partial \mathcal{S}}{\partial r} \right)^2 + \frac{1}{2r^2} \left(\frac{\partial S'}{\partial \vartheta} \right)^2 + V(r) = h .$$

Si osserva subito che il secondo termine deve essere indipendente da ϑ (tutti gli altri lo sono), ovvero S' deve essere lineare in ϑ , $S' = l\vartheta$; il coefficiente di proporzionalità l si usa come costante dell'integrale completo. Si giunge così alla (3.27), ove si sono inseriti tra gli argomenti di S anche le costanti h e l .

e limitandoci per semplicità al segno '+':

$$S(h, l, r) = \int_{r_{\min}(h, l)}^r \sqrt{2(h - \mathcal{V}(x, l))} \, dx .$$

Questa generatrice produce la trasformazione canonica, scritta in variabili miste,

$$p_r = \frac{\partial S}{\partial r} = \sqrt{2(h - \mathcal{V}(r, l))} , \quad \tau = \frac{\partial S}{\partial h} = \int_{r_{\min}(h, l)}^r \frac{dr'}{\sqrt{2(h - \mathcal{V}(r', l))}} , \quad (3.28)$$

cui si aggiungono, riprendendo S completa e denotando con ψ la variabile angolare coniugata a l , le equazioni

$$p_\vartheta = \frac{\partial S}{\partial \vartheta} = l , \quad \psi = \frac{\partial S}{\partial l} = \vartheta - \int_{r_{\min}}^r \frac{\frac{\partial \mathcal{V}}{\partial l}(r', l)}{\sqrt{2(h - \mathcal{V}(r', l))}} \, dr' . \quad (3.29)$$

Le espressioni (3.28) e (3.29) si invertono senza difficoltà, fornendo h , l , τ e ψ in funzione di p_r , p_ϑ , r e ϑ :

$$\begin{aligned} h &= \frac{p_r^2}{2} + \mathcal{V}(r, p_\vartheta) = H(p_r, p_\vartheta, r, \vartheta) \\ l &= p_\vartheta \\ \tau &= \int_{r_{\min}(h, p_\vartheta)}^r \frac{dr'}{\sqrt{2(h - \mathcal{V}(r', p_\vartheta))}} \\ \psi &= \vartheta - \int_{r_{\min}(h, p_\vartheta)}^r \frac{\frac{\partial \mathcal{V}}{\partial l}(r', p_\vartheta)}{\sqrt{2(h - \mathcal{V}(r', p_\vartheta))}} \, dr' , \end{aligned}$$

ove si intende che h , nelle ultime due, è sostituito da $H(p_r, p_\vartheta, r, \vartheta)$. (La dipendenza di H da ϑ è puramente formale, di fatto H non ne dipende.) Corrispondentemente si ha

$$\tilde{H}(h, l, \tau, \psi) = h .$$

Un attimo di riflessione mostra che l'integrale nell'espressione per ψ è l'avanzamento $\vartheta - \vartheta_0$ dell'angolo ϑ nel tempo in cui il raggio passa dal valore r_{\min} del perielio a r : si ha infatti $\frac{\partial \mathcal{V}}{\partial l} = \dot{\vartheta}$, $\sqrt{2h - V} = p_r = \dot{r}$, da cui segue subito

$$\int_{r_{\min}}^r \frac{\dot{\vartheta}}{\dot{r}} \, dr' = \int_{t_0}^t \dot{\vartheta} \, dt = \vartheta(t) - \vartheta_0 ;$$

si è denotato con t_0 l'istante del passaggio al perielio ($r(t_0) = r_{\min}$) e con $\vartheta_0 = \vartheta(t_0)$ il corrispondente valore di ϑ . Si ha così

$$\psi = \vartheta_0 ,$$

costante nel tempo: come d'altra parte deve essere se si guarda alla forma dell'hamiltoniana, in base alla quale solo τ avanza con velocità $\dot{\tau} = 1$ mentre le altre variabili non si muovono.

Come per il pendolo (paragrafo 3.2), possiamo ora utilmente passare dalle variabili di energia–tempo (h, τ) a variabili di azione–angolo (I_1, φ_1), ponendo

$$\varphi_1 = 2\pi \frac{\tau}{T(h, l)} , \quad I_1 = \frac{1}{\pi} \int_{r_{\min}}^{r_{\max}} \sqrt{2(h - \mathcal{V}(r, l))} \, dr ; \quad (3.30)$$

la trasformazione si completa canonicamente sull'altro grado di libertà scrivendo la generatrice (della trasformazione inversa), che denotiamo ancora con S ,

$$S(h, l, \varphi_1, \varphi_2) = \frac{\varphi_1}{\pi} \int_{r_{\min}}^{r_{\max}} \sqrt{2(h - \mathcal{V}(r, l))} \, dr + l \varphi_2 .$$

Questa generatrice dà le (3.30) e inoltre fornisce

$$I_2 = l , \quad \psi = \varphi_2 - \frac{\varphi_1}{\pi} \int_{r_{\min}}^{r_{\max}} \frac{\frac{\partial \mathcal{V}}{\partial p_\vartheta}(r, l)}{\sqrt{2(h - \mathcal{V}(r, l))}} \, dr ;$$

in particolare $I_2 = p_\vartheta$. Procedendo come sopra, si trova che l'integrale vale

$$\int_{r_{\min}}^{r_{\max}} \frac{\dot{\vartheta}}{\dot{r}} \, dr = \frac{\Delta\vartheta}{2} ,$$

ove $\Delta\vartheta$ denota l'avanzamento dell'angolo ϑ tra due passaggi successivi al perielio. Se denotiamo con Θ l'avanzamento del perielio a ogni passaggio, allora $\Delta\vartheta = 2\pi + \Theta$; segue

$$\varphi_2 = \psi + \varphi_1 \left(1 + \frac{\Theta}{2\pi} \right) .$$

Le (3.30), assieme a $I_2 = l = p_\vartheta$ e a questa espressione per φ_2 , sono una scelta possibile di variabili di azione–angolo per il moto centrale. Per ottenere la scelta fatta nel paragrafo 3.3, che è quella più comune, occorre un ultimo ritocco, passando a nuove variabili di azione–angolo (I', φ') così definite:

$$\begin{cases} \varphi'_1 = \varphi_1 \\ \varphi'_2 = \varphi_2 - \varphi_1 \end{cases} , \quad \begin{cases} I'_1 = I_1 + I_2 \\ I'_2 = I_2 \end{cases} .$$

In particolare si ha

$$I'_1 = \frac{1}{\pi} \int_{r_{\min}(h, I'_2)}^{r_{\max}(h, I'_2)} \sqrt{2(h - \mathcal{V}(r, I'_2))} \, dr + I'_2 . \quad (3.31)$$

Rimuovendo gli apici si ottengono precisamente le variabili di azione–angolo introdotte nel paragrafo 3.3 per il moto centrale; il modulo che lì appare nell'espressione di I_1 tiene conto della possibilità che p_ϑ sia negativo.

Quanto all'hamiltoniana si ha evidentemente $H(I, \varphi) = h(I)$, con $h(I)$ definita implicitamente dalla (3.31). Per potenziali V e dunque \mathcal{V} generici non si ottiene un'espressione migliore. Nel caso kepleriano $V(r) = -k/r$ tuttavia tutto si esplicita e si calcola (l'integrale è elementare). Non occorrono nemmeno calcoli complicati, se si accetta di usare il fatto, ereditato dalla meccanica newtoniana, che le orbite sono chiuse e il perielio non avanza ($\Theta = 0$, $\Delta\vartheta = 2\pi$). Infatti dalla (3.31) segue facilmente, derivando rispetto a I_2 ,

$$\left. \frac{\partial I_1}{\partial I_2} \right|_{h=\text{cost}} = -\frac{\Delta\vartheta}{2\pi} + 1 = 0 ,$$

pertanto nel caso kepleriano il secondo membro della (3.31) è indipendente da I_2 (I_1 dipende dall'energia h ma non dal momento angolare), e lo si può calcolare inserendovi il valore di I_2 che si preferisce. In particolare si può prendere il valore di I_2 corrispondente al moto circolare di energia h , $I_2 = k/\sqrt{-2h}$. Ma per i moti circolari $r_{\min} = r_{\max}$, cioè l'integrale si annulla, e in definitiva

$$I_1 = \frac{k}{\sqrt{-2h}} , \quad h = h(I_1) = -\frac{k^2}{2I_1^2} .$$

D. L'equazione di Hamilton–Jacobi dipendente dal tempo.

Accenniamo per completezza alla vera e propria equazione di Hamilton–Jacobi (non ridotta). A questo scopo osserviamo preliminarmente che se, come abbiamo supposto, l'hamiltoniana H sulla quale intendiamo lavorare è indipendente dal tempo, allora la Proposizione 18 si può riformulare nel seguente modo del tutto equivalente:

PROPOSIZIONE 18' *Nelle stesse ipotesi della Proposizione 18, la funzione*

$$S'(\tilde{p}_1, \dots, \tilde{p}_n, q_1, \dots, q_n, t) = S(\tilde{p}_1, \dots, \tilde{p}_n, q_1, \dots, q_n) - t h(\tilde{p}_1, \dots, \tilde{p}_n)$$

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

DIMOSTRAZIONE. Si procede come per la Proposizione 18, ricordando che questa volta S' dipende da t e dunque (proposizione 15) $K = H \circ w + \frac{\partial S'}{\partial t}$. \square

Per capire, riprendiamo il caso dell'oscillatore armonico. La funzione generatrice S' dipendente dal tempo è ora

$$S'(I, q, t) = \int_0^q \sqrt{2\omega I - \omega^2 x^2} dx - \omega I t$$

e corrispondentemente si ha

$$\begin{aligned} p &= \frac{\partial S'}{\partial q} = \sqrt{2\omega I - \omega^2 q^2} \\ \varphi &= \frac{\partial S'}{\partial I} = \int_0^q \frac{dx}{\sqrt{2I/\omega - x^2}} - \omega t = \arcsin \sqrt{\omega/2I} q - \omega t. \end{aligned}$$

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

Alla Proposizione 18' si può dare una formulazione più generale, adatta anche al caso di dipendenza esplicita di H dal tempo. Premettiamo una definizione:

DEFINIZIONE 10 *L'equazione alle derivate parziali*

$$H\left(\frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_n}, q_1, \dots, q_n, t\right) + \frac{\partial S}{\partial t} = 0$$

per l'incognita $S(q_1, \dots, q_n, t)$ è detta equazione di Hamilton–Jacobi associata all'hamiltoniana H . Si dice poi integrale completo dell'equazione ogni famiglia di soluzioni $S(\alpha_1, \dots, \alpha_n, q_1, \dots, q_n, t)$, dipendente dagli n parametri $\alpha_1, \dots, \alpha_n$, tale che risulti

$$\det\left(\frac{\partial^2 S}{\partial \alpha_i \partial q_j}\right) \neq 0.$$

(Avendo a che fare con un'unica generatrice, abbiamo usato il simbolo S anziché S' .) E' allora di immediata verifica la seguente

PROPOSIZIONE 19 *Se la funzione $S(\alpha_1, \dots, \alpha_n, q_1, \dots, q_n, t)$ è un integrale completo dell'equazione di Hamilton–Jacobi relativa a H , allora $S(\tilde{p}_1, \dots, \tilde{p}_n, q_1, \dots, q_n, t)$ genera una trasformazione canonica dipendente dal tempo $(p, q) = w(\tilde{p}, \tilde{q}, t)$ che coniuga H a K identicamente nulla.*

Non ci addentreremo ulteriormente in quest'argomento.

4 Introduzione alla teoria hamiltoniana delle perturbazioni

4.1 Sistemi prossimi a sistemi integrabili

Mentre i sistemi integrabili vanno considerati in qualche modo eccezionali, sono invece abbondanti nel mondo fisico, e interessantissimi per la fenomenologia che presentano anche al di là dell'interesse applicativo, i sistemi “prossimi” a sistemi integrabili. Con questo termine ci si riferisce a sistemi che, nelle variabili di azione–angolo del sistema integrabile di riferimento, si possono scrivere nella forma

$$H_\varepsilon(I, \varphi) = h(I) + \varepsilon f(I, \varphi, \varepsilon) , \quad (4.1)$$

ove ε è un piccolo parametro. In questo contesto h è detta *hamiltoniana imperturbata*, mentre H_ε e εf sono dette rispettivamente *hamiltoniana perturbata* e *perturbazione*. La Meccanica Celeste offre moltissime situazioni di questo genere. Si pensi al Sistema Solare nella sua versione più semplice: un Sole molto massivo e otto pianeti²³ relativamente leggeri. Il rapporto tra la massa del pianeta più grande (Giove) e quella del Sole è piuttosto piccola, circa 0.95×10^{-3} ; una prima ragionevole approssimazione è allora quella di trascurare le masse dei pianeti, riducendo così il sistema a una collezione di otto problemi di Keplero disaccoppiati. Il modello kepleriano puro che ne risulta è con evidenza integrabile, e fornisce un ottimo problema imperturbato di partenza. Tenendo poi conto della massa dei pianeti, o anche solo di alcuni di essi, si ottengono modelli perturbati della forma (4.1). Il sistema si può complicare a piacere, introducendo satelliti, asteroidi, comete, o tenendo conto che pianeti e satelliti non sono punti materiali ma corpi estesi (approssimativamente) rigidi. Ma per avere un problema interessante, ricco di fenomenologia a volte stupefacente, e il cui studio è altamente non banale, non occorrono sistemi complicati: come ben comprese Poincaré, bastano tre soli corpi — ad esempio il Sole, Giove e un altro pianeta — per trovarsi di fronte alla possibilità di moti davvero complessi, non facili nemmeno a immaginarsi, chiamati oggi “moti caotici”. Addirittura è sufficiente che il terzo corpo abbia massa trascurabile — il Sole, Giove e un asteroide non in grado di modificare il moto dei due corpi maggiori — per trovarsi di fronte a domande semplici (ben poste da secoli) cui tuttavia non si sa ancora rispondere compiutamente.

Al di fuori della Meccanica Celeste, troviamo ad esempio il problema del comportamento di un sistema di oscillatori armonici debolmente accoppiati. E' questo uno dei problemi più capillarmente diffusi in ambito fisico: ovunque si trovi qualche cosa che oscilli attorno a un equilibrio stabile, la descrizione di prima approssimazione che si ottiene linearizzando il problema (si riveda su un qualunque testo la trattazione delle piccole oscillazioni in ambito lagrangiano) è quello di un insieme di n oscillatori armonici debolmente accoppiati da termini cubici o di ordine superiore; corrispondentemente nel linguaggio hamiltoniano

$$H(p, q) = H_2(p, q) + V_3(q) + V_4(q) + \dots$$

ove

$$H_2(p, q) = \frac{1}{2} \sum_{i=1}^n \omega_i (p_i^2 + q_i^2) ,$$

mentre V_s , $s \geq 3$, denota un polinomio omogeneo di grado s nelle q . Per studiare il moto in un piccolo intorno di raggio ε del punto di equilibrio $(p, q) = (0, 0)$ conviene effettuare il riscalamento

²³Il 24/8/2006 l'Unione Astronomica Internazionale ha declassato Plutone da “pianeta” a “pianeta nano”, nuova categoria di cui fanno parte anche Cerere (promosso: prima era considerato un asteroide) e pochi altri oggetti trans-nettuniani.

(canonico, ma non strettamente canonico)

$$p = \varepsilon \tilde{p}, \quad q = \varepsilon \tilde{q},$$

che dà alla nuova hamiltoniana $K_\varepsilon(\tilde{p}, \tilde{q}) = \varepsilon^{-2}H(\varepsilon\tilde{p}, \varepsilon\tilde{q})$ la forma

$$K_\varepsilon(\tilde{p}, \tilde{q}) = H_2(\tilde{p}, \tilde{q}) + \varepsilon V_3(\tilde{q}) + \varepsilon^2 V_4(\tilde{q}) + \dots,$$

questa volta in un dominio di diametro indipendente da ε . L'introduzione delle variabili di azione–angolo dei singoli oscillatori conduce evidentemente a una hamiltoniana della forma (4.1).²⁴

Il problema principale di fronte a una hamiltoniana della forma (4.1) riguarda il comportamento, su tempi lunghi, delle variabili di azione. Riprendendo l'esempio del Sistema Solare, e assumendo per semplificare che il sistema sia esattamente piano, questo vuol dire chiedersi cosa avvenga, su tempi lunghi, dell'energia E e del momento angolare M di ciascun pianeta, ovvero, equivalentemente, dei parametri geometrici — semiasse maggiore e minore — delle loro ellissi.²⁵ Si possono ipotizzare, schematicamente, due scenari: uno minimale, in cui queste quantità, non più costanti, si limitano tuttavia a oscillare in modo più o meno regolare con oscillazioni piccole per ε piccolo. L'altro più drammatico, in cui l'accumularsi nel tempo degli effetti della perturbazione produce, in aggiunta a possibili oscillazioni, una deriva magari molto lenta ma sistematica delle costanti del moto, che “nel corso dei secoli” potrebbero subire cambiamenti vistosi: fino alla fuoriuscita di un pianeta dal sistema, o alla sua caduta sul Sole (“fenomeni secolari”, nella terminologia tradizionale della Meccanica Celeste).

Riprendendo l'esempio degli oscillatori armonici debolmente accoppiati da termini nonlineari, una domanda naturale, in effetti cruciale per la Meccanica Statistica classica, riguarda l'efficacia della perturbazione nel produrre rilevanti scambi di energia tra i diversi oscillatori, onde soddisfare uno dei principi cardini della Meccanica Statistica classica, il principio di equipartizione dell'energia (si rimanda per questo ai testi di Meccanica Statistica). Molto interessante è anche il problema della stabilità dei punti lagrangiani triangolari L_4 e L_5 (esercizio 22, appendice E). In questo caso infatti, come si è detto, le pulsazioni ω_i che compaiono nell'hamiltoniana H_2 non hanno segno concorde, e di conseguenza la stabilità dell'equilibrio è strettamente legata alla conservazione separata delle tre energie armoniche.

Problemi analoghi si trovano per altri sistemi prossimi a sistemi integrabili, ad esempio per il corpo rigido di Eulero–Poincot perturbato. Tra le applicazioni più belle, ricordiamo il problema della precessione degli equinozi: il corpo rigido è la Terra, il punto di sospensione è il baricentro, la perturbazione è quella prodotta dal Sole e dalla Luna. E' abbastanza noto²⁶ che in prima appros-

²⁴In questo caso tuttavia la perturbazione non è regolare, come funzione delle azioni, quando una o più di esse si annullano: potenze dispari di q_i comportano infatti la presenza di termini contenenti $\sqrt{I_i}$. E' un fastidio (non una vera difficoltà) cui si rimedia, del quale in queste note non ci occuperemo.

²⁵Più precisamente, delle “ellissi osculatrici”, cioè che si avrebbero se improvvisamente la perturbazione sparisse.

²⁶Il primo a comprenderlo fu Ipparco di Nicea (190 – 120 a.C.; la scoperta della precessione è databile attorno al 130). Ipparco aveva a disposizione un catalogo con le coordinate celesti di alcune stelle fisse, risalente a circa 170 anni prima; compilando un nuovo catalogo osservò uno spostamento sistematico della longitudine delle stelle di circa 2°. Con intuizione straordinaria comprese il fenomeno: una precessione dell'equatore celeste (la proiezione sulla sfera celeste dell'equatore terrestre) relativamente all'eclittica (la curva percorsa dal Sole sulla sfera celeste nel corso di un anno, fissa nella sfera celeste), e conseguentemente la precessione lungo l'eclittica dei punti equinoziali (intersezioni delle due curve). Ipparco fornì per la precessione il valore straordinariamente preciso di 46'' di arco per anno (il valore corretto è 50.26), e corrispondentemente poté calcolare accuratamente la differenza tra *anno sidereo* e *anno tropico*. Si tratta della scoperta astronomica più importante dell'astronomia antica. Sfortunatamente l'opera di Ipparco, *Sullo spostamento dei segni solstiziali ed equinoziali*, è andata perduta, ma il risultato e il metodo da lui adottato (ingegnosissimo: la posizione del Sole sulla sfera celeste è dedotta dalla posizione della Luna in corrispondenza alle eclissi di Luna, quando Sole, Terra e Luna sono allineati) ci sono pervenuti attraverso l'*Almagesto* di Tolomeo.

simazione, per effetto delle perturbazioni congiunte di Sole e Luna, l’asse terrestre e il momento angolare (quasi paralleli) precedono attorno alla perpendicolare al piano dell’orbita, con un periodo di circa 26 000 anni. Come si fa il calcolo? E soprattutto: entro che limiti il fenomeno è compreso? Si possono escludere variazioni significative dell’inclinazione dell’asse terrestre? Sono domande difficili, alcune molto difficili, alle quali la comunità dei fisici matematici e dei meccanici celesti ancora non è in grado di dare risposte esaurienti. L’appendice I riporta la deduzione (in una certa approssimazione) della “formula classica” per la precessione degli equinozi, che conduce alla stima sopra citata del periodo di precessione della Terra.

4.2 La “stima a priori”

Supponiamo di avere un sistema hamiltoniano H_ε prossimo a un sistema imperturbato H_0 ,

$$H_\varepsilon(p, q) = H_0(p, q) + \varepsilon f(p, q) ,$$

ma di non avere particolari informazioni su H_0 (in particolare, di non sapere se H_0 sia o meno integrabile). In generale, la vicinanza *delle hamiltoniane*, e dunque delle equazioni del moto, non garantisce la vicinanza *delle soluzioni* se non per tempi brevi: l’unica informazione di carattere generale infatti deriva (per i sistemi hamiltoniani come per qualunque altro sistema di equazioni differenziali) dal teorema di regolarità delle soluzioni al variare dei parametri: denotando brevemente, per un fissato dato iniziale (p, q) ,

$$d_\varepsilon(t) = \text{dist}(\Phi_{H_\varepsilon}^t(p, q), \Phi_{H_0}^t(p, q)) ,$$

il teorema garantisce stime *a priori* del tipo²⁷

$$d_\varepsilon(t) \leq C \varepsilon e^{\lambda|t|}$$

con C, λ costanti positive (λ è la costante di Lipschitz del problema, e fornisce col suo inverso un’unità di tempo naturale). Tale stima è molto povera, e non dice praticamente più niente dopo il tempo, breve per ogni ε fisicamente sensato,²⁸

$$|t| \sim \lambda^{-1} \log \varepsilon^{-1} . \tag{4.2}$$

Le cose cambiano in modo sostanziale se stiamo perturbando un sistema integrabile, come nella (4.1). Questo è evidente soprattutto se si guarda al comportamento delle azioni. Le equazioni del moto per l’hamiltoniana H_ε sono infatti della forma

$$\begin{aligned} \dot{I}_j &= - \varepsilon \frac{\partial f}{\partial \varphi_j} \\ \dot{\varphi}_j &= \omega_j(I) + \varepsilon \frac{\partial f}{\partial I_j} ; \end{aligned}$$

si vede allora che le azioni nel sistema perturbato evolvono lentamente, con velocità proporzionale a ε : più precisamente, rivoltando in forma integrale l’equazione differenziale per le azioni, cioè scrivendo

$$I_j(t) = I_j(0) - \varepsilon \int_0^t \frac{\partial f}{\partial \varphi_j}(I(t), \varphi(t)) dt , \tag{4.3}$$

²⁷Non sono stime pessimistiche ma realistiche in casi concreti, ad esempio in prossimità del punto di equilibrio instabile del pendolo, se la perturbazione cambia ω in $\omega + \varepsilon$.

²⁸La funzione $\log x$ notoriamente diverge per $x \rightarrow \infty$. Ma per qualunque grandezza x fisicamente immaginabile $\log x$ non è mai grande (se x è il rapporto tra il diametro dell’Universo e il diametro di un protone, $\log x < 100$).

si ottiene facilmente una stima *a priori* del tipo²⁹

$$|I_j(t) - I_j(0)| < C\varepsilon |t| , \quad (4.4)$$

ove C è il massimo di $|\frac{\partial f}{\partial \varphi_j}|$ nel dominio ove si svolge il moto e su j . Dal momento che nel sistema imperturbato si ha $I_j(t) = I_j(0)$, la (4.4) si legge anche come lento scostamento nel tempo della dinamica perturbata da quella imperturbata, limitatamente al comportamento delle azioni. La scala di tempo sulla quale le azioni evolvono in maniera significativa, o equivalentemente il loro moto differisce dal moto imperturbato, è ora

$$|t| \sim \frac{1}{\varepsilon} , \quad (4.5)$$

ben più lunga della (4.2) per ε piccolo. Su scale di tempo più brevi (ma sempre lunghe per ε piccolo), ad esempio $|t| \sim 1/\sqrt{\varepsilon}$, le azioni variano di quantità piccole con ε , $|I_j(t) - I_j(0)| \sim \sqrt{\varepsilon}$.

Le variabili di azione sono dette *variabili lente* del sistema. Al contrario, come mostrano le equazioni del moto, gli angoli (salvo il caso in cui una o più pulsazioni si annullino) evolvono su una scala di tempo breve; anzi, fintantoché le azioni si discostano di poco dal valore iniziale, anche le velocità angolari si mantengono pressoché costanti, e dunque gli angoli ruotano, in prima approssimazione, in modo uniforme. Le variabili angolari sono dette *variabili veloci* del sistema.

- La separazione delle variabili in lente e veloci è però più delicata di quello che sembra. Ad esempio, se per un particolare valore di I si ha $\omega_1 = \omega_2$, allora φ_1 e φ_2 sono veloci, ma la loro differenza $\varphi_1 - \varphi_2$ è lenta; più in generale, se le pulsazioni soddisfano una *relazione di risonanza*

$$k_1\omega_1 + k_2\omega_2 + \dots + k_n\omega_n = 0 , \quad k = (k_1, \dots, k_n) \in \mathbb{Z}^n , \quad (4.6)$$

allora la corrispondente combinazione di angoli $\sum_i k_i \varphi_i$ è lenta. Come si vedrà nel prossimo paragrafo, la presenza di tali variabili lente “nascoste” è una potenziale sorgente di difficoltà nella teoria delle perturbazioni.

ESERCIZIO 37 Si dia una “stima a priori” sugli angoli: $|\varphi_j(t, \varepsilon) - \varphi_j(t, 0)| < \dots$ [Risposta: si trova una stima del tipo $c\varepsilon t$ se ω è costante, $\varepsilon(ct + c't^2)$ se ω dipende da I . Suggerimento per quest'ultimo caso, un po' complicato: si scriva $\dot{\omega}_j(I)$, poi $\omega_j(I(t))$ in forma di integrale in t , infine $\varphi_j(t, \varepsilon)$ come integrale di $\dot{\varphi}$ in t ; interviene un integrale doppio, stimabile con l'area del dominio di integrazione $\sim t^2$ per il massimo della funzione integranda.]

Stime come la (4.4) e la (4.5) si sono ottenute senza davvero lavorare, solo sapendo che il sistema imperturbato è integrabile. E' spontaneo allora chiedersi se con una analisi più approfondita non si possa far di meglio. Ci sono buoni motivi per crederlo: la stima (4.4) infatti si ottiene dall'equazione integrale (4.3) supponendo, pessimisticamente, che gli effetti della perturbazione $\varepsilon \frac{\partial f}{\partial \varphi_j}$ vadano sommandosi; viceversa, è chiaro che sono possibili significative compensazioni: infatti il termine $\frac{\partial f}{\partial \varphi_j}$, essendo la derivata rispetto a un angolo di una funzione periodica, ha necessariamente

²⁹Questa stima, che potrebbe apparere euristica, si rende rigorosa nel modo seguente: sia $D = B \times \mathbb{T}^n$ il dominio di H_ε , e sia

$$\sup_{(I, \varphi) \in D} \max_{1 \leq j \leq n} \left| \frac{\partial f}{\partial \varphi_j} \right| = C < \infty .$$

Per $I(0)$ che disti dal bordo di B più di ε si deduce subito, dalla (4.3), la (4.4), per $|t| \leq 1/(C\varepsilon)$.

media sugli angoli nulla,³⁰ e dunque è presumibile che la perturbazione, nel corso del tempo, abbia segno alterno; quantomeno, è una possibilità da esplorare con attenzione.

Lo studio della dinamica dei sistemi prossimi a sistemi integrabili, su scale di tempo di ordine $1/\varepsilon$ o maggiori, è l’oggetto della *teoria hamiltoniana delle perturbazioni*, che si applica, sostanzialmente, ogniqualvolta vi siano, in un sistema hamiltoniano, variabili che evolvono su scale di tempo ben separate (le idee di fondo tuttavia, come avremo modo di accennare, prescindono dal carattere hamiltoniano del sistema, e sono profonde e più generali). La teoria delle perturbazioni è nata a cavallo tra ’700 e ’800 principalmente ad opera di Lagrange e Laplace, in connessione ai moti planetari. Nel corso dell’800 l’introduzione del formalismo hamiltoniano, e soprattutto della nozione di sistema integrabile da parte di Liouville, ha contribuito in modo decisivo al suo progresso, e si può dire che la teoria sia giunta a una prima maturazione, un secolo dopo la sua nascita, con il grosso lavoro di Poincaré. Un ulteriore salto qualitativo, assai rilevante anche per le sue implicazioni fisiche (ancora non abbastanza esplorate), si è avuto soprattutto a seguito dell’attività della scuola russa, negli anni ’60 e ’70 del novecento.

Procederemo così: nel prossimo paragrafo discuteremo brevemente l’idea principale della teoria delle perturbazioni, introducendo il cosiddetto “principio della media”. Poi tratteremo in dettaglio un esempio elementare “isocrono”, precisamente il caso di due oscillatori armonici debolmente accoppiati da una perturbazione “facile”, generalizzandolo progressivamente a sistemi isocroni più complessi. Poi ancora cercheremo di capire, attraverso un esempio significativo, il caso, più complesso ma anche più interessante, dei sistemi non isocroni. Infine illustreremo alcuni tra i maggiori risultati della teoria.

4.3 Il “principio della media”

Raccogliamo qui l’idea, adombrata verso la fine del paragrafo precedente, che si possa progredire nello studio di un sistema per tempi lunghi ($1/\varepsilon$ o maggiori) tenendo conto di possibili compensazioni negli effetti della perturbazione, più precisamente mettendo in evidenza un eventuale “effetto medio” della perturbazione, dominante su tempi lunghi. Questa idea è nota in letteratura con il nome di *principio della media*, vera anima della teoria delle perturbazioni.³¹ In questo paragrafo cercheremo di introdurre l’idea principale in modo critico, basandoci soprattutto su esempi.

A. Il sistema mediato.

Per maggior chiarezza, conviene impostare il problema in ambito più generale di quello hamiltoniano, che con le sue peculiarità maschera un po’ l’idea centrale. Consideriamo per questo un sistema di equazioni differenziali in $m + n$ variabili della forma

$$\dot{I} = \varepsilon F(I, \varphi) , \quad \dot{\varphi} = \omega(I) + \varepsilon G(I, \varphi) , \quad (4.7)$$

³⁰Se non fosse chiaro: si prenda per semplicità il caso di un solo angolo; la media di $\frac{\partial f}{\partial \varphi}$ è

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{\partial f}{\partial \varphi} d\varphi = \frac{1}{2\pi} [f(2\pi) - f(0)] = 0 .$$

La generalizzazione a n angoli è ovvia.

³¹Non soltanto: c’è una classe molto ampia di problemi fisici in cui si passa da una descrizione microscopica dettagliata a una descrizione su scala più macroscopica, tramite una operazione di media su variabili rapidamente fluttuanti. Si pensi alla nozione di pressione in un gas, o al passaggio dalla meccanica microscopica alle equazioni macroscopiche dell’idrodinamica (come le equazioni di Eulero dei fluidi ideali). Un altro esempio è dato dal procedimento sistematico di decimazione delle variabili nei metodi di tipo “gruppo di rinormalizzazione”, per le transizioni di fase o le teorie quantistiche di campo. Una bella rassegna è in N.G. Van Kampen, *On the elimination of fast variables*, Phys. Rep. **124**, 69–160 (1985).

ove $I = (I_1, \dots, I_m) \in \mathbb{R}^m$, $\varphi = (\varphi_1, \dots, \varphi_n) \in \mathbb{T}^n$. Nel caso di un sistema hamiltoniano quasi integrabile, con hamiltoniana della forma (4.1), si ha evidentemente $m = n$ e

$$\omega = \frac{\partial h}{\partial I}, \quad F = -\frac{\partial f}{\partial \varphi}, \quad G = \frac{\partial f}{\partial I}.$$

La (4.7) mostra che per ε piccolo le variabili I sono lente, mentre gli angoli φ sono veloci. In queste condizioni è naturale scomporre F in una parte media e una parte oscillante:

$$F(I, \varphi) = \mathcal{F}(I) + F_{\text{osc}}(I, \varphi), \quad \mathcal{F}(I) = (2\pi)^{-n} \int_0^{2\pi} \dots \int_0^{2\pi} F(I, \varphi) \, d\varphi_1 \dots d\varphi_n,$$

e pensare che in qualche modo, se ε è piccolo, la parte oscillante F_{osc} dia luogo a compensazioni e perciò contribuisca marginalmente al moto delle variabili lente, e invece il termine medio \mathcal{F} dia un contributo sistematico dominante su tempi lunghi. Ha senso allora considerare, accanto al sistema “vero” (4.7), il *sistema mediato*

$$\dot{J} = \varepsilon \mathcal{F}(J), \tag{4.8}$$

e chiedersi che relazione vi sia tra le soluzioni $J_\varepsilon(t)$ e $I_\varepsilon(t)$ a parità di dato iniziale $J^o = I^o$. Si osservi che il confronto è interessante solo su tempi $t \sim 1/\varepsilon$ o maggiori (infatti per $t \ll 1/\varepsilon$ né I né J si staccano apprezzabilmente dal valore iniziale comune, e dunque banalmente sono anche vicini tra loro), e che su tali tempi lo studio del sistema mediato (4.8) è *molto* più semplice di quello del sistema completo (4.7): non tanto perché (cosa pursempre utile) il sistema mediato contiene solo m variabili, ma soprattutto perché per esso $1/\varepsilon$ è una unità naturale di tempo, mentre per il sistema vero, a causa del moto veloce degli angoli, $1/\varepsilon$ è un tempo molto lungo (si rifletta su questo punto).³² Ci chiediamo allora se, ed eventualmente a quali condizioni, si ha

$$I_\varepsilon(t) - J_\varepsilon(t) \rightarrow 0 \quad \text{per} \quad \varepsilon \rightarrow 0$$

in tutta la scala di tempo $t \sim 1/\varepsilon$.

Se ciò avviene per ogni dato iniziale, si dirà che il sistema “soddisfa il principio della media” al primo ordine in ε (primo ordine perché relativo alla scala di tempo $1/\varepsilon$). Una definizione formale è questa:

DEFINIZIONE 11 Diremo che il sistema (4.7) soddisfa il principio della media al primo ordine in ε , se esiste $t_0 > 0$ tale che per ogni dato iniziale (I^o, φ^o) la soluzione $(I_\varepsilon(t), \varphi_\varepsilon(t))$ e la soluzione $J_\varepsilon(t)$ dell’equazione mediata (4.8) con dato iniziale $J^o = I^o$ sono tali che

$$\max_{|t| \leq t_0/\varepsilon} \|I_\varepsilon(t) - J_\varepsilon(t)\| \rightarrow 0 \quad \text{per} \quad \varepsilon \rightarrow 0.$$

Infine, osserviamo che nel caso hamiltoniano il sistema mediato è banale: per i sistemi hamiltoniani si ha infatti $F = -\frac{\partial H}{\partial \varphi}$ con H periodica in φ , e dunque $\mathcal{F} = 0$, perciò la soluzione dell’equazione mediata non è altro che $J_\varepsilon(t) = I^o$. Il modo appropriato di formulare la questione in ambito hamiltoniano è allora quello di chiedersi se le azioni, che sono costanti del moto per $\varepsilon = 0$, per ε piccolo e $t \sim 1/\varepsilon$ evolvano significativamente, o invece, a somiglianza dell’equazione mediata, restino vicine al valore iniziale I^o . La figura 17 illustra simbolicamente l’ipotesi di lavoro che sta dietro al principio della media, nel caso generale (a) e hamiltoniano (b).

³²Si pensi di dover risolvere numericamente le equazioni differenziali del moto. Come andrebbe preso, per il sistema vero e per quello mediato, il passo di integrazione? Come cambiano il costo e l’affidabilità del risultato?

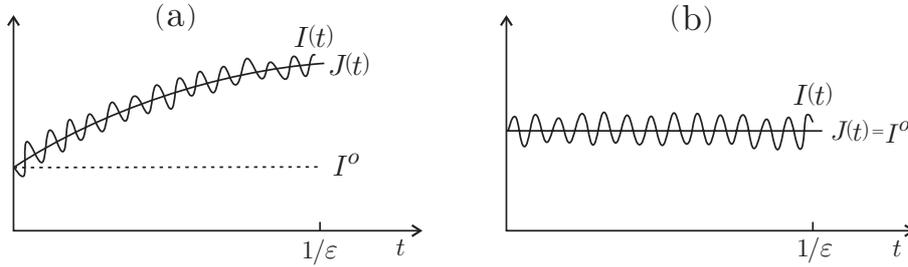


Figura 17: L'ipotesi di lavoro su $I_\varepsilon(t)$ e $J_\varepsilon(t)$ secondo il principio della media, nel caso generale (a) e hamiltoniano (b). Le oscillazioni si schiacciano a zero per $\varepsilon \rightarrow 0$.

B. esempi.

Diamo qui alcuni esempi elementari, presi ad eccezione del primo dal mondo hamiltoniano, utili a valutare il principio della media in modo critico.

(i) Per $m = n = 1$ consideriamo il sistema

$$\dot{I} = \varepsilon(a + b \cos \varphi), \quad \dot{\varphi} = \omega.$$

La soluzione è immediata: si ha infatti $\varphi(t) = \varphi^o + \omega t$, e dunque $\dot{I} = \varepsilon(a + b \cos(\varphi^o + \omega t))$, da cui, se $\omega \neq 0$,

$$I(t) = \varepsilon \int_0^t (a + b \cos(\varphi^o + \omega t')) dt' = I^o + \varepsilon a t + \frac{\varepsilon b}{\omega} (\sin(\varphi^o + \omega t) - \sin \varphi^o).$$

D'altra parte l'equazione mediata $\dot{J} = \varepsilon a$, con dato iniziale $J(0) = I^o$, è risolta da $J(t) = I^o + \varepsilon a t$. Il principio della media è pertanto soddisfatto (per ogni t , non solo fino al primo ordine). Se $\omega = 0$ la soluzione per la I è $I(t) = I^o + \varepsilon a t + \varepsilon b \cos(\varphi^o)t$, e il principio non è soddisfatto: ma in questo caso cade evidentemente il presupposto che l'angolo φ sia una variabile veloce.

(ii) Per $n = 2$ consideriamo il sistema (particolarissimo: parte imperturbata lineare nelle I , perturbazione indipendente da I)

$$H_\varepsilon(I, \varphi) = \omega_1 I_1 + \omega_2 I_2 + \varepsilon \cos(\varphi_2 - \varphi_1).$$

La soluzione (grazie alla particolarità) si scrive esplicitamente: si ha infatti banalmente $\dot{\varphi}_i = \omega_i$, e dunque $\varphi_i(t) = \varphi_i^o + \omega_i t$; sostituendo nell'equazione delle azioni si trova allora

$$\dot{I}_1 = -\varepsilon \sin(\varphi_2^o - \varphi_1^o + (\omega_2 - \omega_1)t), \quad \dot{I}_2 = +\varepsilon \sin(\varphi_2^o - \varphi_1^o + (\omega_2 - \omega_1)t).$$

Queste equazioni si integrano immediatamente e danno

$$I_1(t) = I_1^o + \varepsilon \mathcal{J}(t), \quad I_2(t) = I_2^o - \varepsilon \mathcal{J}(t),$$

con

$$\mathcal{J}(t) = \begin{cases} \frac{\cos(\varphi_2^o - \varphi_1^o + (\omega_2 - \omega_1)t) - \cos(\varphi_2^o - \varphi_1^o)}{\omega_2 - \omega_1} & \text{per } \omega_1 \neq \omega_2 \\ t \sin(\varphi_1^o - \varphi_2^o) & \text{per } \omega_1 = \omega_2 \end{cases}.$$

Si vede così che il principio della media è soddisfatto (alla grande, per tempi infiniti) se $\omega_1 \neq \omega_2$, mentre fallisce per $\omega_1 = \omega_2$. Se ne intuisce il motivo, e con esso la delicatezza della questione: benché gli angoli φ_1 e φ_2 siano individualmente veloci, in condizioni di risonanza, ovvero per $\omega_1 = \omega_2$, la combinazione lineare $\varphi_2 - \varphi_1$ non lo è, e proprio tale combinazione è presente nella perturbazione.

I prossimi esempi sono più sottili e il comportamento è meno intuitivo, perchè il sistema imperturbato non è più lineare e dunque le pulsazioni e le loro relazioni di risonanza dipendono dalle azioni; soprattutto, *in essi discuteremo del principio della media senza conoscere la soluzione esatta* (usare come sopra la soluzione esatta è un po' barare, è una situazione troppo eccezionale).

(iii) Per $n = 2$ consideriamo il sistema

$$H_\varepsilon(I, \varphi) = \frac{1}{2}(I_1^2 + I_2^2) - \varepsilon \cos(\varphi_2 - \varphi_1) \quad (4.9)$$

(due rotatori, con un potenziale che è minimo per $\varphi_1 = \varphi_2$). Non è difficile vedere che il principio della media è sempre soddisfatto, addirittura per ogni $t \in \mathbb{R}$. Una via per dimostrarlo è quella di osservare (procedendo come nel problema a due corpi) che la somma $J = I_1 + I_2$ (analoga alla quantità di moto del baricentro) resta costante; posto poi $p = I_2 - I_1$, $\vartheta = \varphi_2 - \varphi_1$, si vede subito che p, ϑ seguono le equazioni di un pendolo di pulsazione $\sqrt{2\varepsilon}$: infatti

$$\dot{p} = \dot{I}_2 - \dot{I}_1 = -2\varepsilon \sin(\varphi_2 - \varphi_1) = -2\varepsilon \sin \vartheta, \quad \dot{\vartheta} = \dot{\varphi}_2 - \dot{\varphi}_1 = I_2 - I_1 = p.$$

Segue che le oscillazioni di p , e dunque di $I_1 = \frac{1}{2}(J + p)$, $I_2 = \frac{1}{2}(J - p)$, sono limitate da $\text{cost} \sqrt{\varepsilon}$ (si tracci il ritratto in fase).³³

ESERCIZIO 38 *Si deduca, per il pendolo descritto dalle variabili p, ϑ , la disuguaglianza $|p(t) - p(0)| < 4\sqrt{2\varepsilon}$ per le librazioni, $|p(t) - p(0)| < 2\sqrt{2\varepsilon}$ per le rotazioni.*

(iv) Per $n = 2$ consideriamo il sistema

$$H_\varepsilon(I, \varphi) = \frac{1}{2}(I_1^2 - I_2^2) - \varepsilon \cos(\varphi_2 - \varphi_1). \quad (4.10)$$

La differenza rispetto al caso precedente (un segno meno nella parte cinetica) può apparire marginale, invece è molto significativa. Si vede infatti che *esistono particolari moti che non soddisfano il principio della media*. Per scovarli, osserviamo innanzitutto che (come del resto nell'esempio precedente) la somma $I_1 + I_2$ è una costante del moto: lo si verifica subito, ed è ovvio data l'invarianza di H per traslazione degli angoli. Le rette

$$I_1 + I_2 = c$$

³³Un altro modo di vederlo è quello di introdurre una trasformazione canonica adattata al problema, in cui la combinazione angolare $\varphi_2 - \varphi_1$ presente nell'hamiltoniana coincida con uno dei nuovi angoli, ad esempio $\tilde{\varphi}_1 = \varphi_1$, $\tilde{\varphi}_2 = \varphi_2 - \varphi_1$; la trasformazione completa $(I, \varphi) = w(\tilde{I}, \tilde{\varphi})$ si scrive

$$\begin{cases} \varphi_1 = \tilde{\varphi}_1 \\ \varphi_2 = \tilde{\varphi}_1 + \tilde{\varphi}_2 \end{cases}, \quad \begin{cases} I_1 = \tilde{I}_1 - \tilde{I}_2 \\ I_2 = \tilde{I}_2 \end{cases},$$

e la nuova hamiltoniana si trova subito essere

$$\tilde{H}(\tilde{I}, \tilde{\varphi}) = h(\tilde{I}) - \varepsilon \cos \tilde{\varphi}_2, \quad \text{con} \quad h(\tilde{I}) = \frac{1}{2}(\tilde{I}_1^2 + 2\tilde{I}_2^2 - 2\tilde{I}_1\tilde{I}_2) = \left(\tilde{I}_2 - \frac{1}{2}\tilde{I}_1\right)^2 + \frac{1}{4}\tilde{I}_1^2.$$

Si vede allora che \tilde{I}_1 resta esattamente costante, mentre $\tilde{I}_2, \tilde{\varphi}_2$ si muovono come in un pendolo di pulsazione $\sqrt{2\varepsilon}$, solo con \tilde{I}_2 che oscilla attorno a $\frac{1}{2}\tilde{I}_1$ anziché attorno allo zero. La conclusione, come sopra, è immediata.

sono pertanto invarianti e il moto si svolge obbligatoriamente su una di esse. Tra queste c'è la *retta risonante* $I_1 + I_2 = 0$, sulla quale la combinazione di angoli $\varphi_2 - \varphi_1$ non evolve: infatti $\dot{\varphi}_2 - \dot{\varphi}_1 = -I_2 - I_1 = 0$. Per dati iniziali su questa retta dunque si resta sulla retta, e si ha

$$\varphi_2(t) - \varphi_1(t) = \varphi_2^o - \varphi_1^o ;$$

segue immediatamente $\dot{I}_1 = -\dot{I}_2 = \varepsilon \sin(\varphi_2^o - \varphi_1^o)$, e

$$I_1(t) = -I_2(t) = I^o + \varepsilon t \sin(\varphi_2^o - \varphi_1^o) . \quad (4.11)$$

Il principio della media non è soddisfatto. La soluzione si completa banalmente con

$$\varphi_1(t) = \varphi^o + I^o t + \frac{1}{2} \varepsilon t^2 , \quad \varphi_2(t) = \varphi_1(t) + \frac{\pi}{2} ; \quad (4.12)$$

per esercizio si può verificare che le (4.11), (4.12) risolvono le equazioni di Hamilton.

- Il punto cruciale è che in questo esempio, a differenza del precedente che pure ha le medesime rette invarianti, il vettore \dot{I} è parallelo anziché perpendicolare alla retta risonante, perciò il moto delle azioni mantiene la relazione di risonanza. Il segno meno nell'hamiltoniana imperturbata gioca per questo un ruolo essenziale.

ESERCIZIO 39 *Si riprenda l'esempio (iii), e si mostri che per dati iniziali sulla retta risonante (qui $I_1 - I_2 = 0$) il moto delle I è trasverso alla retta. Si verifichi che il risultato si mantiene anche se a $\varphi_2 - \varphi_1$ si sostituisce, ad argomento del coseno, qualunque altra combinazione lineare $k \cdot \varphi$. Si verifichi che il risultato si mantiene sostituendo alla parte cinetica una diversa funzione convessa di I_1 e I_2 , o anche una funzione come $I_1 + I_2^2$.*

ESERCIZIO 40 *Si consideri*

$$H_\varepsilon(I, \varphi) = \frac{1}{2}(I_1^2 - 4I_2^2) - \varepsilon \cos(k \cdot \varphi) ;$$

si determini k in modo tale che non sia soddisfatto il principio della media. Si mettano in evidenza altre hamiltoniane imperturbate quadratiche in I_1 e I_2 per cui, in corrispondenza a un opportuno k nella perturbazione, il principio della media non è soddisfatto.

Infine, un bell'esempio non hamiltoniano in cui vale il principio della media è l'equazione di Van der Pol; si veda per questo l'appendice H.

4.4 Sistemi isocroni perturbati: un passo perturbativo

Affrontiamo qui lo studio di un problema semplice, ma assolutamente non banale, attraverso il quale illustriamo uno dei procedimenti più usati in teoria hamiltoniana delle perturbazioni per studiare la stabilità delle azioni su tempi lunghi.

A. Il problema.

Consideriamo un sistema di n oscillatori armonici debolmente accoppiati, già scritto in variabili di angolo azione: precisamente

$$H_\varepsilon(I, \varphi) = h(I) + \varepsilon f(I, \varphi) , \quad (I, \varphi) \in D = B \times \mathbb{T}^n , \quad (4.13)$$

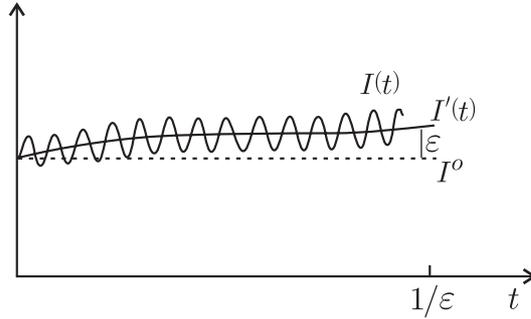


Figura 18: Le variabili \tilde{I} evolvono con $\frac{d\tilde{I}}{dt} \sim \varepsilon^2$ e a $t = 1/\varepsilon$ si scostano da $I(0)$ al più di $\mathcal{O}(\varepsilon)$. Le variabili I istantaneamente sono più rapide ($\frac{dI}{dt} \sim \varepsilon$), ma restano agganciate entro ε alle \tilde{I} . Il principio della media è soddisfatto.

con h lineare,

$$h(I) = \omega \cdot I . \quad (4.14)$$

Pur non essendo la situazione più comune, per evitare complicazioni che oscurerebbero le idee che andiamo a esporre supporremo che f sia regolare in tutte le variabili (e dunque non contenga radici delle azioni, come invece avviene genericamente nei problemi di piccole oscillazioni). Seguendo uno dei metodi perturbativi più tipici, ci proponiamo di “riassorbire” per quanto possibile la perturbazione εf in un cambiamento di variabili. Più precisamente, cerchiamo una trasformazione canonica prossima all’identità:

$$(I, \varphi) = w_\varepsilon(\tilde{I}, \tilde{\varphi}) , \quad \begin{cases} I = \tilde{I} + \varepsilon \dots \\ \varphi = \tilde{\varphi} + \varepsilon \dots \end{cases} ,$$

tale che la nuova hamiltoniana $\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) = H_\varepsilon(w_\varepsilon(\tilde{I}, \tilde{\varphi}))$ sia “più integrabile di H ”, ad esempio sia integrabile a meno di termini di ordine ε^2 :

$$\tilde{H}_\varepsilon = \omega \cdot \tilde{I} + \varepsilon g(\tilde{I}) + \varepsilon^2 \tilde{f}(\tilde{I}, \tilde{\varphi}, \varepsilon) ,$$

con opportuna g . Se questo programma va in porto, allora immediatamente si deduce che le nuove azioni \tilde{I} evolvono, al più, con velocità ε^2 , dunque per una variazione significativa ci vuole almeno il tempo $1/\varepsilon^2$, mentre si ha

$$\|\tilde{I}(t) - \tilde{I}(0)\| < \text{cost } \varepsilon \quad \text{per } |t| \leq \frac{\text{cost}}{\varepsilon} .$$

Ma poiché w_ε è prossima all’identità, sia $I(t)$ che $I(0)$ sono ε -vicine rispettivamente a $\tilde{I}(t)$ e $\tilde{I}(0)$: dunque anche per esse si può scrivere

$$\|I(t) - I(0)\| < \text{cost } \varepsilon \quad \text{per } |t| \leq \frac{\text{cost}}{\varepsilon} ,$$

mentre un’evoluzione significativa può avvenire, al più, su tempi $t \sim 1/\varepsilon^2$. Si veda la figura 18. In queste condizioni evidentemente il principio della media è soddisfatto.

Per produrre la trasformazione canonica che ci interessa si può procedere, indifferentemente, con il metodo delle funzioni generatrici in variabili miste oppure con il metodo di Lie. Seguiamo quest’ultimo, e pertanto cerchiamo una hamiltoniana generatrice, che denotiamo $\chi(I, \varphi)$, tale che il suo flusso Φ_χ^ε al tempo ε sia la trasformazione con le proprietà cercate.

Per ogni funzione $F : D \rightarrow \mathbb{R}$ si ha, come già sappiamo (paragrafo 2.2B, formula (2.32)), lo sviluppo

$$F \circ \Phi_\chi^\varepsilon = F + \varepsilon\{F, \chi\} + \frac{\varepsilon^2}{2}\{\{F, \chi\}, \chi\} + \dots$$

Avendo deciso di trascurare i contributi di ordine ε^2 , basta usare per h lo sviluppo troncato al primo ordine, mentre per f (che già è moltiplicata per il piccolo parametro ε) basta addirittura il troncamento all'ordine zero. Scriveremo pertanto

$$h \circ \Phi_\chi^\varepsilon = h + \varepsilon\{h, \chi\} + \mathcal{O}(\varepsilon^2), \quad f \circ \Phi_\chi^\varepsilon = f + \mathcal{O}(\varepsilon), \quad (4.15)$$

e la nuova hamiltoniana sarà corrispondentemente della forma

$$H_\varepsilon \circ \Phi_\chi^\varepsilon = h + \varepsilon(\{h, \chi\} + f) + \mathcal{O}(\varepsilon^2) \quad (4.16)$$

(il termine di ordine ε^2 , all'occorrenza, si scrive facilmente troncando gli sviluppi di $h \circ \Phi_\chi^\varepsilon$ e $f \circ \Phi_\chi^\varepsilon$ all'ordine successivo; si veda oltre). Il nostro programma pertanto si realizza se riusciamo a trovare χ e g tali che

$$\{h, \chi\}(\tilde{I}, \tilde{\varphi}) + f(\tilde{I}, \tilde{\varphi}) = g(\tilde{I}),$$

ovvero, poiché $\{h, \chi\} = -\omega \cdot \frac{\partial \chi}{\partial \varphi}$, se riusciamo a risolvere l'equazione alle derivate parziali

$$\omega \cdot \frac{\partial \chi}{\partial \varphi} = f - g, \quad (4.17)$$

ove f è nota mentre χ e g sono incognite; a g si richiede che dipenda solo da \tilde{I} . Quest'ultima richiesta in realtà fa sì che la funzione g resti determinata: prendendo infatti la media sugli angoli a entrambi i membri della (4.17), e ricordando che la derivata rispetto a un angolo di una funzione periodica ha necessariamente media nulla, si deduce subito che deve essere

$$g = \langle f \rangle_\varphi. \quad (4.18)$$

L'equazione da risolvere è così

$$\omega \cdot \frac{\partial \chi}{\partial \varphi} = f - \langle f \rangle_\varphi. \quad (4.19)$$

- E' un utile esercizio provare a procedere, anziché col metodo di Lie, cercando la funzione generatrice in variabili miste $S(\tilde{I}, \varphi)$ che realizza la trasformazione w con le proprietà cercate. Brevemente: posto $S = \tilde{I} \cdot \varphi + \varepsilon \mathcal{S}(\tilde{I}, \varphi)$, le equazioni della trasformazione sono

$$I(\tilde{I}, \varphi) = \tilde{I} + \varepsilon \frac{\partial \mathcal{S}}{\partial \varphi}(\tilde{I}, \varphi), \quad \tilde{\varphi}(\tilde{I}, \varphi) = \varphi + \varepsilon \frac{\partial \mathcal{S}}{\partial \tilde{I}}(\tilde{I}, \varphi),$$

e all'ordine ε si invertono immediatamente:

$$\varphi(\tilde{I}, \tilde{\varphi}) = \tilde{\varphi} - \varepsilon \frac{\partial \mathcal{S}}{\partial \tilde{I}}(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^2), \quad \text{e poi} \quad I(\tilde{I}, \tilde{\varphi}) = \tilde{I} + \varepsilon \frac{\partial \mathcal{S}}{\partial \varphi}(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^2);$$

sostituendo in H si trova allora

$$\begin{aligned} \tilde{H}(\tilde{I}, \tilde{\varphi}) &= h(I(\tilde{I}, \tilde{\varphi})) + \varepsilon f(I(\tilde{I}, \tilde{\varphi}), \varphi(\tilde{I}, \tilde{\varphi})) \\ &= h(\tilde{I}) + \varepsilon \omega \cdot \frac{\partial \mathcal{S}}{\partial \varphi}(\tilde{I}, \tilde{\varphi}) + \varepsilon f(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^2). \end{aligned}$$

L'equazione da risolvere per eliminare $\tilde{\varphi}$ all'ordine ε è così

$$-\omega \cdot \frac{\partial \mathcal{S}}{\partial \varphi} = f - g, \quad (4.20)$$

cioè esattamente la stessa di prima, con la corrispondenza $\mathcal{S} = -\chi$. Un minimo di riflessione mostra che questa equazione altro non è che l'equazione di Hamilton–Jacobi per l'hamiltoniana $H(I, \varphi) = h(I) + \varepsilon f(I, \varphi)$, approssimata (troncata) all'ordine ε . Infatti, l'equazione di Hamilton–Jacobi per l'hamiltoniana H^ε si scrive (utilizzando il simbolo \tilde{h} al posto di h per la costante, per evitare conflitti)

$$\omega \cdot \frac{\partial S}{\partial \varphi} + \varepsilon f\left(\frac{\partial S}{\partial \varphi}, \varphi, \varepsilon\right) = \tilde{h}.$$

Per trovare un integrale completo approssimato, che risolva l'equazione a meno di termini di ordine ε^2 , è spontaneo porre

$$S(\alpha, \varphi) = \alpha \cdot \varphi + \varepsilon \mathcal{S}(\alpha, \varphi), \quad \tilde{h}(\alpha) = \omega \cdot \alpha + \varepsilon g(\alpha);$$

sostituendo nell'equazione di Hamilton–Jacobi e raccogliendo i termini di ordine ε , si ottiene allora la (4.20).

B. Il caso $n = 2$, f con sviluppo di Fourier finito, ω non risonante.

Veniamo allora alla soluzione della (4.17). Ci restringiamo per semplicità al caso di due soli oscillatori armonici (la generalizzazione poi non è difficile). Per evitare problemi di convergenza supponiamo qui che la serie di Fourier della perturbazione sia finita, ovvero che per opportuno $\mathcal{K} \subset \mathbb{Z}^2$ finito sia

$$f(I, \varphi) = \sum_{k=(k_1, k_2) \in \mathcal{K}} \hat{f}_k(I) e^{ik \cdot \varphi} \quad (4.21)$$

(la generalizzazione, cioè la rimozione di questa ipotesi, non è però altrettanto immediata). Guardando la (4.17), si vede che serve χ capace, per così dire, di compensare tutti i termini della serie di Fourier di f , tranne il valor medio \hat{f}_0 (che già non dipende dagli angoli). Dobbiamo allora cercare χ della forma

$$\chi(I, \varphi) = \sum_{k \in \mathcal{K}} \hat{\chi}_k(I) e^{ik \cdot \varphi},$$

al fine di soddisfare poi la (4.17) componente per componente. Poiché

$$\omega \cdot \frac{\partial \chi}{\partial \varphi} = \sum_{k \in \mathcal{K}} (ik \cdot \omega) \chi_k(I) e^{ik \cdot \varphi},$$

si vede bene che per accomodare la componente $k = 0$ bisogna e basta porre

$$g(I) = f_0(I), \quad (4.22)$$

mentre per $k \neq 0$ si deve trovare χ_k tale che

$$(ik \cdot \omega) \hat{\chi}_k(I) = \hat{f}_k(I); \quad (4.23)$$

il termine $\hat{\chi}_0$ resta libero e lo si può prendere per esempio nullo. La (4.22) coincide, si osservi, con la (4.18).

Perché la (4.23) si possa risolvere per ogni k , è necessario che risulti

$$k_1\omega_1 + k_2\omega_2 \neq 0 \quad \text{per } k \in \mathcal{K} \setminus \{0\} ; \quad (4.24)$$

questa condizione, detta *condizione di non risonanza* (limitatamente a $k \in \mathcal{K}$) esclude la presenza di rapporti razionali tra le componenti ω_1 e ω_2 di ω , precisamente

$$\frac{\omega_1}{\omega_2} \neq -\frac{k_2}{k_1} \quad \text{per } k \in \mathcal{K} .$$

Se questa condizione è soddisfatta, allora la (4.17) è risolta da

$$\chi(I, \varphi) = \sum_{k \in \mathcal{K} \setminus \{0\}} \frac{\hat{f}_k(I)}{ik \cdot \omega} e^{ik \cdot \varphi} \quad (4.25)$$

(si è preso $\hat{\chi}_0 = 0$) e la nuova hamiltoniana ha la forma voluta

$$\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) = h(\tilde{I}) + \varepsilon f_0(\tilde{I}) + \mathcal{O}(\varepsilon^2) . \quad (4.26)$$

In conclusione: il nostro programma di rendere il sistema integrabile “a vista” a meno di termini di ordine ε^2 funziona, purché sia soddisfatta la condizione di non risonanza (4.24). La (4.26) è detta *forma normale non risonante* di H_ε , del primo ordine in ε . Essa mostra che *la prima correzione all'hamiltoniana integrabile, dovuta alla perturbazione, è data dal valor medio f_0 della perturbazione sugli angoli*; il nuovo termine non muove le azioni, ma corregge un po' le frequenze del sistema e lo rende debolmente anarmonico.

- Quanto abbiamo visto si può riformulare, con linguaggio più algebrico, in questo modo: sia $\mathcal{F}_\mathcal{K}$ lo spazio delle funzioni con sviluppo di Fourier finito, precisamente con $k \in \mathcal{K}$, e sia L_ω l'operatore $\omega \cdot \frac{\partial}{\partial \varphi}$, $L_\omega: \mathcal{F}_\mathcal{K} \rightarrow \mathcal{F}_\mathcal{K}$. L'equazione (4.17) da risolvere si riscrive con questa notazione

$$L_\omega \chi = f - \langle f \rangle . \quad (4.27)$$

Ora si ha $L_\omega e^{ik \cdot \varphi} = i(k \cdot \omega) e^{ik \cdot \varphi}$, pertanto le funzioni $u_k = e^{ik \cdot \varphi}$, $k \in \mathcal{K}$, sono per $\mathcal{F}_\mathcal{K}$ una base di autofunzioni di L_ω , con autovalori $\lambda_k = ik \cdot \omega$. Perciò se è soddisfatta la relazione di non risonanza (4.24) il nucleo di L_ω si riduce alle sole funzioni costanti; ma allora $f - \langle f \rangle$ non ha componenti sul nucleo, dunque L_ω nella (4.27) si inverte e si ha

$$\chi = L_\omega^{-1}(f - \langle f \rangle) = \sum_{k \in \mathcal{K} \setminus \{0\}} \frac{\hat{f}_k}{\lambda_k} u_k ,$$

che è la (4.25). Si vede bene che lo sviluppo di Fourier di f è naturale, in questo problema, perchè nella base delle armoniche di Fourier $e^{ik \cdot \omega}$ l'operatore L_ω è diagonale.

Quanto abbiamo visto fino a qui si può riassumere nella proposizione seguente:

PROPOSIZIONE 20 *Sia data l'hamiltoniana a due gradi di libertà*

$$H_\varepsilon(I, \varphi) = \omega \cdot I + \varepsilon f(I, \varphi) ,$$

e sia f della forma (4.21). Se le frequenze soddisfano la relazione di non risonanza (4.24), allora esiste una trasformazione canonica ε -vicina all'identità tale che la nuova hamiltoniana ha la forma (4.26); corrispondentemente per ogni moto di H si ha

$$\|I(t) - I(0)\| < \text{cost } \varepsilon \quad \text{per } |t| \leq \frac{\text{cost}}{\varepsilon} .$$

All'occorrenza, il termine $\mathcal{O}(\varepsilon^2)$ lasciato indicato nella (4.26) si scrive facilmente: troncando infatti gli sviluppi (4.15) di $h \circ \Phi_\chi^\varepsilon$ e $f \circ \Phi_\chi^\varepsilon$ rispettivamente all'ordine due e uno, anziché uno e zero, si trova subito per questo termine un'espressione del tipo

$$\varepsilon^2 f^{(2)}(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^3),$$

con

$$f^{(2)} = \frac{1}{2} \{ \{h, \chi\}, \chi \} + \{f, \chi\} = \frac{1}{2} \{f + f_0, \chi\} \quad (4.28)$$

(si è tenuto conto che $\{h, \chi\} = f_0 - f$).

C. La costruzione risonante

Consideriamo ora il caso in cui, per un qualche $k^* \in \mathcal{K}$, risulta $k^* \cdot \omega = 0$. In tal caso la (4.17) non ha soluzione. Possiamo tuttavia ancora ottenere un successo parziale, eliminando dall'espressione (4.25) di χ i termini con k parallelo a k^* ; in questo modo alcune componenti di Fourier di f sono rimosse all'ordine ε , mentre altre (quelle con $k \parallel k^*$) sopravvivono, e la nuova hamiltoniana ha la forma, più debole della (4.26),

$$\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) = h(\tilde{I}) + \varepsilon g(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^2), \quad \text{con} \quad g(\tilde{I}, \tilde{\varphi}) = \sum_{k \in \mathcal{K}, k \parallel k^*} \hat{f}_k(\tilde{I}) e^{ik \cdot \tilde{\varphi}}. \quad (4.29)$$

Questa espressione è detta *forma normale risonante* di H_ε , relativa alla risonanza $k^* \cdot \omega = 0$. Ad esempio, se $k^* = (0, 1)$, cioè $\omega = (\omega_1, 0)$ e $h(I) = \omega_1 I_1$, allora

$$\tilde{H}_\varepsilon = h(\tilde{I}) + \varepsilon g(\tilde{I}, \tilde{\varphi}_2) + \mathcal{O}(\varepsilon^2);$$

all'ordine ε l'hamiltoniana dipende ancora dall'angolo φ_2 , dunque \tilde{I}_1 resta quasi costante fino al tempo $t = 1/\varepsilon$, mentre \tilde{I}_2 nello stesso intervallo di tempo evolve, in generale, in modo significativo.

Per k^* generico si vede subito che le singole azioni evolvono significativamente già a $t \sim 1/\varepsilon$, muovendosi (quasi esattamente) in direzione di k^* : infatti, come si verifica immediatamente scrivendo le equazioni di Hamilton, $\dot{\tilde{I}}$ è parallelo a k^* a meno di $\mathcal{O}(\varepsilon^2)$. Tuttavia una combinazione lineare delle azioni, precisamente l'hamiltoniana imperturbata $h(\tilde{I}) = \omega \cdot \tilde{I}$, su questa stessa scala di tempo $1/\varepsilon$ (quasi) si conserva: infatti (trascurando i termini $\mathcal{O}(\varepsilon^2)$) si ha $\dot{h} = \omega \cdot \dot{\tilde{I}}$, ma $\dot{\tilde{I}} \parallel k^*$ e dunque h è costante. Il fatto che le azioni evolvano mostra che ora il principio della media, a causa della risonanza, non appare più soddisfatto.³⁴

La *forma normale risonante* tuttavia ugualmente è molto utile, perché mostra che (fintantoché si può trascurare il resto di ordine ε^2) nel sistema c'è una quantità conservata in aggiunta all'energia, con tutte le implicazioni che ne seguono (in particolare, il sistema troncato è ancora integrabile; si rifletta su questo punto delicato).

Il caso di k^* generico, volendo, si riporta al caso $k^* = (0, 1)$, solo apparentemente particolare, con un cambiamento di variabili del tipo introdotto nell'appendice E in connessione all'esercizio 31.

- La costruzione risonante si può interpretare da un punto di vista algebrico nel modo seguente: nel caso risonante, l'operatore L_ω ha un nucleo più grande, comprendente le funzioni il cui

³⁴Appare: ovvero non è più garantito che il principio sia soddisfatto, e tutto lascia supporre che in generale non lo sia; ma il fallimento del metodo che stiamo impiegando non implica per sé che il principio non sia soddisfatto.

sviluppo di Fourier contiene, oltre alle funzioni delle sole azioni, le armoniche risonanti, cioè con k tale che $k \cdot \omega = 0$. L'equazione che si può risolvere in questo caso è

$$L_\omega \chi = f - \Pi_N f ,$$

e la nuova hamiltoniana corrispondentemente risulta essere

$$\tilde{H} = h + \varepsilon \Pi_N f + \mathcal{O}(\varepsilon^2) .$$

Queste espressioni, si osservi, unificano il caso risonante e quello non risonante, che differiscono solo per il diverso nucleo N .

D. Generalizzazione: $n > 2$

Per quanto riguarda la generalizzazione al caso $n > 2$, basta osservare che non abbiamo mai veramente usato alcuna peculiarità di $n = 2$ (si ripercorra quanto visto sopra), se non al momento di interpretare la risonanza come condizione di razionalità del rapporto ω_1/ω_2 . Anche per $n > 2$ si ottiene allora che, se ω soddisfa la condizione di non risonanza

$$k_1 \omega_1 + \cdots + k_n \omega_n \neq 0 \quad \text{per} \quad k \in \mathcal{K} \setminus \{0\} ,$$

allora la dipendenza dagli angoli si rimuove per intero al primo ordine in ε e la nuova hamiltoniana assume la forma normale non risonante (4.26); la proposizione 20 si estende pertanto in modo naturale a n qualunque.

Se invece sono presenti una o più relazioni di risonanza

$$k \cdot \omega = 0 \quad \text{per} \quad k \in \mathcal{K} \setminus \{0\} ,$$

allora una parziale dipendenza dagli angoli necessariamente sopravvive all'ordine ε . La situazione diventa variegata a seconda del numero di risonanze presenti; qualche informazione in più si trova nei prossimi esercizi.

ESERCIZIO 41 *Si dimostri che l'insieme dei $k \in \mathbb{Z}^n$ tali che $k \cdot \omega = 0$ è un sottogruppo \mathcal{R} di \mathbb{Z}^n .*

Il sottogruppo \mathcal{R} cui si fa cenno nell'esercizio è detto *reticolo risonante* di ω ; la sua dimensionalità, ovvero il numero di relazioni di risonanza indipendenti, è detto *molteplicità* della risonanza.

ESERCIZIO 42 *Per $n = 3$ si faccia un esempio di ω con una risonanza semplice e uno con risonanza doppia. Come è fatta la forma normale risonante in questi esempi? E in generale? Quali combinazioni lineari $\mathcal{J} = u \cdot I$ (quasi) si conservano per $t \sim 1/\varepsilon$, nonostante la presenza degli angoli nella forma normale all'ordine ε ? [Risposta: per \mathcal{R} generico si ha*

$$g(I, \varphi) = \sum_{k \in \mathcal{R}} \hat{f}_k(I) e^{ik \cdot \varphi} ;$$

si conservano le combinazioni $\mathcal{J} = u \cdot I$ per $u \perp \mathcal{R}$, in particolare si conserva l'energia imperturbata $\omega \cdot I$.]

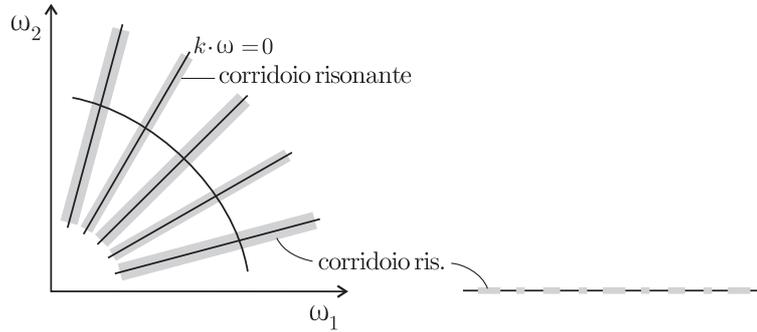


Figura 19: I “corridoi risonanti” per $n = 2$.

E. Generalizzazione: f non polinomiale.

La generalizzazione a f con un numero infinito di componenti di Fourier è molto meno ovvia e anzi assai delicata. Se ripercorriamo lo stesso ragionamento fatto sopra per f polinomio di Fourier, senza difficoltà arriviamo alle equazioni (4.22) e (4.23), quest’ultima però per infiniti $k \in \mathbb{Z}^n$ e non più per un numero finito di essi. Per risolvere la (4.23) per ogni singolo k è sufficiente la condizione di non risonanza

$$k \cdot \omega \neq 0 \quad \forall k \in \mathbb{Z}^n \setminus \{0\} ,$$

che per $n = 2$ equivale a ω_1/ω_2 irrazionale. Ma poi ci sono problemi non banali di convergenza della serie di Fourier per χ

$$\chi(I, \varphi) = \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{\hat{f}_k}{ik \cdot \omega} e^{ik \cdot \varphi} .$$

Infatti i denominatori $k \cdot \omega$, benché tutti diversi da zero, in assenza di ulteriori ipotesi su ω , per opportuno k grande possono avvicinarsi arbitrariamente a zero, impedendo la convergenza della serie a una funzione regolare.

Proseguire l’analisi del problema ci condurrebbe in questioni delicate di tipo aritmetico, nelle quali in queste note non è sensato addentrarsi. Al solo scopo di far comprendere come sia sottile la questione, possiamo ricordare un’ipotesi tipica su ω che, congiuntamente a ipotesi di regolarità su f , consente di giungere a risultati *molto* interessanti (ben più che costruire una forma normale di primo ordine in ε). Tale ipotesi, nota in letteratura come “condizione diofantea” di Siegel (1942), è che $k \cdot \omega$ non solo sia diverso da zero, ma in qualche modo sia staccato via da zero, precisamente che esista una costante $\gamma > 0$ tale che

$$|k \cdot \omega| \geq \frac{\gamma}{|k|^n} \quad \forall k \in \mathbb{Z}^n , k \neq 0 , \tag{4.30}$$

ove si è denotato

$$|k| = |k_1| + \dots + |k_n| .$$

Limitiamoci per semplicità a $n = 2$. La condizione è violata con evidenza da un insieme denso di frequenze $\omega \in \mathbb{R}^2$: tutte quelle che stanno sulle rette del piano ω_1, ω_2 a pendenza razionale (figura 19); per esse infatti con opportuno k si ha $k_1\omega_1 + k_2\omega_2 = 0$ esattamente. Ma variando di poco ω certamente si ha $|k \cdot \omega| < \gamma/|k|^2$, dunque la condizione è violata in un insieme, oltre che denso, aperto (i corridoi grigi in figura 19). Ciononostante, se prendiamo una palla B_R di raggio R nello spazio delle frequenze centrata nell’origine, e rimuoviamo da essa i corridoi, l’insieme $B_{R,\gamma}$ che resta (pur con interno vuoto) ha misura di Lebesgue grande per γ piccolo: precisamente,

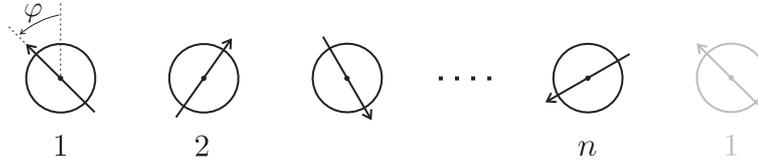


Figura 20: Il modello dei rotatori

PROPOSIZIONE 21 *Esiste $C > 0$ tale che*

$$\text{mes}(B_R \setminus B_{R,\gamma}) < C R^{-1} \gamma \text{mes}(B_R) . \quad (4.31)$$

La dimostrazione, non difficile, si ottiene calcolando il volume del corridoio corrispondente a un generico $k \in \mathbb{Z}^2$, e poi sommando su tutti i $k \in \mathbb{Z}^2$.

ESERCIZIO 43 *Si consideri l'intervallo $(0, 1)$. Come è fatto un suo sottoinsieme aperto denso, ma di misura piccola? Si facciano degli esempi (tale insieme corrisponde all'insieme dei corridoi, tagliati da una curva regolare come in figura 19).*

F. *La forma normale del secondo ordine in ε (cenno).*

Una domanda spontanea, a questo punto, è se dopo aver messo il sistema in forma normale fino all'ordine ε non si possa fare di meglio, ovvero passare all'ordine $\varepsilon^2, \varepsilon^3 \dots$ ottenendo per un qualche $r > 1$ la *forma normale di ordine r*

$$\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) = h(\tilde{I}) + \varepsilon g_1(\tilde{I}) + \dots + \varepsilon^r g_r(\tilde{I}) + \mathcal{O}(\varepsilon^{r+1}) .$$

L'interesse è ovvio: se così avviene, si ottiene un buon controllo sulle azioni per tempi più lunghi, $t \sim 1/\varepsilon^r$. La questione non è facile, e ci limitiamo a un breve cenno relativo a $r = 2$, limitatamente al caso in cui f è un polinomio di Fourier.

Con evidenza, il lavoro da fare è analogo al precedente: l'hamiltoniana su cui lavorare (abolendo la tilde) è ora

$$H_\varepsilon(I, \varphi) = h(I) + \varepsilon f_0(I) + \varepsilon^2 f'(I, \varphi) + \mathcal{O}(\varepsilon^3) ,$$

con f' dato dalla (4.28); ma poiché sia f che χ sono polinomi di Fourier, è un polinomio di Fourier (benché di grado più elevato) anche f' , e il gioco si ripete: se non intervengono nuove risonanze, si trova una trasformazione canonica $(I, \varphi) = \Phi_\chi^{\varepsilon^2}(\tilde{I}, \tilde{\varphi})$, piccola con ε^2 , che elimina tutte le componenti di Fourier di f' , e dà alla nuova hamiltoniana $\tilde{H} = H \circ \Phi_\chi^{\varepsilon^2}$ la “forma normale del secondo ordine”

$$\tilde{H}(\tilde{I}, \tilde{\varphi}) = h(\tilde{I}) + \varepsilon f_0(\tilde{I}) + \varepsilon^2 f'_0(\tilde{I}) + \mathcal{O}(\varepsilon^3) ,$$

ove si è denotata con f'_0 la media di f' . In presenza di risonanze si ottiene invece una forma normale più debole.

4.5 Sistemi non isocroni; il modello dei rotatori

Studieremo qui il caso di sistemi (strettamente) non isocroni, precisamente sistemi hamiltoniani quasi integrabili della consueta forma (4.13), assumendo però, al posto della (4.14), che le frequenze $\omega = \frac{\partial h}{\partial I}$ dipendano in modo sostanziale dalle azioni, precisamente che risulti

$$\det \left(\frac{\partial \omega_i}{\partial I_j} \right) = \det \left(\frac{\partial^2 h}{\partial I_i \partial I_j} \right) \neq 0 . \quad (4.32)$$

In queste condizioni si ha un diffeomorfismo locale tra lo spazio delle frequenze e lo spazio delle azioni, che associa varietà regolari in uno spazio a varietà regolari nell'altro. In particolare, *ai piani risonanti* $k \cdot \omega = 0$ nello spazio delle frequenze restano associate varietà risonanti di equazione

$$k \cdot \omega(I) = 0$$

nello spazio delle azioni. Le frequenze — è bene sottolinearlo — non sono più parametri del modello, non dobbiamo distinguere tra modelli risonanti e non risonanti: *le risonanze ora sono luoghi nello spazio delle azioni*.

Non studieremo il caso generale (cui tuttavia accenneremo più avanti), ma studieremo in dettaglio un modello particolare, una catena di rotatori debolmente accoppiati, che come vedremo racchiude l'essenziale. L'hamiltoniana del modello è

$$H(I, \varphi) = \sum_{j=1}^n \frac{I_j^2}{2A} + \varepsilon f(\varphi) \quad (4.33)$$

con, nel caso più semplice,

$$f(\varphi) = - \sum_{j=1}^n \cos(\varphi_{j+1} - \varphi_j), \quad \varphi_{n+1} = \varphi_1$$

(condizioni periodiche agli estremi). Si ha $\omega_i(I) = A^{-1}I_i$, mentre l'hessiana è $h''(I) = A^{-1}\mathbb{I}$; le curve risonanti sono così piani per l'origine anche nello spazio delle azioni.

La perturbazione, si osservi, ha un numero finito di componenti di Fourier, inoltre queste sono del tipo particolare $\pm k^j$ ove

$$k^j = (0, \dots, 0, -1, 1, 0, \dots, 0), \quad (4.34)$$

con la coppia $-1, 1$ in posizione $j-1, j$. Il carattere locale dell'interazione aiuta a comprendere meglio i risultati, ma non è realmente importante, come anche il fatto che f sia indipendente dalle azioni e abbia media nulla. La forma della parte imperturbata

$$h = \sum_{j=1}^n \frac{I_j^2}{2A}$$

invece è importante: comprendere fino a che punto la si può generalizzare non è banale.

Nel seguito, per alleggerire le espressioni, porremo³⁵ $A = 1$. Come dominio nello spazio delle azioni prenderemo la palla B_R di raggio R .

Due osservazioni preliminari:

- (i) Le condizioni periodiche hanno il vantaggio che tutti i siti sono equivalenti, lo svantaggio che c'è una quantità conservata, il momento totale $J = I_1 + \dots + I_n$. Non è tuttavia restrittivo supporre $J^o = J(0) = 0$: infatti per ogni dato iniziale I^o con $J^o \neq 0$ possiamo eseguire la trasformazione w dipendente dal tempo

$$I_j = \frac{J^o}{n} + \tilde{I}_j, \quad \varphi_j = \frac{J^o}{n} t + \tilde{\varphi}_j, \quad j = 1, \dots, n$$

³⁵formalmente è, come al solito, un riscalamento canonico: $I = \sqrt{A}\tilde{I}$, $\varphi = \tilde{\varphi}$, con valenza $1/\sqrt{A}$ che però si può togliere riscalandolo t .

(sistema “corotante” col moto medio; è come mettersi a cavallo del baricentro per i sistemi che conservano il momento lineare); per questa trasformazione (puntuale, $f = 0$) è facile vedere che si ha $K_0(\tilde{I}, \tilde{\varphi}) = -\frac{J^o}{n}(\tilde{I}_1 + \dots + \tilde{I}_n)$, e dunque la nostra nuova hamiltoniana $K = H \circ w + K_0$ coincide con la vecchia a meno di una inessenziale costante additiva. Il dato iniziale \tilde{I}^o corrispondente a I^o ora ha momento totale nullo.

- (ii) E' equivalente considerare il modello (4.33), con il piccolo parametro ε di fronte a f e come si è detto $I \in B_R$, oppure considerare un modello privo di ε di fronte a f ,

$$H(I, \varphi) = \sum_{j=1}^n \frac{I_j^2}{2} + f(\varphi), \quad (4.35)$$

ma in un dominio delle azioni grande con ε , precisamente

$$I \in B_{R/\sqrt{\varepsilon}}.$$

Infatti, il riscalamento

$$I = \frac{\tilde{I}}{\sqrt{\varepsilon}}, \quad \varphi = \tilde{\varphi}$$

riporta il dominio delle azioni a B_R indipendente da ε , e dà alla nuova hamiltoniana $K(\tilde{I}, \tilde{\varphi}) = \sqrt{\varepsilon}H(\tilde{I}/\sqrt{\varepsilon}, \tilde{\varphi})$ la forma

$$K(\tilde{I}, \tilde{\varphi}) = \frac{1}{\sqrt{\varepsilon}} \left[\sum_{j=1}^n \frac{\tilde{I}_j^2}{2} + \varepsilon V(\tilde{\varphi}) \right];$$

il riscalamento di tempo $t = \tilde{t}\sqrt{\varepsilon}$ elimina il fattore³⁶ $1/\sqrt{\varepsilon}$ di fronte a K , e ci si ritrova con l'hamiltoniana $\sqrt{\varepsilon}K$ che è la (4.33).

Allo stesso modo si comprende che se dentro a B_R consideriamo il dominio piccolo

$$I \in B_{R\sqrt{\varepsilon}},$$

la presenza di ε di fronte a f è illusoria: infatti se riportiamo il dominio delle azioni a B_R con il riscalamento, questa volta,

$$I = \sqrt{\varepsilon}\tilde{I}, \quad \varphi = \tilde{\varphi},$$

e passiamo al tempo $\tilde{t} = t\sqrt{\varepsilon}$, il parametro ε scompare del tutto dall'hamiltoniana, che prende la forma (4.35), e dunque il modello non è in alcun modo prossimo a integrabile. Questo è un elemento cruciale, che sarà importante nel seguito e sul quale è bene riflettere. Un aiuto alla riflessione è dato dall'esercizio seguente:

ESERCIZIO 44 *Si prendano in esame i diversi casi considerati al punto (ii): presenza o assenza di ε a fronte di f , raggio r del dominio delle azioni pari a R , $R/\sqrt{\varepsilon}$, $R\sqrt{\varepsilon}$; si raffrontino la velocità di evoluzione degli angoli e quella delle azioni, quest'ultima rapportata al raggio r dove le azioni vivono. [Risposta: sono ben separate da un fattore ε nel caso di debole accoppiamento (azioni molto più lente degli angoli), dello stesso ordine nel caso di forte accoppiamento.]*

³⁶Stiamo usando $\Phi_K^t = \Phi_{\sqrt{\varepsilon}K}^{t/\sqrt{\varepsilon}}$.

Nel seguito del paragrafo applicheremo al modello dei rotatori metodi noti di teoria delle perturbazioni, eseguendo un passo perturbativo fuori risonanza e in risonanza, e poi illustreremo i risultati ottenuti confrontandoli con una semplice indagine numerica.

A. *Studio analitico: un passo perturbativo, caso non risonante.* Eseguiamo un passo perturbativo con il metodo di Lie, ovvero con $w = \Phi_\chi^\varepsilon$. Procedendo in modo *identico* a quanto fatto nel paragrafo 4.4-A, si trova subito che se χ soddisfa l'equazione fondamentale (4.19), che essendo f a media nulla si scrive ora

$$\omega(I) \cdot \frac{\partial \chi}{\partial I} = f ,$$

allora la perturbazione si elimina all'ordine ε e la nuova hamiltoniana assume la forma normale

$$\tilde{H} = h + \varepsilon^2 \tilde{f} + \mathcal{O}(\varepsilon^3) , \quad \tilde{f} = \frac{1}{2} \{f, \chi\} .$$

L'equazione per χ è risolta da

$$\chi(I, \varphi) = - \sum_{j=1}^n \frac{\sin(\varphi_j - \varphi_{j-1})}{I_j - I_{j-1}} ,$$

ovviamente dove i denominatori non si annullano, e con pochi calcoli si trova poi

$$\begin{aligned} \tilde{f}(I, \varphi) &= \frac{1}{2} \sum_{j=1}^n \frac{1 - \cos 2(\varphi_j - \varphi_{j-1})}{(I_j - I_{j-1})^2} + \frac{1}{4} \sum_{j=1}^n \left[\cos(\varphi_{i+1} - \varphi_{j-1}) \right. \\ &\quad \left. - \cos(\varphi_{i+1} - 2\varphi_j + \varphi_{j-1}) \right] \left[\frac{1}{(I_{i+1} - I_j)^2} - \frac{1}{(I_j - I_{j-1})^2} \right] + \mathcal{O}(\varepsilon) , \end{aligned}$$

avendo qui mantenuto la notazione (I, φ) per le nuove variabili. Si osservi che

- (i) in \tilde{f} compaiono nuove armoniche, ma in numero finito, precisamente quelle del tipo

$$k = \pm k^j \pm k^{j-1} ;$$

- (ii) χ contiene i piccoli divisori

$$k^j \cdot \omega(I) = I_j - I_{j-1} ,$$

corrispondentemente \tilde{f} contiene i loro quadrati (com'è ovvio, dovendo fare le derivate di χ rispetto alle azioni). Formalmente dunque la forma normale si può scrivere ovunque tranne che sulle risonanze, che sono qui i piani

$$I_j - I_{j-1} = 0 , \quad j = 1, \dots, n$$

(due rotatori consecutivi girano alla stessa velocità; la combinazione $\varphi_j - \varphi_{j-1}$ non avanza). Ma il guadagno di un ordine, in tutto un intorno $\mathcal{O}(\sqrt{\varepsilon})$ dei piani risonanti, è illusorio. E' infatti un guadagno vero lontano dai piani risonanti, quando i denominatori sono grandi: ma se ci si avvicina ai piani risonanti i denominatori di \tilde{f} rimpiccioliscono e il guadagno progressivamente si perde; *se poi uno dei denominatori diviene piccolo come $\sqrt{\varepsilon}$, la nuova perturbazione ha norma* (qualunque sia la norma scelta) $\varepsilon^2 \|\tilde{f}\| \sim \varepsilon$, *cioè tal quale la perturbazione di partenza, e non si è guadagnato nulla.* (In situazioni intermedie si guadagna qualche cosa, ma meno di un ordine in ε : ad esempio, a distanza $\mathcal{O}(\varepsilon^{1/4})$ dalla risonanza si ha $\varepsilon^2 \|\tilde{f}\| \sim \varepsilon^{3/2}$.) Si rifletta su questo punto: algebricamente abbiamo il fattore ε^2 di fronte alla nuova perturbazione, ma a causa dei denominatori $\mathcal{O}(\varepsilon)$, \dot{I} (per le I coinvolte nella o nelle risonanze) è piccolo come ε , non come ε^2 , come se il passo perturbativo non lo avessimo fatto.

B. Moti entro una risonanza semplice. Supponiamo ora che sia $|I_2 - I_1| = \mathcal{O}(\sqrt{\varepsilon})$ e invece $|I_j - I_{j-1}| = \mathcal{O}(1)$ per $j \neq 2$. La miglior forma normale che siamo in grado di costruire si ottiene sfilando via da χ il termine col denominatore piccolo, ovvero prendendo

$$\chi = - \sum_{j \neq 2} \frac{\sin(\varphi_j - \varphi_{j-1})}{I_j - I_{j-1}} ;$$

corrispondentemente la nuova hamiltoniana assume la *forma normale risonante*

$$\tilde{H}(\tilde{I}, \tilde{\varphi}) = \frac{1}{2}(\tilde{I}_1^2 + \tilde{I}_2^2) - \varepsilon \cos(\tilde{\varphi}_2 - \tilde{\varphi}_1) + \frac{1}{2} \sum_{j>2} \tilde{I}_j^2 + \mathcal{O}(\varepsilon^2) . \quad (4.36)$$

Fintantoché il resto di ordine ε^2 si può trascurare, il moto dei primi due rotatori si disaccoppia dal resto della catena ed è governato dall'hamiltoniana ridotta

$$\mathcal{H}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2) = \frac{1}{2}(\tilde{I}_1^2 + \tilde{I}_2^2) - \varepsilon \cos(\tilde{\varphi}_2 - \tilde{\varphi}_1) . \quad (4.37)$$

Questa è una vecchia conoscenza, infatti altro non è che la (4.9) con notazioni appena differenti; sappiamo allora che il momento complessivo $J = \tilde{I}_1 + \tilde{I}_2$ resta costante, il moto medio si disaccoppia, e il sistema, in buona sostanza, è un pendolo: \tilde{I}_1 e \tilde{I}_2 eseguono oscillazioni di ampiezza $\mathcal{O}(\sqrt{\varepsilon})$, con periodo di ordine $1/\sqrt{\varepsilon}$, attorno alla media $J = \frac{1}{2}(\tilde{I}_1 + \tilde{I}_2)$. Per capire ancor meglio, possiamo supporre — senza perdita di generalità, come si è visto sopra — che sia $J = 0$; in tal caso abbiamo $\tilde{I}_1 \simeq \tilde{I}_2 \simeq \sqrt{\varepsilon}$, e conviene eseguire un riscaldamento (uno *zoom*)

$$\tilde{I}_j = \sqrt{\varepsilon} \hat{I}_j , \quad \tilde{\varphi}_j = \hat{\varphi}_j , \quad t = \frac{1}{\sqrt{\varepsilon}} \hat{t} . \quad (4.38)$$

Il dominio delle variabili riscalate ora è di ordine uno, e la nuova hamiltoniana³⁷ $\hat{\mathcal{H}} = \varepsilon^{-1} \mathcal{H}$ diviene indipendente da ε :

$$\hat{\mathcal{H}} = \frac{1}{2}(\hat{I}_1^2 + \hat{I}_2^2) - \cos(\hat{\varphi}_2 - \hat{\varphi}_1) .$$

Il moto delle variabili $\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2$ sotto l'hamiltoniana ridotta (4.37) altro non è che questo moto indipendente da ε — un rotatore libero corrispondente al moto medio, un pendolo che esegue oscillazioni limitate — semplicemente riscaldato nelle azioni e nel tempo di fattori $\sqrt{\varepsilon}$ e $1/\sqrt{\varepsilon}$. Fintantoché il resto $\mathcal{O}(\varepsilon^2)$ si può trascurare, questo moto dei primi due rotatori si ritrova per la forma normale (4.36), con \tilde{I}_1, \tilde{I}_2 che compiono oscillazioni $\mathcal{O}(\sqrt{\varepsilon})$ nella scala di tempo $\mathcal{O}(1/\sqrt{\varepsilon})$, mentre intanto $\tilde{I}_3, \dots, \tilde{I}_n$ restano costanti e $\tilde{\varphi}_1, \dots, \tilde{\varphi}_n$ avanzano uniformemente. Il tutto si ritrova infine anche per le variabili di partenza, a meno della deformazione piccola con ε dovuta alla trasformazione canonica che ha dato al sistema la forma normale. *Il principio della media è chiaramente soddisfatto.*

In questo caso di risonanza singola la forma normale risonante, trascurando il resto $\mathcal{O}(\varepsilon^2)$, risulta integrabile (facilmente, passando alle variabili di azione–angolo del pendolo che abbiamo incontrato, si scriverebbero tutte le sue variabili di azione–angolo). Non è questo però il punto: come vedremo nel prossimo paragrafo *C*, nel caso di almeno due risonanze (coinvolgenti siti consecutivi) l'integrabilità si perde e i moti si fanno sostanzialmente più complessi; *ugualmente, con identico meccanismo, il principio della media resterà soddisfatto.*

³⁷Un fattore $\varepsilon^{-1/2}$ è dovuto al riscaldamento delle azioni, un altro fattore $\varepsilon^{-1/2}$ è dovuto al riscaldamento del tempo.

C. Moti entro una risonanza multipla. Per capire, è sufficiente comprendere il caso di due risonanze contigue, ad esempio

$$|I_2 - I_1|, |I_3 - I_2| = \mathcal{O}(\sqrt{\varepsilon}); \quad (4.39)$$

la generalizzazione al caso di risonanze multiple poi è immediata.

Procedendo come sopra per il caso di risonanza singola, è facile ottenere una forma normale adattata alla risonanza considerata, precisamente

$$\tilde{H}(\tilde{I}, \tilde{\varphi}) = \mathcal{H}(\tilde{I}, \tilde{\varphi}) + \frac{1}{2} \sum_{j>3} \tilde{I}_j^2 + \mathcal{O}(\varepsilon^2), \quad (4.40)$$

questa volta con (eliminando la tilde)

$$\mathcal{H}(I_1, I_2, I_3, \varphi_1, \varphi_2, \varphi_3) = \frac{1}{2} \sum_{j=1}^3 I_j^2 - \varepsilon [\cos(\varphi_2 - \varphi_1) + \cos(\varphi_3 - \varphi_2)]. \quad (4.41)$$

Come sopra, ignoriamo il resto $\mathcal{O}(\varepsilon^2)$ e analizziamo il moto di \mathcal{H} . Possiamo supporre anche qui che il momento angolare totale $I_1 + I_2 + I_3$ sia nullo; ma allora si ha $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\varepsilon}$, e il riscaldamento (4.38) dà l'hamiltoniana

$$\hat{\mathcal{H}} = \frac{1}{2} \sum_{j=1}^3 \hat{I}_j^2 - [\cos(\hat{\varphi}_2 - \hat{\varphi}_1) + \cos(\hat{\varphi}_3 - \hat{\varphi}_2)]$$

in un dominio indipendente da ε . Questo è un sistema hamiltoniano fortemente accoppiato indipendente da ε , con due gradi di libertà effettivi, che non è integrabile né prossimo a integrabile, non esegue moti quasi periodici ma è in grado di eseguire moti profondamente diversi e sostanzialmente più complessi, detti comunemente *moti caotici*.³⁸ Nel prossimo paragrafo, dedicato all'esplorazione numerica del modello dei rotatori, ne vedremo degli esempi. Si osservi però che *grazie alla conservazione dell'energia, tali moti sono limitati* (infatti, la parte cinetica è definita positiva mentre la parte potenziale è limitata) *e si sviluppano su una loro scala di azione e di tempo, che evidentemente non dipende da ε* . Corrispondentemente, *per ε piccolo, \mathcal{H} ammette moti limitati* (in generale complicati, comunque piccolo sia ε) *che si sviluppano su una scala di azione $\sqrt{\varepsilon}$ e su una scala di tempo $1/\sqrt{\varepsilon}$* . Fintantoché il resto $\mathcal{O}(\varepsilon^2)$ si può ignorare, i moti caotici esistono anche nel sistema di partenza, precisamente appaiono come *moti caotici localizzati entro la risonanza doppia, limitati da una disuguaglianza del tipo*

$$|I_j(t) - I_j^o| < \text{cost} \sqrt{\varepsilon},$$

che si sviluppano sulla scala di tempo $1/\sqrt{\varepsilon}$.

In questo modello particolare, il comportamento caotico è localizzato nella catena di rotatori, nel senso che coinvolge tre soli di essi mentre gli altri avanzano regolarmente (a meno di un piccolo effetto dovuto al fatto che \tilde{I}_j e I_j non coincidono). Come annunciato sopra, *pur in una situazione sostanzialmente più complessa della precedente, resta soddisfatto il principio della media.*

³⁸Qui ci accontentiamo dell'intuizione. Tutto però si precisa in modo pulito e si dimostra nell'ambito della teoria dei sistemi dinamici; lo stesso termine *caotico*, qui usato in modo necessariamente vago in contrapposizione a integrabile, si formalizza in un termine tecnico, e lo si usa, in buona sostanza, per quei sistemi in cui la divergenza esponenziale delle traiettorie vicine — quella che si ha quando la "stima a priori" non è migliorabile — interessa un insieme di dati iniziali di misura positiva. A questa instabilità esponenziale si accompagna una sostanziale imprevedibilità dei moti.

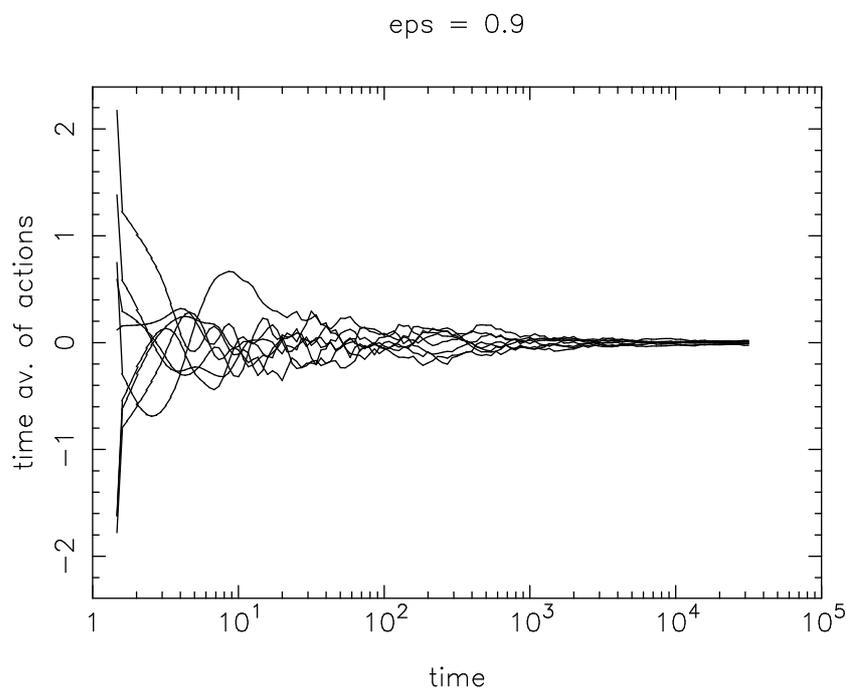


Figura 21: $\bar{I}_j(T)$ in funzione di T , per $\varepsilon = 0.9$.

D. Studio numerico del modello dei rotatori: le quantità calcolate. Lo studio numerico di questo o qualunque altro sistema dinamico, hamiltoniano o no, si effettua risolvendo approssimativamente le equazioni del moto mediante un opportuno algoritmo in cui si discretizza il tempo e si procede a “passi di integrazione” finiti. La letteratura sull’argomento è *enorme*, e in parte intrecciata con lo studio perturbativo.³⁹ Trarre conclusioni chiare sull’affidabilità dei metodi di integrazione numerica è difficile; il loro impiego richiede esperienza e sensibilità. Noi naturalmente non entreremo in nessun modo in questa discussione, e con atteggiamento, per una volta, non critico, prenderemo per buoni i risultati numerici senza ulteriori indagini.

Studieremo qui una catena di $n = 8$ rotatori, con hamiltoniana (4.33). Le quantità che andremo ad esaminare sono, per ciascun moto,

- (a) La *media temporale* delle azioni fino al tempo T , precisamente

$$\bar{I}_j(T) = \frac{1}{T} \int_0^T I_j(t) dt .$$

In pratica, se s è la discretizzazione temporale che interviene nell’intergrazione numerica, quello che si calcola è

$$\bar{I}_j(Ns) = \frac{1}{N} \sum_{l=0}^{N-1} I_j(ls) .$$

³⁹Si intuisce forse che la discretizzazione si può riguardare come una perturbazione delle equazioni del moto originarie, anche se non è facile precisare in che senso.

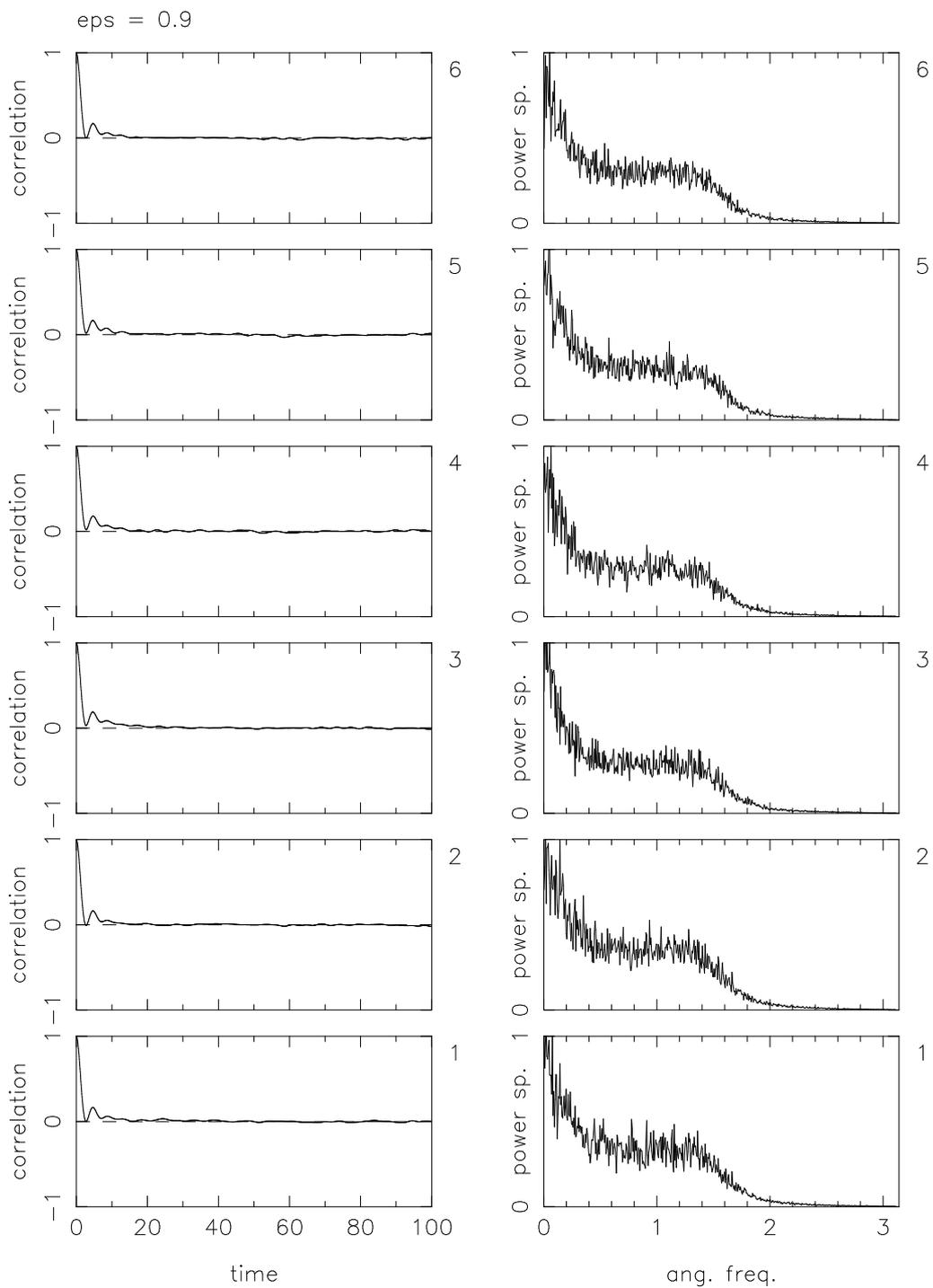


Figura 22: γ_j (sinistra) e W_j (destra), $j = 1, \dots, 6$ dal basso in alto, per $\varepsilon = 0.9$.

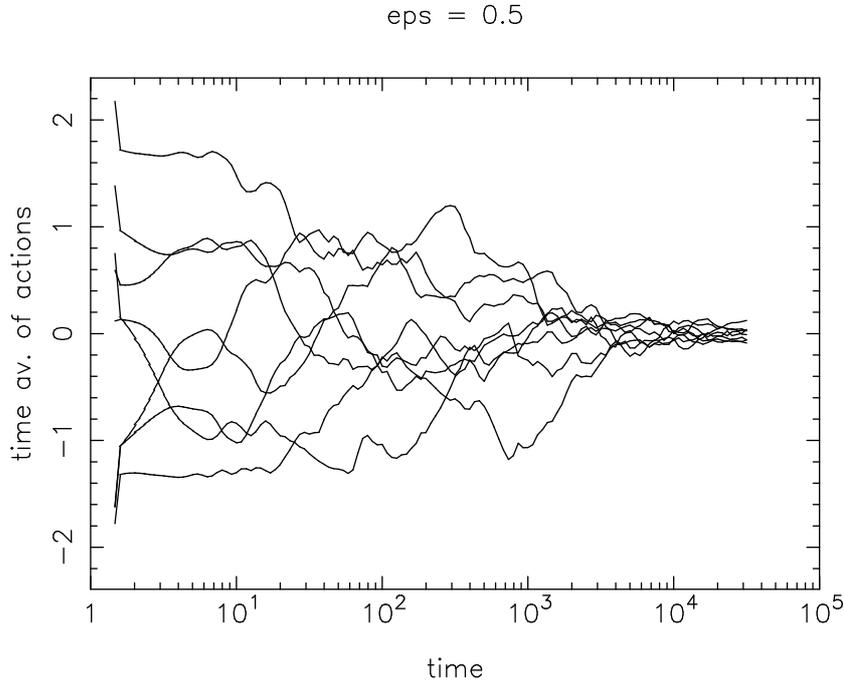


Figura 23: Come in figura 21, per $\varepsilon = 0.5$.

In linea di principio si vorrebbe calcolare la media infinita⁴⁰ $\overline{I_j}(\infty) = \lim_{T \rightarrow \infty} \overline{I_j}(T)$; in pratica ovviamente il calcolo è troncato a $T_{\max} = sN_{\max}$ grande. In condizioni di forte interazione (ε di ordine uno) ci si aspetta che tutti i rotatori si comportino, su tempi lunghi, allo stesso modo, dunque che la media $\overline{I_j}(\infty)$ sia la stessa per tutti e, in assenza di moto medio, sia nulla.⁴¹ Viceversa, in condizioni di debole accoppiamento (ε piccolo, assenza di risonanze) ci si aspetta che ogni rotatore vada sostanzialmente per conto proprio, mantenendo, a meno di piccole oscillazioni, la I_j di partenza.

(b) La *funzione di autocorrelazione* $\gamma_j(\tau)$ delle I_j , definita nel modo seguente:

$$\gamma_j(\tau) = \frac{\Gamma_j(\tau)}{\sigma^2}, \quad \Gamma_j(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I_j(t + \tau) I_j(t) dt - \overline{I_j}(\infty)^2,$$

dove $\sigma^2 = \Gamma_j(0)$ è la varianza temporale. In pratica il calcolo è discretizzato e troncato come

⁴⁰Un teorema dovuto a Birkhoff garantisce l'esistenza delle medie temporali infinite per qualunque funzione (misurabile) per moti generici, precisamente con la possibile eccezione di un insieme di dati iniziali di misura di Lebesgue nulla. È un teorema non banale di grande importanza.

⁴¹Un tale comportamento è anche quello che prevede la Meccanica Statistica, nella quale alla media temporale si sostituisce una opportuna distribuzione di probabilità (la cosiddetta *misura microcanonica*, uniforme sulla superficie di energia costante, per sistemi che conservino l'energia); per la simmetria del modello, tutti i rotatori non possono che avere comportamento statistico identico.

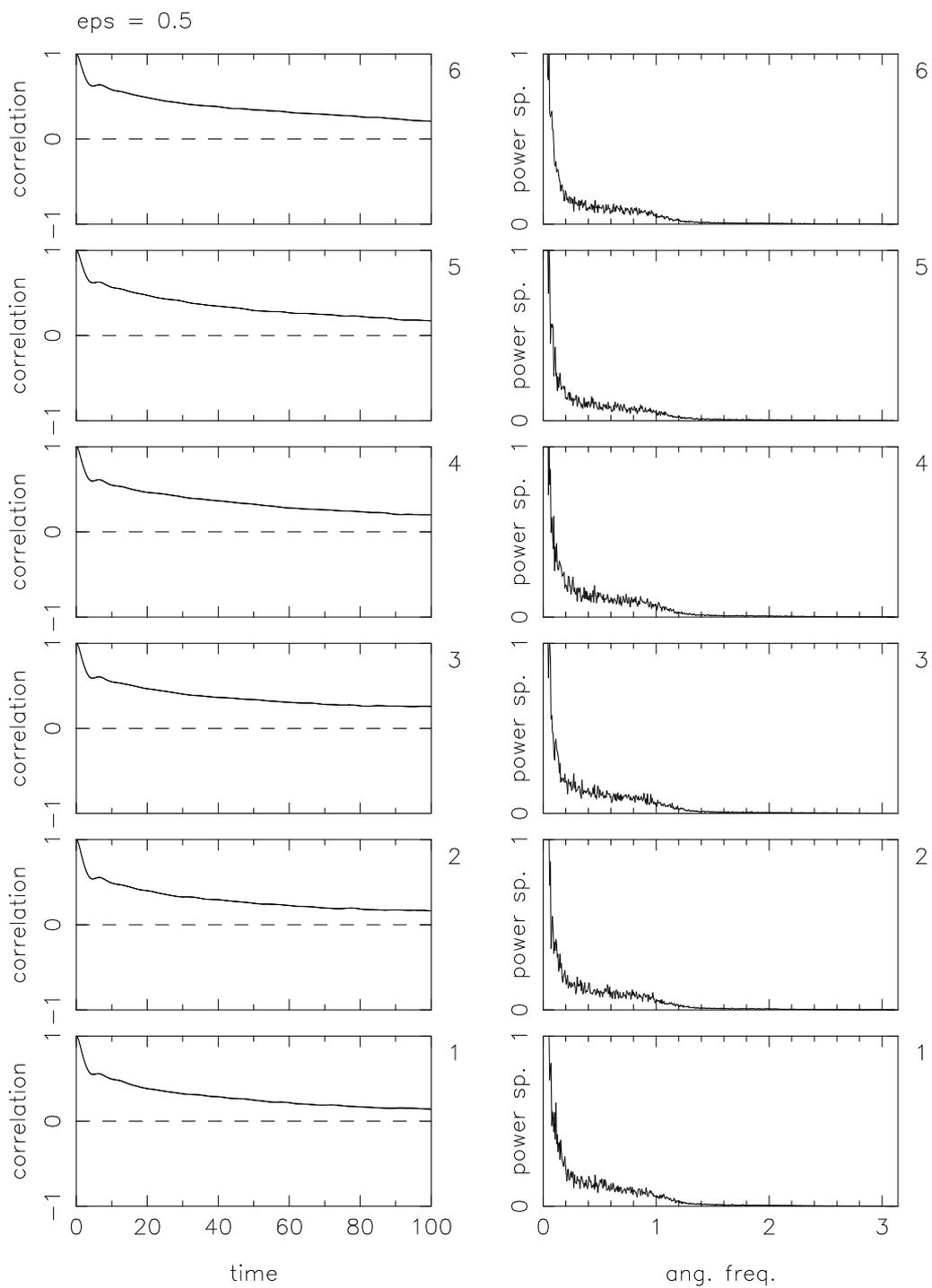


Figura 24: Come in figura 22, per $\varepsilon = 0.5$.

sopra. Per una generica $f(t)$, la funzione di autocorrelazione⁴²

$$\gamma_f(\tau) = \frac{\Gamma_f(\tau)}{\sigma^2}, \quad \Gamma_f(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t+\tau)f(t) dt - \bar{f}(\infty)^2,$$

rappresenta la “memoria” al tempo $t + \tau$ del valore di f al tempo t : se $\gamma_f(\tau)$ è positiva (negativa) vuol dire che tipicamente, se $f(t)$ sta sopra o sotto la media temporale $\bar{f}(\infty)$, lo stesso (l’opposto) fa $f(t + \tau)$, ovvero una misura di f al tempo t dà qualche informazione su f al tempo $t + \tau$. Se $\gamma_f(\tau) = 0$, il comportamento di f al tempo $t + \tau$ è invece indipendente (scorrelato) dal comportamento di f al tempo t . In Meccanica Statistica, il tempo di decadimento a zero (delle funzioni in qualche modo importanti) è identificato col tempo di raggiungimento dell’equilibrio.

Per le nostre I_j , in condizioni di forte accoppiamento, ci si aspetta che $\gamma_j(\tau)$ decada a zero su una qualche scala di tempo non troppo lunga; in condizioni di debole accoppiamento ci si aspetta al contrario che $\gamma_j(\tau)$ resti diversa da zero per tempi τ lunghi.

(c) Lo spettro di potenza W_j di I_j , la cui definizione è la seguente:⁴³ se

$$I_j(t) = \int_{-\infty}^{\infty} J_j(\nu) e^{i\nu t} d\nu,$$

allora $W_j(\nu) = |J(\nu)|^2$. Anche qui naturalmente si discretizza e si tronca. Lo spettro di potenza risulta essere la trasformata di Fourier della funzione di correlazione, pertanto esso non fornisce un’informazione indipendente, ma è ugualmente utile prenderlo in considerazione.

In condizione di debole accoppiamento il moto di ciascuna j ci si aspetta sia, con buona approssimazione, quasi periodico; corrispondentemente lo spettro di potenza sarà “a righe”, cioè discreto con un numero finito di frequenze. Viceversa, nel caso di forte accoppiamento e moto caotico, ci si aspetta di trovare spettro continuo (infinite frequenze in gioco).

E. I risultati numerici. Tutti i risultati che esporremo si riferiscono a dati iniziali con momento totale $I_1 + \dots + I_n = 0$.

La figura 21 mostra il comportamento delle medie $\bar{I}_j(T)$ in funzione di T (su scala logaritmica), per grande accoppiamento, precisamente per $\varepsilon = 0.9$. Il risultato è consistente con l’idea che il sistema, in queste condizioni, abbia “un buon comportamento statistico”; come atteso si ha $\bar{I}_j(\infty) = 0$ per ogni j . La figura (22) mostra nelle stesse condizioni, più precisamente per lo stesso movimento, la funzione di correlazione (a sinistra) e lo spettro di potenza (a destra); per ragioni grafiche la figura si limita a soli sei rotatori. Come si vede, tutte le correlazioni decadono rapidamente, in poche unità di tempo (tutte le velocità angolari sono di ordine uno, perciò poche unità di tempo vuol dire pochi giri di ciascun rotatore). Si osservi che lo spettro ha l’aspetto di

⁴²Si possono considerare correlazioni incrociate di due funzioni f e g , con ovvia definizione

$$\gamma_{f,g} = \frac{\Gamma_{f,g}(\tau)}{\sigma^2}, \quad \Gamma_{f,g}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t+\tau)g(t) dt - \bar{f}(\infty)\bar{g}(\infty).$$

Quando $g = f$ si usa il termine “autocorrelazione” (spesso sostituito, nella pratica, dal semplice “correlazione”).

⁴³La comprensione qui richiede un minimo di familiarità con la *trasformata di Fourier*, generalizzazione della serie di Fourier quando l’intervallo di definizione delle funzioni è tutto \mathbb{R} , e corrispondentemente la frequenza, qui denotata ν , anziché essere multipla di una frequenza fondamentale è una variabile continua, $\nu \in \mathbb{R}$. Chi non avesse tale familiarità può senza danno saltare qui e più avanti tutte le considerazioni relative allo spettro di potenza.

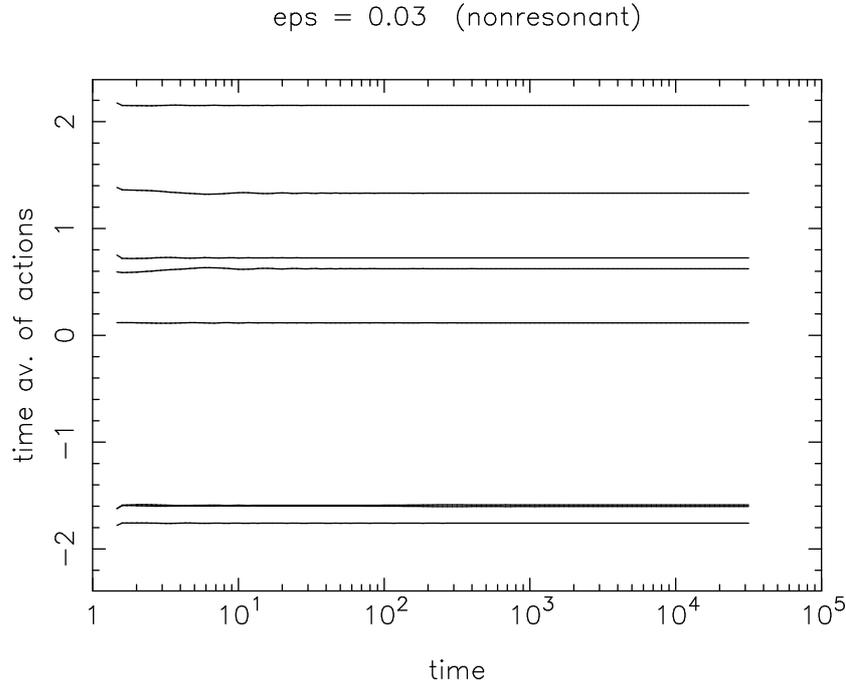


Figura 25: Come nelle figure 21 e 23, per $\varepsilon = 0.03$, con dati iniziali non risonanti.

una banda continua, che si estende da zero a un valore ν_{\max} prossimo a π , e che tutti i rotatori si comportano allo stesso modo. Il moto, con evidenza, non ha nulla a che vedere con un moto quasi periodico.

Le figure (23) e (24) riportano le medesime quantità per accoppiamento più basso, $\varepsilon = 0.5$. L'aspetto qualitativo non è molto diverso, ma la scala di tempo sulla quale si manifestano i diversi fenomeni si è dilatata; corrispondentemente la banda continua dello spettro è più stretta (le frequenze in gioco sono più basse). Le figure successive (25) e (26) si riferiscono invece a un valore di ε sensibilmente più basso, precisamente $\varepsilon = 0.03$, con dati iniziali scelti in modo da evitare la presenza di risonanze. La situazione ora è alquanto diversa: le azioni appaiono stabili, le correlazioni non decadono, lo spettro si mostra discreto.⁴⁴

Le successive figure (27) e (28) si riferiscono al caso risonante, precisamente alla risonanza doppia $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\varepsilon}$, con dati iniziali tali che $I_1 + I_2 + I_3 = 0$; il valore di ε è ancora 0.03. Con evidenza i rotatori 1, 2 e 3 formano un sottosistema separato interagente: la media delle loro azioni è nulla, la correlazione decade, lo spettro di potenza mostra una banda stretta ma continua, ben diversa dalle righe del caso precedente; nel contempo gli altri rotatori avanzano regolarmente, sostanzialmente come nel caso precedente. La funzione di correlazione, ad esempio tramite il suo primo zero, fornisce una elementare scala di tempo associata al moto caotico. In base all'analisi teorica sopra riportata, questo tempo deve riscalare come $1/\sqrt{\varepsilon}$, mentre corrispondentemente l'ampiezza della banda in W_j deve risultare proporzionale a $\sqrt{\varepsilon}$. Una verifica di questa proprietà di riscaldamento — rilevante, perché coglie l'aspetto più essenziale del moto risonante — si può fare riducendo ε , ad esempio, di un fattore 4, e controllando che il tempo di decorrelazione e l'ampiezza della banda continua

⁴⁴Si potrebbe vedere che le linee nello spettro del j -esimo rotatore sono esattamente a $\nu_j^+ = I_{j+1} - I_j$ e $\nu_j^- = I_j - I_{j-1}$; sono queste in effetti le frequenze che intervengono nella trasformazione canonica che determina l'oscillazione di I_j attorno a \tilde{I}_j (quasi costante).

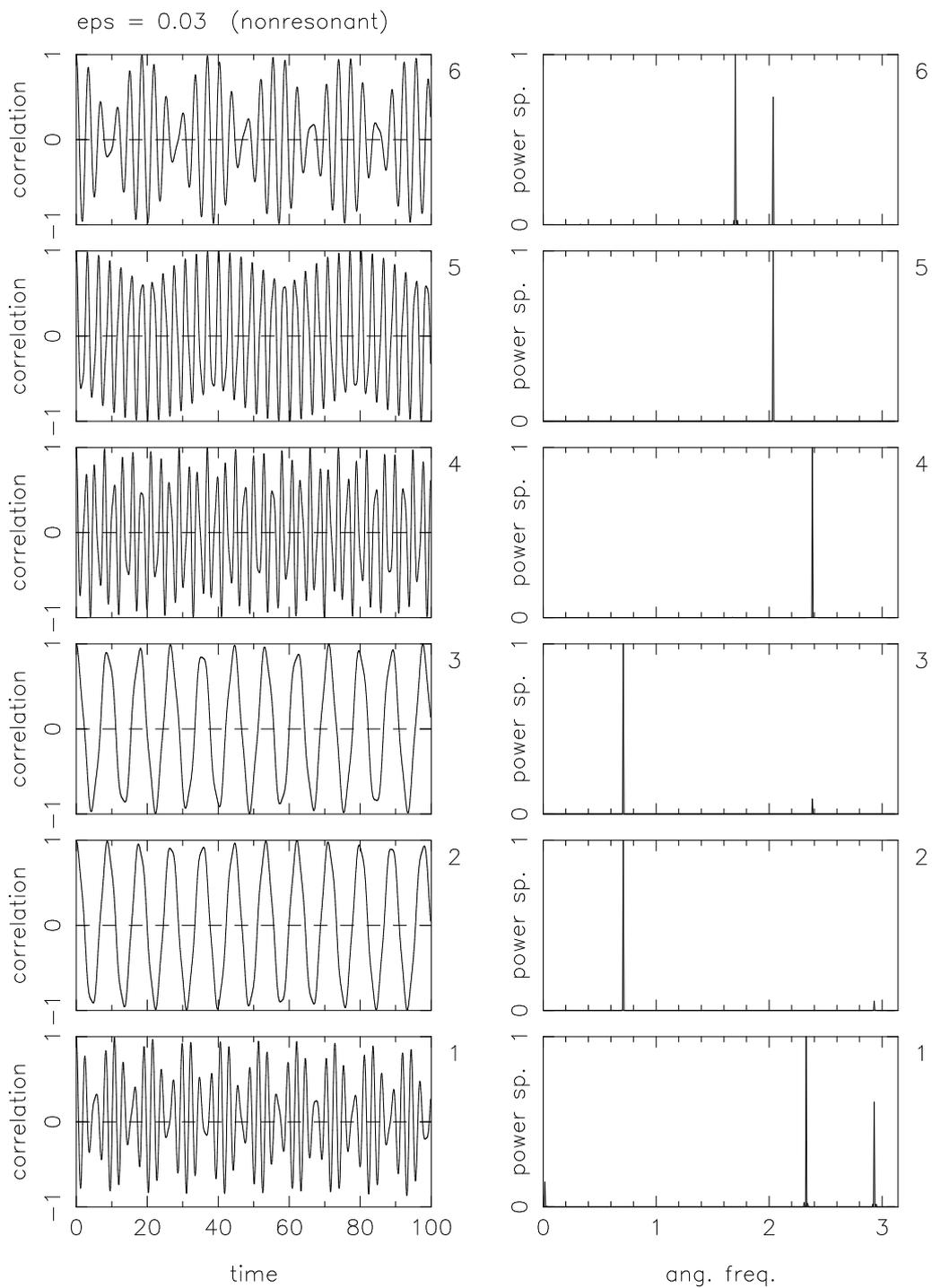


Figura 26: Come nelle figure 22 e 24, per $\epsilon = 0.03$, con dati iniziali risonanti.

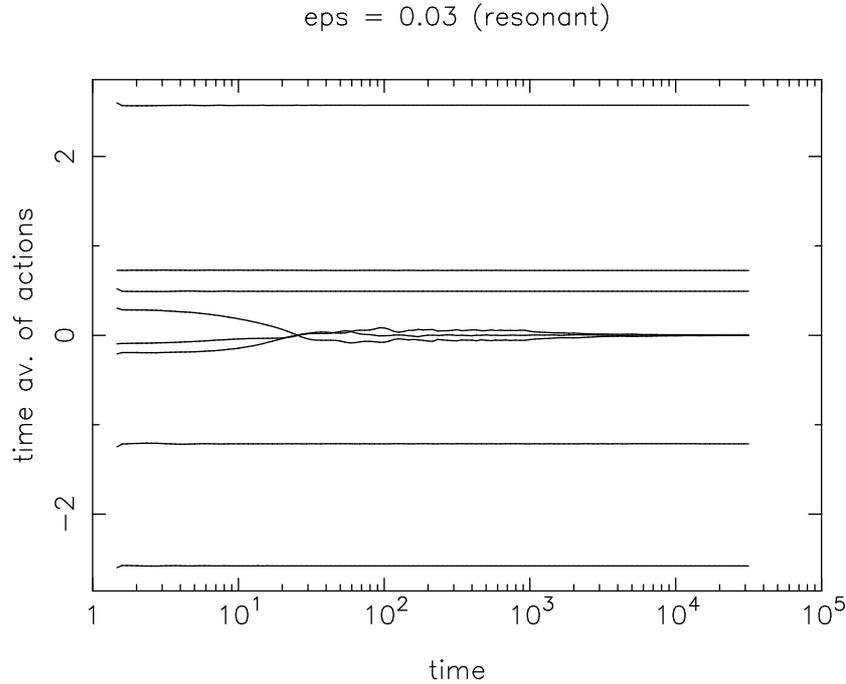


Figura 27: Come nella figura 22, per il medesimo $\varepsilon = 0.03$, ma dentro alla risonanza $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\varepsilon}$.

siano rispettivamente aumentati e diminuiti di un fattore 2. La verifica è riportata in figura 29. La parte superiore si riferisce a $\varepsilon = 0.03$, quella inferiore a $\varepsilon = 0.0075$. Per quest'ultimo valore di ε la scala di tempo è raddoppiata rispetto al caso $\varepsilon = 0.03$, mentre la scala delle frequenze è dimezzata; come si vede, le figure si assomigliano in modo significativo (per rendere più leggibili le figure con lo spettro, la scala delle frequenze è stata ulteriormente dilatata di un fattore cinque, allo stesso modo per i due valori di ε).

- Le figure 21 – 29 sono servite a illustrare la teoria sviluppata nel paragrafo precedente. Ma vanno oltre la teoria: infatti la scala di tempo su cui si osservano, per ε piccolo, i moti regolari, o anche i moti caotici dei primi tre rotatori ma disaccoppiati dalla parte restante della catena, è troppo grande perché la si possa considerare spiegata dall'analisi elementare di primo ordine sviluppata sopra. Un'analisi di ordine più elevato sarebbe necessaria. Non possiamo qui andare oltre; si vedano però più avanti, per dare l'idea della potenzialità dei metodi perturbativi, il teorema KAM e il teorema di Nekhoroshev.

F. Oltre il modello dei rotatori. Il modello dei rotatori che abbiamo sopra studiato ha molte peculiarità che ci hanno semplificato lo studio analitico e — grazie soprattutto alla localizzazione spaziale delle variabili d'azione e all'interazione anch'essa localizzata nella catena — hanno facilitato l'esposizione dei risultati e la loro visualizzazione fisica. In realtà niente di tutto questo è realmente necessario, tutto si generalizza in maniera abbastanza naturale (anche il passaggio da polinomio a serie di Fourier è abbastanza indolore), ad eccezione di un punto: *il meccanismo di confinamento, per il quale abbiamo utilizzato la conservazione dell'energia*. Un'estensione facile è quella a generiche hamiltoniane imperturbate convesse, non necessariamente quadratiche né in alcun senso localizzate,

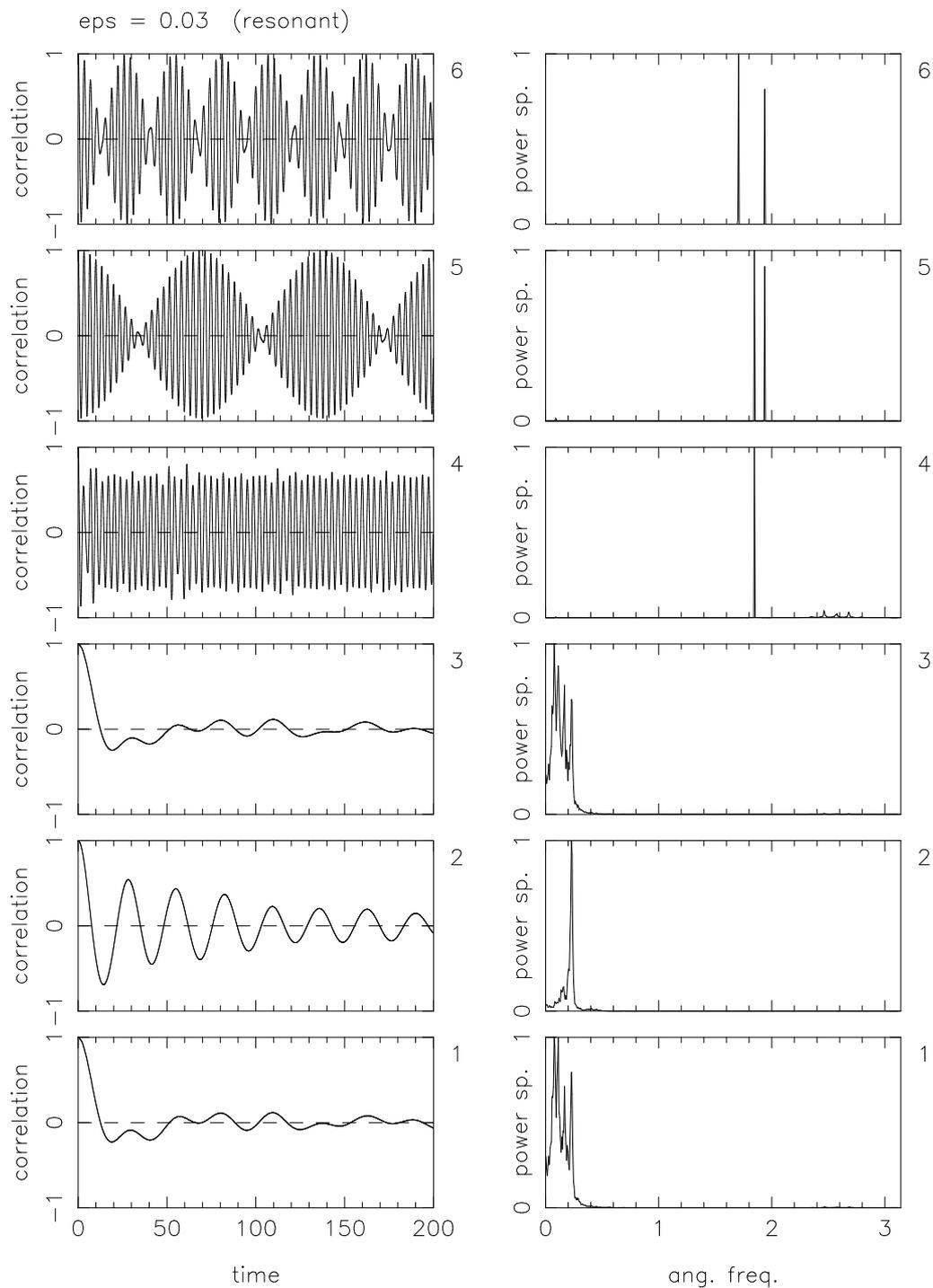


Figura 28: Come nella figura 26, per lo stesso $\varepsilon = 0.03$, ma dentro alla risonanza $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\varepsilon}$; dati iniziali tali che $I_1^0 + I_2^0 + I_3^0 = 0$.

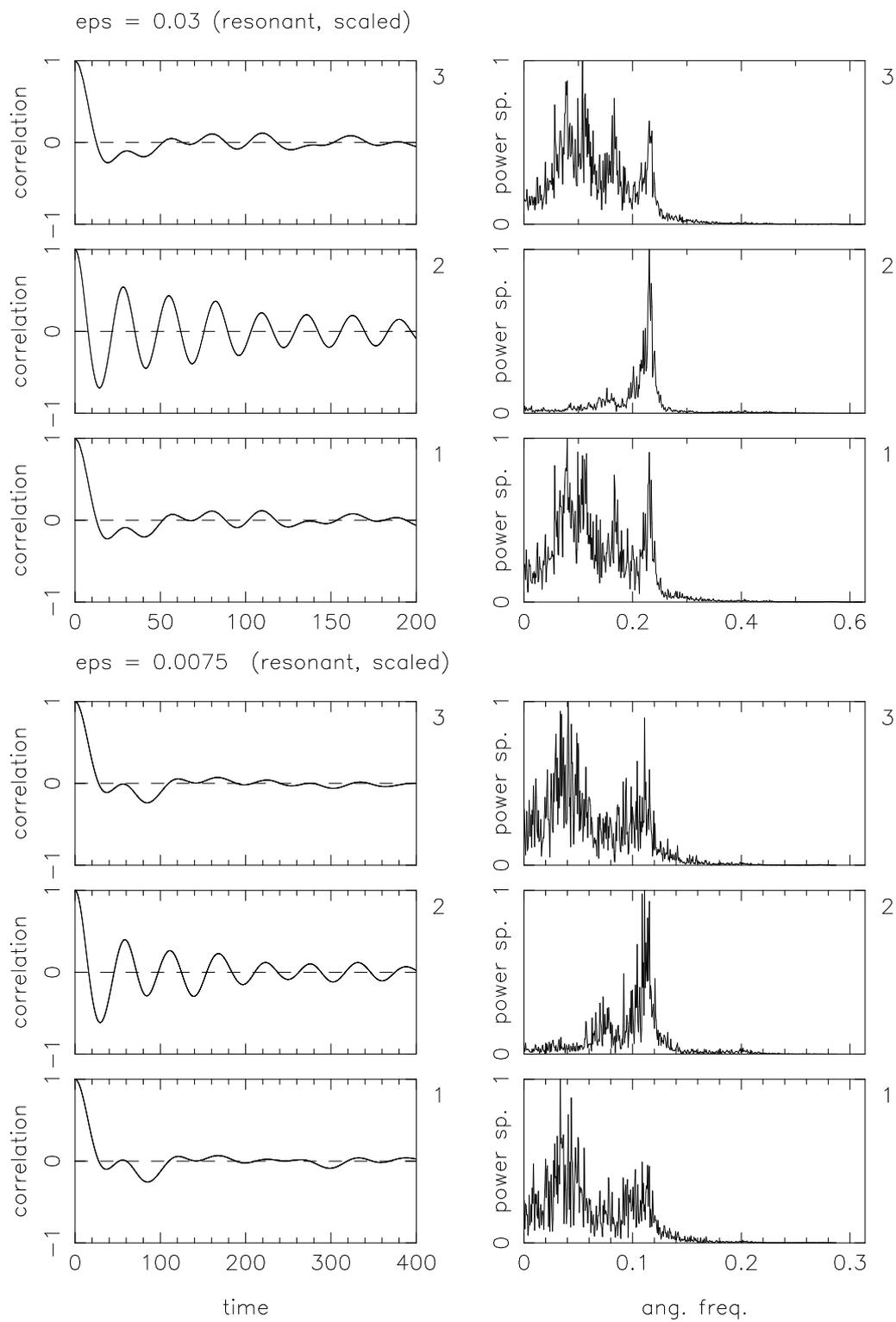


Figura 29: Correlazione e spettro di potenza di I_1, I_2, I_3 , nella risonanza $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\varepsilon}$, a $\varepsilon = 0.03$ (alto) e $\varepsilon = 0.0075$ (basso).

per le quali la conservazione dell'energia fornisce ancora il confinamento. Ma la comprensione del problema del confinamento, e delle ipotesi minime o “naturali” su h che lo garantiscono, è di quelli difficili. Qualche ipotesi su h più forte della (4.32) è sicuramente necessaria, se si vuole che le azioni restino confinate (e valga il principio della media), come mostra il controesempio (4.10). Capire a fondo il problema non è cosa da poco; l'idea fondamentale è stata introdotta nel 1976 da Nekhoroshev, con la nozione di *steepness*, ma il problema — molto rilevante per la fisica: i sistemi importanti tipicamente non sono convessi — è ancora allo studio. Ci fermiamo qui; uno studio poco più approfondito si trova nell'appendice J.

4.6 I grandi teoremi della teoria hamiltoniana delle perturbazioni

Esponiamo qui, senza nemmeno accennare alla dimostrazione, alcuni risultati importanti di carattere generale che si sono ottenuti per sistemi hamiltoniani prossimi a sistemi integrabili. Caratteristica comune a tali risultati è la capacità di dominare gli effetti della perturbazione, e garantire la stabilità delle azioni, o per tempi infiniti o per tempi che per ε piccolo crescono più rapidamente di qualunque potenza inversa di ε :

$$t \sim e^{(\varepsilon_*/\varepsilon)^a} \quad \text{con } \varepsilon_*, a > 0. \quad (4.42)$$

A. Il teorema “KAM”.

Il teorema KAM (Kolmogorov 1954, Arnol'd 1962, Moser 1961) è il più classico dei risultati moderni. Non è esagerato dire che sia il risultato per sé, sia i metodi introdotti nella dimostrazione, hanno segnato una svolta profonda nella teoria hamiltoniana delle perturbazioni (e non solo). Più che un singolo teorema è una collezione di teoremi e metodi di lavoro, che formano una vera e propria teoria (ancora in accrescimento, in particolare nelle sue estensioni a sistemi infinito dimensionali). Qui ci limiteremo a un singolo enunciato di base.

Il teorema riguarda sistemi hamiltoniani quasi integrabili

$$H_\varepsilon(I, \varphi) = h(I) + \varepsilon f(I, \varphi), \quad (I, \varphi) \in B \times \mathbb{T}^n, \quad (4.43)$$

con h non degenere, precisamente tale che sia

$$\det h''(I) \neq 0 \quad \text{ove } h'' = \left(\frac{\partial^2 h}{\partial I_i \partial I_j} \right). \quad (4.44)$$

E' la situazione in cui la mappa $I \mapsto \omega(I)$ è localmente invertibile, e corrispondentemente i punti non risonanti, come quelli con un numero arbitrario di risonanze $r \leq n - 1$, sono densi in B .

In buona sostanza, il teorema assicura che la struttura di sistema integrabile, precisamente la decomposizione dello spazio delle fasi in tori \mathbb{T}^n invarianti, sui quali i moti sono quasi periodici, persiste nel sistema perturbato riuscendone soltanto deformata, purché si escluda dallo spazio delle fasi un insieme di dati iniziali di misura di Lebesgue piccola, $\mathcal{O}(\sqrt{\varepsilon})$.

Diamo qui un enunciato, fra i diversi possibili, che più si adatta al nostro contesto (Lazutkin 1975; Chierchia e Gallavotti 1981; Pöschel 1981).

PROPOSIZIONE 22 Consideriamo l'hamiltoniana H_ε della forma (4.43). Supponiamo che:

- i) H_ε sia analitica e limitata in un intorno complesso D_{ext} del dominio $D = B \times \mathbb{T}^n$;

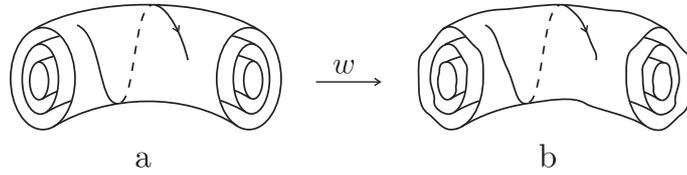


Figura 30: (a) I tori invarianti nelle nuove variabili; (b) gli stessi tori nelle vecchie variabili, “deformati” dalla trasformazione w prossima all’identità.

- ii) h sia non degenera, precisamente soddisfi la (4.44) in B ;
- iii) ε sia piccolo.

Allora esistono:

- (a) una trasformazione canonica prossima all’identità $(I, \varphi) = w_\varepsilon(\tilde{I}, \tilde{\varphi})$, $w_\varepsilon : D \rightarrow w_\varepsilon(D) \subset D_{\text{ext}}$, di classe C^∞ , definita nel dominio D ;
- (b) un’hamiltoniana integrabile \tilde{h}_ε prossima a h in classe C^∞ , $\tilde{h}_\varepsilon : B \rightarrow \mathbb{R}$;
- (c) un sottoinsieme $B_\varepsilon \subset B$ di misura di Lebesgue grande:

$$\text{mes } B \setminus B_\varepsilon \sim \sqrt{\varepsilon} ,$$

tali che la nuova hamiltoniana $\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) = H_\varepsilon(w_\varepsilon(\tilde{I}, \tilde{\varphi})) : D \rightarrow \mathbb{R}$ coincide con \tilde{h}_ε per $\tilde{I} \in B_\varepsilon$:

$$\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) \stackrel{B_\varepsilon}{\equiv} \tilde{h}_\varepsilon(\tilde{I}) ,$$

ove il simbolo $\stackrel{B_\varepsilon}{\equiv}$ denota uguaglianza di funzioni e derivate limitatamente a $\tilde{I} \in B_\varepsilon$. La frequenza $\omega = \frac{\partial \tilde{h}_\varepsilon}{\partial \tilde{I}}$, per $\tilde{I} \in B_\varepsilon$, soddisfa le relazione diofantea (4.30) con $\gamma \sim \sqrt{\varepsilon}$.

Si rifletta bene su questa non comune situazione:

- la trasformazione canonica, e con essa $\tilde{H}_\varepsilon = H_\varepsilon \circ w$, sono definiti in tutto il dominio D ;
- solo su un certo sottoinsieme di misura grande si garantisce che \tilde{H} è indipendente dagli angoli, e corrispondentemente il sistema, ristretto a questo sottoinsieme, si comporta come se fosse integrabile.

L’immediata conseguenza della proposizione è che le nuove azioni \tilde{I} , se prese inizialmente nell’insieme grande B_ε , restano costanti; corrispondentemente i tori $\{\tilde{I}\} \times \mathbb{T}^n$ sono invarianti e su di essi il moto degli angoli è lineare e quasi periodico (con frequenze diofantee). Complessivamente questi moti formano un insieme di misura grande. L’immagine di ciascun toro nelle vecchie variabili è ancora un toro invariante, solo un po’ deformato dalla trasformazione w (si veda la figura 30). Le vecchie azioni I , dovendo restare prossime alle \tilde{I} , oscillano di poco. Il principio della media è soddisfatto per tempi infiniti, *condizionatamente a prendere il dato iniziale su uno dei tori invarianti*.

L’aspetto delicato del risultato (*non facile da capire*) è che l’insieme B_ε , così come è costruito nella dimostrazione, pur avendo misura grande ha interno vuoto, più precisamente il suo complementare $\overline{B_\varepsilon} = B \setminus B_\varepsilon$ è aperto denso. Inoltre, per $n \geq 3$, $\overline{B_\varepsilon}$ è anche connesso: come una rete o un intreccio (“Arnol’d web”) di “massa” piccola, ma sparso ovunque. Da un lato dunque per la grande

maggioranza *in misura* di dati iniziali il sistema si comporta come se fosse integrabile. Dall'altro *per nessun aperto* si può garantire che le azioni resteranno prossime al valore iniziale, e la diffusione lungo la rete (“Arnol’d diffusion”) non si può escludere. Si conoscono esempi particolarissimi di sistemi in cui la diffusione si dimostra esistere. Fino a che punto il fenomeno sia generico, e come avvenga — in particolare per sistemi fisicamente importanti come il Sistema Solare — è una cosa ancora non ben compresa.

B. Il teorema di Nekhoroshev.

Il prossimo risultato che esponiamo, dovuto a Nekhoroshev (1976, un allievo di Arnol’d) è in un certo senso complementare al teorema KAM. Riguarda anch’esso hamiltoniane della forma (4.43), ma a h si richiede che soddisfi un’ipotesi geometrica la cui espressione più semplice, anche se non la più generale, è la convessità:⁴⁵

$$u \cdot h''(I)u \neq 0 \quad \forall u \in \mathbb{R}^n \setminus \{0\}, \quad I \in B \quad (4.45)$$

($u \cdot h''u$ deve così essere sempre positivo o sempre negativo). Il risultato è, come per il teorema KAM, un risultato di stabilità delle azioni. Ma a differenza del teorema KAM:

- il risultato vale per ogni dato iniziale;
- la scala di tempo su cui si assicura la stabilità delle azioni non è infinita, ma “solo” molto lunga come nella (4.42).

L’uniformità del risultato nello spazio delle fasi lo rende molto adatto alle applicazioni fisiche; la limitazione a tempi lunghi ma non infiniti non è molto penalizzante fisicamente, dal momento che ogni modello fisico è buono in realtà solo entro una scala di tempo lunga ma finita.⁴⁶ Le applicazioni fisiche (alla Meccanica Celeste, alla Meccanica Statistica, alla fisica del plasma, alla fisica degli acceleratori di particelle) sono state coltivate soprattutto in Italia.

PROPOSIZIONE 23 *Consideriamo H_ε della forma (4.43). Supponiamo che:*

- i) H_ε sia analitico e limitato in un intorno complesso del dominio $D = B \times \mathbb{T}^n$;*
- ii) h soddisfi l’ipotesi di convessità (4.45);*
- iii) ε sia piccolo.*

Allora esistono costanti positive ε_ , T_* , I_* , a , b tali che per ogni dato iniziale in D e per*

$$|t| \leq T_* e^{(\varepsilon_*/\varepsilon)^a}$$

risulta

$$|I_i(t) - I_i(0)| \leq I_* \varepsilon^b .$$

Valori possibili di a , b sono $a = b = 1/(2n)$.

⁴⁵Di solito si prende come ipotesi la cosiddetta quasi-convessità: precisamente che risulti $u \cdot h''u(I) \neq 0$ per tutti gli $u \in \mathbb{R}^n$ tali che $u \cdot \omega(I) = 0$. Un esempio di funzione quasi-convessa ma non convessa è $h(I_1, I_2) = I_1 + I_2^2$. La proprietà più generale sotto la quale il teorema vale, detta *steepness*, è troppo complicata per essere qui descritta.

⁴⁶Ad esempio, modelli hamiltoniani del Sistema Solare hanno senso solo per tempi (qualcosa come 10^9 anni) in cui gli effetti dissipativi, dovuti principalmente alle maree, non si fanno sentire.

- Vale la pena di osservare che il controesempio (iv) al principio della media, riportato nel paragrafo 4.3, non soddisfa l'ipotesi di convessità (4.45), mentre la soddisfa l'esempio (iii).

C. *Stime esponenziali per sistemi isocroni perturbati.*

L'ultimo risultato che raccontiamo riguarda sistemi isocroni perturbati della forma già considerata nel paragrafo precedente, con ω che soddisfa la condizione diofantea (4.30).

PROPOSIZIONE 24 *Sia data l'hamiltoniana*

$$H_\varepsilon(I, \varphi, \varepsilon) = \omega \cdot I + \varepsilon f(I, \varphi, \varepsilon), \quad I \in B \subset \mathbb{R}^n, \quad \varphi \in \mathbb{T}^n.$$

Supponiamo che:

- f sia analitica e limitata in un intorno complesso di $B \times \mathbb{T}^n$;
- l' n -pla $(\omega_1, \dots, \omega_n)$ delle frequenze soddisfi la relazione diofantea

$$|k \cdot \omega| \geq \frac{\gamma}{|k|^n} \quad \forall k \in \mathbb{Z}^n \setminus \{0\}$$

per un opportuno $\gamma > 0$;

- ε sia sufficientemente piccolo.

Allora esistono costanti positive T_* , I_* , ε_* tali che per ogni dato iniziale in $B \times \mathbb{T}^n$, e per la scala di tempo

$$|t| \leq T_* e^{(\varepsilon_*/\varepsilon)^{1/n}},$$

risulta

$$|I_j(t) - I_j(0)| \leq I_* \varepsilon^{1/n}.$$

Il risultato si colloca nel solco della teoria di Nekhoroshev ma è stato enucleato per la prima volta (in forma poco più debole) da Gallavotti (1984). Quello che si dimostra, in verità, è che esiste una trasformazione canonica $(I, \varphi) = w(\tilde{I}, \tilde{\varphi})$ piccola con ε tale che la nuova hamiltoniana ha la forma

$$\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) = \omega \cdot \tilde{I} + \varepsilon g(\tilde{I}, \varepsilon) + \varepsilon e^{(\varepsilon_*/\varepsilon)^{1/n}} f'(\tilde{I}, \tilde{\varphi}),$$

con g e f' limitate; la stima sulla stabilità delle azioni è allora un corollario.

Una variante del risultato precedente riguarda sistemi hamiltoniani della forma

$$H(p, q) = \sum_{i=1}^n E_i(p_i, q_i) + V_3(p, q) + V_4(p, q) + \dots, \quad (4.46)$$

ove $E_i = \frac{1}{2}\omega_i(p_i^2 + q_i^2)$ mentre V_s , $s \geq 3$, denota un polinomio omogeneo di grado s in p, q ; il dominio D è un piccolo intorno dell'origine:

$$D_\varepsilon = \{(p, q) \in \mathbb{R}^{2n} : E_i(p_i, q_i) < \varepsilon\}.$$

Questa classe di hamiltoniane, come si è già commentato nel paragrafo 4.1, comprende innanzitutto i sistemi hamiltoniani nell'intorno di una posizione di equilibrio stabile (p e q sono le variabili associate ai modi normali di oscillazione), ma comprende anche, si riveda l'esercizio 22, l'hamiltoniana

dell'asteroide attorno ai punti lagrangiani L_4 e L_5 : questo caso differisce dal precedente perché le frequenze ω_i non sono concordi e la stabilità del punto di equilibrio non è nota a priori. Nel trattare l'hamiltoniana (4.46) non ci si può ridurre alla proposizione precedente perché, come già si è osservato, passando alle variabili di azione–angolo intervengono le radici quadrate delle azioni e si perde la regolarità della perturbazione. La difficoltà tuttavia è solo tecnica, e si può dimostrare la seguente proposizione:

PROPOSIZIONE 25 *Sia H come sopra. Se l' n -pla delle frequenze soddisfa la condizione diofantea (4.30) e il parametro ε che caratterizza il dominio D_ε è sufficientemente piccolo, allora esistono costanti T_* , E_* , ε_* , tali che per ogni dato iniziale in D_ε e per la scala di tempo*

$$|t| < T_* e^{(\varepsilon_*/\varepsilon)^{1/n}}$$

risulta

$$|E_i(t) - E_i(0)|/\varepsilon < E_* \varepsilon^{1/n}, \quad i = 1, \dots, n,$$

ove si è usata la notazione agile $E_i(t)$ per $E_i(p_i(t), q_i(t))$.

Nel caso delle piccole oscillazioni attorno a una posizione di equilibrio stabile, la proposizione ci dice che se le frequenze dei modi normali di oscillazione sono diofantee (cosa che avviene con grande probabilità, se le frequenze sono prese a caso con riferimento alla misura di Lebesgue), allora i modi normali sono sostanzialmente disaccoppiati per tempi lunghi (per il sistema vero, non solo per quello linearizzato). Per il problema di L_4 , L_5 ci viene garantita, in più, la stabilità non perpetua ma per tempi lunghi.

- I sistemi quasi integrabili fisicamente interessanti, ad esempio quelli che si presentano naturalmente in meccanica celeste, di rado soddisfano le ipotesi del teorema KAM o del teorema di Nekhoroshev, che difficilmente sono direttamente applicabili. Si tratta tuttavia di teoremi (o meglio teorie) abbastanza flessibili, che si applicano magari magari dopo un lavoro preliminare *ad hoc* sull'hamiltoniana, oppure sfruttando qualche peculiarità del sistema che compensa il non pieno soddisfacimento di un'ipotesi. Questo è vero in particolare per il teorema di Nekhoroshev, che ben si adatta a operare, tra gli altri, sui sistemi degeneri come il corpo rigido, nei quali non tutte le azioni sono presenti nell'hamiltoniana imperturbata.

4.7 Invarianti adiabatici

Consideriamo un sistema hamiltoniano $H(p, q, \lambda)$ dipendente da un parametro λ . Ci chiediamo che cosa avviene se λ varia lentamente nel tempo: diciamo

$$\lambda = \lambda(\varepsilon t),$$

con ε piccolo e $\lambda(\cdot)$ funzione assegnata. Il problema comincia a essere interessante per $t = \mathcal{O}(1/\varepsilon)$ o maggiore, cioè quando si lascia al parametro λ il tempo di variare di una quantità che non è piccola per ε piccolo. La domanda spontanea è se, pur variando il parametro in modo significativo, non ci sia qualche cosa che si conserva; l'energia, si ricordi, non si conserva, per via della dipendenza dell'hamiltoniana da t . Una definizione appropriata allora è la seguente:

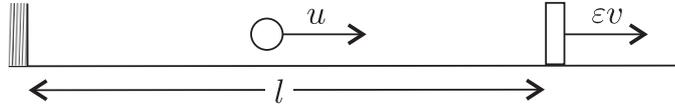


Figura 31: Un modello elementare di gas in una dimensione.

DEFINIZIONE 12 Una funzione $F(p, q, \lambda)$ si dice essere un invariante adiabatico per il sistema di hamiltoniana $H(p, q, \lambda(\varepsilon t))$, se esiste $\tau_0 > 0$ tale che per ogni moto del sistema risulti

$$F(p(t), q(t), \lambda(\varepsilon t)) - F(p(0), q(0), \lambda(0)) \rightarrow 0 \quad \text{per } \varepsilon \rightarrow 0$$

in tutto l'intervallo $0 \leq t \leq \tau_0/\varepsilon$.

Il tempo $\tau = \varepsilon t$ è detto “tempo breve” del sistema; lo studio è dunque per $\varepsilon \rightarrow 0$ in un intervallo di tempo breve fissato.

Ci ripromettiamo di studiare la presenza di un invariante adiabatico in casi semplici, applicando un metodo perturbativo. Prima però, per illustrare la nozione, studiamo un modellino del tutto elementare ove la presenza di un invariante adiabatico si mette in evidenza con un calcolo esplicito.

A. Un modello di gas in una dimensione.

Consideriamo un modello brutalmente semplificato di gas in una dimensione, composto da un solo punto materiale che rimbalza avanti e indietro tra una parete fissa e un pistone che si muove lentamente (figura 31). Detta l la distanza del pistone dalla parete fissa, prendiamo per semplicità la legge lineare

$$l = l_0 + \varepsilon v t ,$$

ma quanto andiamo a dire si estende a $l(\varepsilon t)$ regolare qualsiasi.

Denotiamo con u il modulo della velocità del punto materiale; si verifica facilmente che il prodotto lu è un invariante adiabatico del sistema, e un possibile valore di τ_0 è $\tau_0 = l_0/(2v)$. Per vederlo, denotiamo con l_j la posizione del pistone al momento della j -esima collisione e con u_j il modulo della velocità del punto materiale tra la j -esima e la $j+1$ -esima collisione. La collisione successiva avviene dopo il tempo t tale che $u_j t = 2l_j + \varepsilon v t$, ovvero $t_j = 2l_j/(u_j - \varepsilon v)$, e in quel momento il pistone è nella posizione

$$l_{j+1} = l_j + \frac{2\varepsilon v l_j}{u_j - \varepsilon v} ;$$

il nuovo valore della velocità all'uscita della collisione è poi

$$u_{j+1} = u_j - 2\varepsilon v .$$

Si ha così una mappa $(l_{j+1}, u_{j+1}) = \Phi(l_j, u_j)$ da una collisione all'altra, ed è immediato vedere che tale mappa conserva il prodotto lu al primo ordine in ε : infatti

$$\begin{aligned} l_{j+1} u_{j+1} &= l_j u_j - 2\varepsilon v l_j + 2\varepsilon v l_j \frac{u_j}{u_j - \varepsilon v} + 4\varepsilon^2 v l_j \frac{v}{u_j - \varepsilon v} \\ &= l_j u_j + \mathcal{O}(\varepsilon^2) . \end{aligned}$$

E' necessario che $u_j = u_0 - 2j\varepsilon v$ resti staccato da zero, e per questo imponiamo, ad esempio, $u_j \geq u_0/2$. Ciò limita il numero di urti a $n_0 = u_0/(4\varepsilon v)$. Il tempo trascorso è dato dalla somma

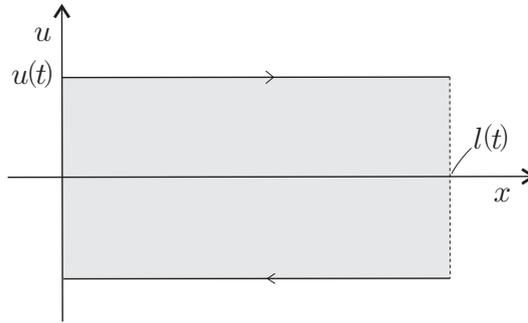


Figura 32: L 'invariante adiabatico è l'area nel piano uv , ovvero l'azione del sistema integrabile a parametro fissato.

dei t_j fino a n_0 , e poichè per ogni j si ha $t_j > 2l_0/u_0$, si trova un intervallo di tempo superiore a τ_0/ε con

$$\tau_0 = \frac{l_0}{2v} ; \quad (4.47)$$

nell'intervallo τ_0/ε dunque è avvenuto un numero di urti non superiore a $n_0 = \mathcal{O}(1/\varepsilon)$, e corrispondentemente il prodotto lu è variato di $\mathcal{O}(\varepsilon)$. Il prodotto lu è pertanto un invariante adiabatico, con τ_0 dato dalla (4.47).

- Quella che abbiamo trovato è la legge delle adiabatiche in una dimensione, ove si interpreti l come volume e u^2 (a meno di una costante moltiplicativa) come temperatura. Infatti per un gas monoatomico in dimensione $d = 1$ la costante γ delle adiabatiche vale $(d + 2)/d = 3$, perciò la legge delle adiabatiche $TV^{\gamma-1} = \text{cost}$ si scrive $u^2 l^2 = \text{cost}$ che è proprio quello che abbiamo trovato.
- Vale la pena di osservare che il prodotto ul , ovvero l'invariante adiabatico in questo problema, è a meno di un fattore due l'area nel piano di fase ux sottesa alle traiettorie (discontinue) del sistema a l costante (figura 32), ovvero è l'azione del sistema integrabile a parametro fissato.

ESERCIZIO 45 *Si stimi al meglio la variazione $\mathcal{O}(\varepsilon)$ del prodotto lu al tempo τ_0/ε .*

ESERCIZIO 46 *Si generalizzi la dimostrazione al caso in cui il pistone si muove di moto vario, cioè per $l(\tau) \geq L > 0$ regolare qualsiasi.*

B. *L'oscillatore armonico con frequenza variabile.*

La teoria degli invarianti adiabatici è semplice e ben stabilita per sistemi a un grado di libertà che a parametro bloccato eseguono moti periodici. Qui ci limitiamo a descrivere un esempio facile per il quale si scrive tutto esplicitamente, precisamente l'oscillatore armonico con frequenza variabile

$$H(p, q, \varepsilon t) = \frac{1}{2}(p^2 + \omega(\varepsilon t)^2 q^2) , \quad (4.48)$$

ma il risultato si generalizza opportunamente a sistemi oscillanti qualsiasi, ad esempio al pendolo (pur di restare lontani dalle separatrici, in modo che la frequenza resti sempre staccata da zero). Precisamente, dimostriamo che

PROPOSIZIONE 26 Per il sistema (4.48), se $\omega(\tau) \neq 0$ per $0 \leq \tau \leq \tau_0$, l'azione

$$I(p, q, \omega) = \frac{1}{2\omega}(p^2 + \omega^2 q^2) \quad (4.49)$$

è un invariante adiabatico.

DIMOSTRAZIONE. Possiamo scomporre concettualmente la dimostrazione in tre passi.

Primo passo, elementare: passiamo al sistema autonomo a due gradi di libertà nello spazio esteso,

$$H_\varepsilon^{\text{ext}}(p, q, A, \tau) = \frac{1}{2}(p^2 + \omega(\tau)^2 q^2) + \varepsilon A ,$$

ove A, τ sono variabili coniugate; poiché $\dot{\tau} = \varepsilon$, si verifica immediatamente che i moti di $H_\varepsilon^{\text{ext}}$ con $\tau(0) = 0$ coincidono, per quanto riguarda p e q , con i moti di H . Nel seguito eseguiremo alcune trasformazioni canoniche, ma tutte tali da non cambiare τ che mantiene così il significato di tempo breve εt .

Secondo passo, un po' laborioso: introduciamo le variabili di azione–angolo dell'oscillatore, estendendole in modo opportuno alla coppia (J, τ) . Precisamente, cerchiamo una trasformazione

$$(p, q, A, \tau) = w(I, \varphi, J, \tau')$$

in cui (i) τ non cambia, cioè $\tau = \tau'$; (ii) I, φ sono quelle che scriveremmo se ω fosse fissato, ovvero

$$p = \sqrt{2I\omega(\tau)} \cos \varphi , \quad q = \sqrt{2I/\omega(\tau)} \sin \varphi , \quad (4.50)$$

cosicché l'azione I è data dalla (4.49); infine (iii) l'espressione di A in funzione delle nuove variabili è tale da rendere complessivamente canonica w . (In altre parole: per p e q scriviamo la trasformazione canonica “a tempo bloccato” che avremmo scritto per ω costante, poi la estendiamo canonicamente alla coppia (A, τ) curando però che sia $\tau = \tau'$.) È facile verificare che una buona scelta di A esiste ed è

$$A = J + a(I, \varphi, \tau) , \quad (4.51)$$

con

$$a(I, \varphi, \tau) = \frac{\omega'(\tau)}{2\omega(\tau)} I \sin 2\varphi , \quad (4.52)$$

ove ω' denota la derivata di ω rispetto a τ . È facile verificarlo, guardando alle parentesi di Poisson; un po' più difficile è inventarsi le (4.51), (4.52). Un modo spontaneo di procedere⁴⁷ è quello di scrivere la funzione generatrice in variabili miste $\mathcal{S}(p, \varphi, \tau)$ che per ogni τ fissato dà la (4.50), o meglio la sua inversa; già sappiamo (esercizio 10) che la scelta buona è

$$\mathcal{S}(p, \varphi, \tau) = \frac{p^2}{2\omega(\tau)} \tan \varphi .$$

⁴⁷Chi non ama le funzioni generatrici può far leva sulle parentesi di Poisson: perché risulti $\{\tau, A\} = \frac{\partial A}{\partial J} = 1$, l'espressione di A deve essere necessariamente del tipo (4.51); la forma (4.52) di a si deduce poi facilmente imponendo che $\{A, p\}$ e $\{A, q\}$ (le uniche parentesi di Poisson non banali) si annullino. Così facendo vengono equazioni lineari per $\frac{\partial a}{\partial I}$ e $\frac{\partial a}{\partial \varphi}$ che si risolvono subito e danno

$$\frac{\partial a}{\partial I} = \frac{\omega'(\tau)}{2\omega(\tau)} \sin 2\varphi , \quad \frac{\partial a}{\partial \varphi} = \frac{\omega'(\tau)}{\omega(\tau)} I \cos 2\varphi ;$$

segue la (4.52).

Questa generatrice si estende all'altro grado di libertà "aggiungendo l'identità", cioè ponendo

$$S(p, \varphi, A, \tau') = \mathcal{S}(p, \varphi, \tau') + A\tau' ;$$

la scelta assicura che τ non cambi, infatti si ha $\tau = \frac{\partial S}{\partial A} = \tau'$. A questo punto si scrive $J = \frac{\partial S}{\partial \tau'}$, e il risultato, togliendo l'apice a τ' , è l'accoppiata (4.51), (4.52).

La nuova hamiltoniana $\mathcal{H}_\varepsilon = H_\varepsilon^{\text{ext}} \circ w$ risulta essere

$$\mathcal{H}_\varepsilon(I, J, \varphi, \tau) = h(I, \tau) + \varepsilon J + \varepsilon a(I, \varphi, \tau) ,$$

con

$$h(I, \tau) = \omega(\tau) I .$$

E' interessante osservare che la lenta dipendenza di ω dal tempo si è tradotta, al momento di passare alle variabili di azione–angolo, nella comparsa del termine perturbativo $\varepsilon a(I, \varphi, \tau)$. E' cruciale, si vede, aver supposto $\omega(\tau) \neq 0$.

Terzo passo: Su questa hamiltoniana lavoriamo ora con metodi perturbativi, mostrando che

$$|I(t) - I^0| < \text{cost } \varepsilon \quad \text{per } |t| \leq \frac{\tau_0}{\varepsilon} .$$

La presenza della variabile τ , che non è un'azione, nell'hamiltoniana imperturbata, apparentemente rende il problema diverso da quelli visti precedentemente; vedremo però che lo stesso schema perturbativo funziona ugualmente senza difficoltà.

Eseguiamo allora un passo perturbativo con il metodo di Lie, ovvero poniamo

$$(I, \varphi, J, \tau) = \Phi_\chi^\varepsilon(\tilde{I}, \tilde{\varphi}, \tilde{J}, \tilde{\tau}) ,$$

avendo però l'accortezza di prendere χ indipendente da J , in modo che nella dinamica di χ (non in quella vera!) sia $\dot{\tau} = 0$, e così nella trasformazione canonica risulti $\tau = \tilde{\tau}$. Come risultato troviamo la nuova hamiltoniana $\tilde{\mathcal{H}}_\varepsilon = \mathcal{H}_\varepsilon \circ \Phi_\chi^\varepsilon$ della forma

$$\tilde{\mathcal{H}}_\varepsilon = h + \varepsilon \{h, \chi\} + \varepsilon(J + a) + \mathcal{O}(\varepsilon^2) \quad \text{con} \quad h = \omega(\tau) I ,$$

e si deve determinare χ in modo che il termine di ordine ε perda la dipendenza dall'angolo φ . Grazie al fatto che χ non dipende dalla variabile J coniugata a τ , la parentesi di Poisson $\{h, \chi\}$ è la stessa che si avrebbe per ω costante, e per questo la presenza di τ in h non disturba.

Osservando che la media della perturbazione è J (a infatti ha media in φ nulla), l'equazione da risolvere risulta essere

$$\omega(\tau) \frac{\partial \chi}{\partial \varphi} = a(I, \varphi, \tau) ,$$

ovvero

$$\frac{\partial \chi}{\partial \varphi} = \frac{\omega'(\tau)}{2\omega(\tau)^2} I \sin 2\varphi ,$$

ed evidentemente la soluzione c'è sempre,

$$\chi(I, \varphi, \tau) = -\frac{\omega'(\tau)}{4\omega(\tau)^2} I \cos 2\varphi ;$$

si osservi che grazie alla presenza di un solo grado di libertà veloce non ci sono piccoli divisori, ma solo ω a denominatore, che fino al tempo breve τ_0 è garantita essere diversa da zero.

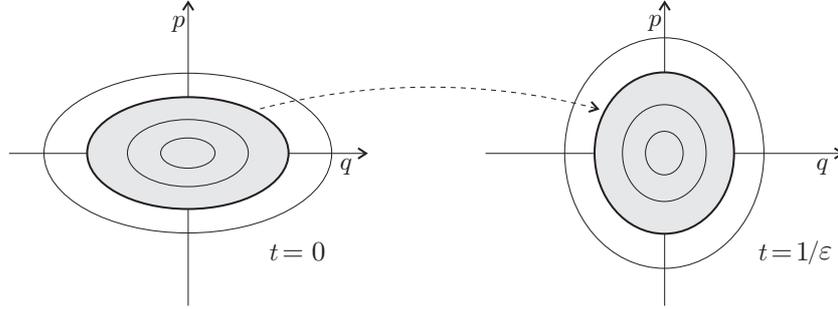


Figura 33: Ad illustrazione dell'invarianza adiabatica dell'azione. La corrispondenza tra moti a $t = 0$ e $t \sim 1/\varepsilon$ è a parità di azione (di area), non di energia.

La trasformazione canonica Φ_χ^ε porta così l'hamiltoniana $\tilde{\mathcal{H}}_\varepsilon$ nella forma

$$\tilde{\mathcal{H}}_\varepsilon(\tilde{I}, \tilde{\varphi}, \tilde{J}, \tau) = \omega(\tau) \tilde{I} + \varepsilon \tilde{J} + \mathcal{O}(\varepsilon^2),$$

e corrispondentemente si hanno le solite stime del tipo

$$|\tilde{I}(t) - \tilde{I}(0)| < \text{cost } \varepsilon \quad \text{per } 0 \leq t \leq \tau_0/\varepsilon.$$

D'altra parte, la trasformazione canonica Φ_χ^ε è ε -vicina all'identità, e in particolare si ha

$$|I(t) - \tilde{I}(t)|, |I(0) - \tilde{I}(0)| < \text{cost } \varepsilon;$$

segue allora

$$|I(t) - I(0)| < \text{cost } \varepsilon \quad \text{per } 0 \leq t \leq \tau_0/\varepsilon,$$

ovvero l'invarianza adiabatica dell'azione. □

Si osservi che per l'oscillatore armonico che stiamo considerando l'azione è legata all'energia E da $I = E/\omega(\tau)$. L'invarianza adiabatica di I ci dice così immediatamente che per lente variazioni di ω l'energia E varia proporzionalmente a ω . La figura 33 illustra il risultato: dopo il tempo t di ordine $1/\varepsilon$ la pulsazione $\omega(\varepsilon t)$ è cambiata, e corrispondentemente è cambiato il ritratto in fase del sistema (il ritratto, si intende, a parametro bloccato); moti che a $t = 0$ si svolgevano su una certa curva di livello, trascorso il tempo t si svolgono non sulla curva di livello con la medesima energia, ma su quella che ha (quasi esattamente) la medesima azione, ovvero sulla curva che sottende la medesima area.

Il risultato, come si è detto, si trasporta sostanzialmente senza variazioni a tutti i sistemi hamiltoniani $H(p, q, \lambda)$ a un grado di libertà che per valori fissi del parametro λ abbiano moti periodici (il ritratto in fase per ogni λ fissato ha curve chiuse):

PROPOSIZIONE 27 *Con riferimento a tali hamiltoniane, sia $I(p, q, \lambda)$ l'azione a λ fissato, e sia $\omega(I, \lambda)$ la corrispondente frequenza. Se $\omega(I, \lambda(\tau))$ resta staccata da zero per $0 \leq \tau \leq \tau_0$, allora I è un invariante adiabatico.*

- Il caso di un sistema con separatrici che si muovono al variare del parametro, come il pendolo a frequenza variabile

$$H(p, q, \omega(\varepsilon t)) = \frac{1}{2}p^2 - \omega(\varepsilon t)^2 \cos q,$$

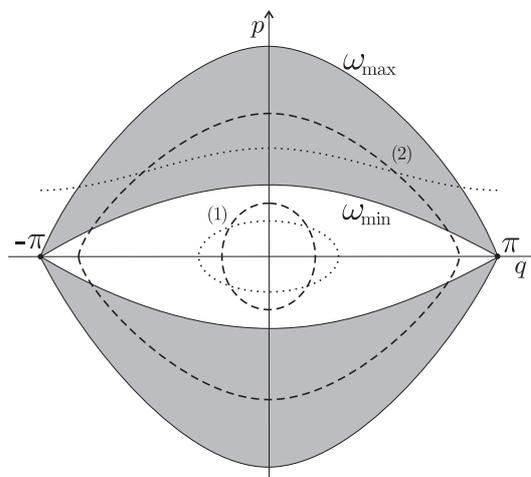


Figura 34: In grigio la zona spazzata dalle separatrici lentamente pulsanti. Per i moti sempre interni o esterni (al variare di ω) c'è un invariante adiabatico, con situazione non dissimile dall'oscillatore armonico; nella zona grigia l'invarianza adiabatica è interrotta dall'attraversamento delle separatrici.

si presenta in realtà più ricco e complesso dell'oscillatore armonico visto sopra. Prendiamo, per fissare le idee,

$$\omega(\tau) = (2 + \cos \Omega\tau) \omega_0 ;$$

al variare di τ le separatrici (del sistema a ω bloccata) spazzano un'area come quella indicata in grigio in figura 34. Al di fuori di quest'area, nella regione sempre interna o sempre esterna alle separatrici lentamente pulsanti, il metodo di prova utilizzato per il pendolo funziona perfettamente (salvo che molte espressioni, a cominciare da $\mathcal{S}(p, \varphi, \tau)$, restano implicite: sono ben definite, ma non si sbrigliano in funzioni elementari). L'invariante adiabatico, come nei casi visti sopra, è l'azione, ovvero l'area sottesa da una traiettoria del sistema a ω bloccata. E' il caso della coppia di traiettorie (1) in figura: quella tratteggiata corrisponde a $\omega_{\max} = 3\omega_0$, e si deforma nel mezzo periodo $\tau = \pi/\Omega$ nella traiettoria punteggiata corrispondente a $\omega_{\min} = \omega_0$; le due curve sottendono (quasi esattamente) la stessa area. Dopo un periodo completo $2\pi/\Omega$ le traiettorie riprendono (quasi esattamente) l'aspetto iniziale.

L'invarianza adiabatica viceversa si interrompe se un moto attraversa la separatrice, passando dalla regione (momentaneamente) interna a quella esterna o viceversa. E' il caso della coppia di curve (2): come sopra, la curva tratteggiata è una traiettoria corrispondente a ω_{\max} , interna alla separatrice; nel corso del tempo la separatrice attraversa la curva, e la traiettoria diventa esterna. Prima e dopo il passaggio l'invariante adiabatico si conserva (l'area sottesa resta quasi costante); nel passaggio invece l'invariante adiabatico subisce un salto.⁴⁸

In realtà (sempre per sistemi con un solo grado di libertà) si dimostra di più della proposizione 27, ad esempio:

⁴⁸Il salto si può calcolare; A.I. Neishtadt ha lavorato molto in questo campo, studiando anche la statistica possibile dei successivi salti in un tempo lungo e la possibilità che li si possano descrivere come processo casuale diffusivo (cosa che non risulta sempre vera: salti successivi, in certe condizioni, sono correlati e ci sono compensazioni.)

- se $H(p, q, \lambda(\varepsilon t))$ è analitica, l'azione è un invariante adiabatico per tempi molto lunghi,

$$|t| \leq \text{cost } e^{\varepsilon_*/\varepsilon} ,$$

ove ε_* è una costante positiva.

- Sempre nel quadro analitico, se $\lambda(\tau)$ tende a limiti λ_{\pm} per $\tau \rightarrow \pm\infty$, allora I è un invariante adiabatico per ogni t ,

$$|I(t) - I(0)| < \text{cost } \varepsilon \quad \forall t \in \mathbb{R} ,$$

inoltre la variazione asintotica di I tra $t = -\infty$ e $t = +\infty$ è molto piccola: precisamente, esistono i limiti $I_{\pm\infty} = \lim_{t \rightarrow \pm\infty} I(t)$ e si ha

$$|I_{+\infty} - I_{-\infty}| < \text{cost } e^{-\varepsilon_*/\varepsilon} .$$

- Infine, se H dipende da εt in modo periodico l'azione è invariante adiabatico per tutti i tempi.

Ulteriori estensioni, rilevanti per la fisica, comprendono il caso in cui al posto della dipendenza lenta da t vi sia nel sistema un secondo grado di libertà lento (si pensi ad esempio a un elettrone che avanza lentamente lungo le linee di un campo magnetico, ruotando attorno ad esse con la frequenza di Larmor).

La nozione di invariante adiabatico fu formulata da P. Ehrenfest. C'è una connessione con la quantizzazione: grazie al fatto che l'azione è la variabile quantizzata, la lenta variazione di λ avviene senza variazione dello stato quantico. (Più precisamente, per ε piccolo la probabilità di transizione dallo stato quantico di partenza ad un altro in un tempo $t = 1/\varepsilon$ è piccola.)

APPENDICI

A Introduzione al formalismo geometrico.

Nel primo paragrafo di questa appendice si richiama il formalismo geometrico che sta alla base della nozione di equazione differenziale su una varietà. “Si richiama” vuol dire che si suppone già una minima familiarità con le idee di base (la nozione stessa di varietà con le sue carte e l’atlante; la nozione di spazio tangente e di campo vettoriale; la nozione di differenziale...). Anche chi fosse del tutto a digiuno, tuttavia, un’idea se la può fare.

Il secondo paragrafo è invece orientato alla nozione di sistema hamiltoniano su una varietà. Il punto di vista geometrico non è veramente necessario a capire le due idee centrali su cui si sviluppano queste note — in buona sostanza, la nozione di sistema integrabile e di sistema prossimo a integrabile, con le diverse scale di tempo associate al moto — ma di certo costituisce, anche se solo intuito e non studiato a fondo, un punto di vista più maturo. Il linguaggio è un po’ informale; lo spirito è soprattutto quello di rivisitare da un punto di vista geometrico quanto già visto nel testo nel linguaggio delle coordinate, con poche estensioni (seppur significative, in particolare quando si introduce la 2-forma canonica cui nel testo non si fa cenno).

Una parola sulle notazioni. E’ qui particolarmente importante distinguere gli oggetti globali, geometrici, definiti sulla varietà, dai loro rappresentativi in una carta di coordinate locali. Le notazioni necessarie saranno introdotte di volta in volta, ma l’idea è di usare maiuscole calligrafiche come \mathcal{F} per le funzioni sulla varietà e la corrispondente comune minuscola f per la rappresentativa locale; per oggetti a n componenti — i punti stessi della varietà, vettori, 1-forme — si useranno notazioni del tipo x, X, ϑ per gli oggetti globali, e gli stessi simboli sottolineati $\underline{x}, \underline{X}, \underline{\vartheta}$ per le n -ple che li rappresentano localmente. Per le 2-forme, si userà similmente una notazione del tipo ω per l’oggetto geometrico, $\underline{\omega}$ per la matrice che lo rappresenta localmente.

A.1 Funzioni, campi vettoriali, forme

Coordinate locali.

Sia M una varietà differenziabile n -dimensionale. Denoteremo le sue carte con (U, φ) , ove $U \subset M$ e $\varphi : U \rightarrow V$ biunivoca, V aperto di \mathbb{R}^n . Ogni carta costituisce un sistema di coordinate locali. Se φ invia $x \in U$ in $\underline{x} = (x_1, \dots, x_n) \in V$, allora $\varphi_i : x \mapsto x_i \in \mathbb{R}$ — in notazione più spiccia $x_i(x)$ — è detta i -esima (funzione) coordinata, mentre le immagini inverse delle consuete linee coordinate di \mathbb{R}^n , ristrette a V , costituiscono le linee coordinate locali su M .

Date due carte (U, φ) e $(\tilde{U}, \tilde{\varphi})$, posto $U_0 = U \cap \tilde{U} \neq \emptyset$, $V_0 = \varphi(U_0)$, $\tilde{V}_0 = \tilde{\varphi}(U_0)$, la funzione di transizione

$$\psi = \varphi \circ \tilde{\varphi}^{-1}, \quad \psi : \tilde{V}_0 \rightarrow V_0$$

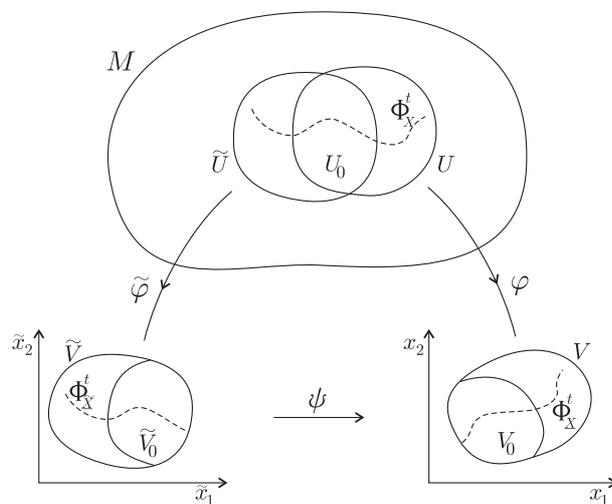


Figura 35: La varietà M , due sue carte con la funzione di transizione $\psi : \tilde{V}_0 \rightarrow V_0$; un moto (linea tratteggiata) sulla varietà e nelle carte.

rappresenta un cambiamento locale di coordinate (figura 35). Nel seguito scriveremo semplicemente $\psi : \tilde{V} \rightarrow V$, intendendo $\psi : \tilde{V}_0 \rightarrow V_0$. La classe di regolarità delle funzioni di transizione è per definizione la classe di regolarità di M . Analogamente per tutti gli oggetti geometrici (funzioni, campi vettoriali, 1 e 2-forme) che introdurremo, la classe di regolarità sarà, per definizione, quella dei loro rappresentativi locali. Per semplicità, e senza farne esplicita menzione, supporremo sempre tutto di classe C^∞ .

Funzioni a valori in \mathbb{R} .

Sia \mathcal{F} una funzione a valori reali (o 0-forma) su M , $\mathcal{F} : M \rightarrow \mathbb{R}$; per ogni carta (U, φ) resta definita la *rappresentativa locale*

$$f = \mathcal{F} \circ \varphi^{-1} : V \rightarrow \mathbb{R}, \quad V = \varphi(U).$$

Se f, \tilde{f} sono rappresentative locali di una stessa \mathcal{F} in due diverse carte, e $\psi : \tilde{V} \rightarrow V$ è come sopra la funzione di transizione da una carta all'altra, allora si ha

$$\tilde{f} = f \circ \psi, \quad \text{ovvero} \quad \tilde{f}(\tilde{x}) = f(\psi(\tilde{x})); \quad (\text{A.1})$$

viceversa, se si ha una famiglia di funzioni locali $f : V \rightarrow \mathbb{R}$, una per ogni carta dell'atlante, e sono soddisfatte le relazioni di compatibilità (A.1) per ogni coppia di carte (con intersezione non vuota tra i domini), resta definita una funzione intrinseca $\mathcal{F} : M \rightarrow \mathbb{R}$. La funzione \tilde{f} è anche detta *pull-back* di f sotto ψ (ψ va da \tilde{V} a V ; il *pull-back* porta funzioni f su V "indietro" a funzioni \tilde{f} su \tilde{V}).

Campi vettoriali; equazioni differenziali.

Come è noto, in ogni punto $x \in M$ resta definito uno *spazio tangente* $T_x M$, $\dim T_x M = n$. La collezione di tutti gli spazi tangenti ha una struttura naturale di varietà, ereditata dalla struttura di M , e costituisce il *fibrato tangente* TM , $\dim TM = 2n$; ogni punto di TM è una coppia (x, X) con $x \in M$ e $X \in T_x M$. Un campo vettoriale su M è una mappa che ad ogni punto $x \in M$ associa

un vettore $X(x) \in T_x M$, formalmente è una mappa $M \rightarrow TM$ che a ogni $x \in M$ associa una coppia $(x, X(x)) \in TM$. Con piccolo abuso denoteremo il campo stesso con X . Per ogni carta (U, φ) , e per ogni $x \in U$, resta definita una base in $T_x M$ formata dai vettori tangenti alle linee coordinate, usualmente denotati $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ (o anche $\partial_{x_1}, \dots, \partial_{x_n}$); il campo in x si scriverà allora

$$X(x) = \sum_{i=1}^n X_i(x) \frac{\partial}{\partial x_i} .$$

La notazione introduce una piacevole confusione tra il campo e l'operazione di derivazione associata, la cosiddetta derivata di Lie definita sotto.

La mappa tangente a $\varphi : U \rightarrow V$, denotata $D\varphi$, manda lo spazio tangente $T_x M$, $x \in U$, nel tangente a V in $\underline{x} = \varphi(x)$, ovvero in \mathbb{R}^n , e associa alla base $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ la base canonica di \mathbb{R}^n : $D\varphi \frac{\partial}{\partial x_i} = e_i$, $i = 1, \dots, n$. Per ogni campo vettoriale X su M , e per ogni carta (U, φ) , resta allora definito un campo vettoriale su $V = \varphi(U)$, che denotiamo \underline{X} :

$$\underline{X}(\underline{x}) = (X_1(\underline{x}), \dots, X_n(\underline{x}))$$

(con piccolo abuso del simbolo X_i , usato qui al posto di $X_i \circ \varphi^{-1}$); \underline{X} si dice campo rappresentativo locale di X . Se \underline{X} e $\tilde{\underline{X}}$ sono rappresentativi locali di uno stesso campo X su M in due carte diverse, con funzione di transizione ψ dall'una all'altra come sopra, allora si ha la legge di trasformazione

$$\underline{X}(\underline{x}) = J(\tilde{\underline{x}}) \tilde{\underline{X}}(\tilde{\underline{x}}) , \quad \tilde{\underline{x}} = \psi^{-1}(\underline{x}) , \quad J = \left(\frac{\partial \psi_i}{\partial \tilde{x}_j} \right) . \quad (\text{A.2})$$

Viceversa, se in tutte le carte dell'atlante sono definiti campi locali che (se i domini si intersecano) soddisfano a due a due questa relazione di compatibilità, allora resta definito un campo X sulla varietà. Il campo $\tilde{\underline{X}}$ definito dalla (A.2) si dice *pull-back*⁴⁹ di \underline{X} . La relazione che lega tra loro le due basi su $T_x M$ è

$$\frac{\partial}{\partial \tilde{x}_i} = \sum_{j=1}^n J_{ij}^T \frac{\partial}{\partial x_j}$$

(come se si trattasse di derivate parziali).

Dato un campo vettoriale X su M restano definite un'equazione differenziale sulla varietà e un'equazione differenziale locale in ogni carta, che si scrivono rispettivamente

$$\dot{x} = X(x) , \quad \dot{\underline{x}} = \underline{X}(\underline{x}) ;$$

le diverse soluzioni locali $\Phi_{\underline{X}}^t$ (riportate su M) “si incollano bene” nell'intersezione dei domini delle carte, e in questo modo resta definita una soluzione in grande su M , detta *flusso* del campo X e denotata Φ_X^t . Se M è compatta, o comunque se Φ_X^t è prolungabile per $t \in \mathbb{R}$ (cosa che supporremo sempre), allora $\{\Phi_X^t, t \in \mathbb{R}\}$ è un gruppo:⁵⁰

$$\Phi_X^{t+s} = \Phi_X^t \circ \Phi_X^s , \quad (\Phi_X^t)^{-1} = \Phi_X^{-t} , \quad \Phi_X^0 = \mathbb{I} .$$

⁴⁹La definizione di *pull-back* di un campo è ben data grazie al fatto che ψ è invertibile. Per funzioni e forme invece l'invertibilità, pur presente nel nostro contesto, non è necessaria perchè il *pull-back* sia per sé ben definito.

⁵⁰Quando, come nel primo capitolo di queste note, si ragiona non su una varietà ma in aperti, per dar senso alla nozione di flusso si deve supporre, con fastidiosa restrizione, che la soluzione non esca mai dal dominio. In una prospettiva geometrica invece non c'è alcun problema se la soluzione passa da una carta all'altra attraverso l'intersezione dei domini.

Derivata di Lie di funzioni; parentesi di Lie.

A ogni campo vettoriale X su M resta associato un operatore di derivazione L_X che agisce sulle funzioni $M \rightarrow \mathbb{R}$, detto *derivata di Lie*. La definizione si può dare in coordinate (paragrafo 1.1), introducendo in ogni singola carta l'operatore di derivazione

$$L_{\underline{X}} = \sum_{i=1}^n X_i \frac{\partial}{\partial x_i}, \quad \text{ovvero} \quad (L_{\underline{X}}f)(\underline{x}) = \sum_{i=1}^n X_i(\underline{x}) \frac{\partial f}{\partial x_i}(\underline{x})$$

(la sottile differenza è che ora $\frac{\partial}{\partial x_i}$ ha il consueto significato di derivata). Si riconosce immediatamente che si ha

$$(L_{\underline{X}}f)(\underline{x}) = \left. \frac{d}{dt} f(\Phi_{\underline{X}}^t(\underline{x})) \right|_{t=0},$$

e che le diverse definizioni locali sono compatibili, cosicché resta definito un operatore L_X agente sulle funzioni $M \rightarrow \mathbb{R}$; l'operatore ammette anche la definizione intrinseca

$$(L_X \mathcal{F})(x) = \left. \frac{d}{dt} \mathcal{F}(\Phi_X^t(x)) \right|_{t=0}.$$

Siano ora X e Y campi vettoriali qualsiasi su M e L_X, L_Y le corrispondenti derivate di Lie. Il prodotto $L_X L_Y$ non è una derivata di Lie corrispondente ad alcun campo vettoriale (appena ci si mette in coordinate applicandolo a una funzione f spuntano le derivate seconde di f), ma lo è il commutatore

$$[L_X, L_Y] = L_X L_Y - L_Y L_X,$$

precisamente

PROPOSIZIONE 28 *Risulta $[L_X, L_Y] = L_Z$, con Z rappresentato, in una carta qualsiasi, da*

$$Z_i = L_X Y_i - L_Y X_i. \tag{A.3}$$

DIMOSTRAZIONE. Calcolando in coordinate si vede subito che i termini di derivata seconda si semplificano e si ha

$$[L_{\underline{X}}, L_{\underline{Y}}] f = \sum_{i=1}^n \left(\sum_{j=1}^n \left(X_j \frac{\partial Y_i}{\partial x_j} - Y_j \frac{\partial X_i}{\partial x_j} \right) \right) \frac{\partial f}{\partial x_i} = \sum_{i=1}^n (L_X Y_i - L_Y X_i) \frac{\partial f}{\partial x_i},$$

da cui segue l'espressione (A.3) di Z . L'indipendenza dalle coordinate si trova immediatamente sostituendo $X_i = \sum_k J_{ik} \tilde{X}_k$, $Y_i = \sum_k J_{ik} \tilde{Y}_k$. □

DEFINIZIONE 13 *Il campo vettoriale Z su M definito in questo modo è detto parentesi di Lie di X e Y , e si denota*

$$Z = [X, Y].$$

Si ha dunque, per definizione di parentesi di Lie,

$$[L_X, L_Y] = L_{[X, Y]}.$$

L'operazione è con evidenza antisimmetrica e in particolare $[X, X] = 0$. Con un po' di pazienza si verifica l'identità di Jacobi:

$$[[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X] = 0, \tag{A.4}$$

figlia della corrispondente relazione tra le derivate di Lie L_X, L_Y, L_Z .

La derivata di Lie $[X, Y]$ è anche detta *commutatore* dei campi vettoriali X e Y ; quando è nulla si dice che i campi vettoriali commutano. Alla commutazione dei campi corrisponde quella dei rispettivi flussi:

PROPOSIZIONE 29 *Si ha*

$$\Phi_X^s \circ \Phi_Y^t = \Phi_Y^t \circ \Phi_X^s \quad \forall s, t \in \mathbb{R}$$

se e solo se $[X, Y] = 0$.

La dimostrazione è identica a quella riportata nell'appendice C per il caso hamiltoniano.

1-forme su M .

Sia ora T_x^*M lo spazio duale a T_xM , cioè lo spazio vettoriale delle mappe lineari $T_xM \rightarrow \mathbb{R}$, detto *spazio cotangente* a M in $x \in M$. Ancora risulta che la collezione di tutti gli spazi cotangenti ha la struttura naturale di varietà, e si dice formare il *fibrato cotangente* T^*M , $\dim T^*M = 2n$. I punti di T^*M sono coppie (x, ϑ) , con $x \in M$ e ϑ mappa lineare: $T_xM \rightarrow \mathbb{R}$. Una 1-forma su M è (similmente a un campo vettoriale) una mappa $M \rightarrow T^*M$ che a ogni $x \in M$ associa un punto $(x, \vartheta(x)) \in T^*M$; la 1-forma sarà denotata semplicemente con ϑ . Data una 1-forma ϑ e un campo vettoriale X su M , resta definita una funzione $\mathcal{F} : M \rightarrow \mathbb{R}$ nel modo ovvio $\mathcal{F}(x) = (\vartheta(x))(X(x))$. Ovvero: una 1-forma è una mappa che manda linearmente campi vettoriali su M in funzioni $M \rightarrow \mathbb{R}$, e questa potrebbe essere usata come definizione.

Una base in ciascuno spazio cotangente T_x^*M è fornita dai differenziali delle funzioni coordinate $d\varphi_1, \dots, d\varphi_n$, in notazione più spiccia dx_1, \dots, dx_n , perciò si potrà scrivere

$$\vartheta(x) = \sum_{i=1}^n \vartheta_i(x) dx_i, \quad \vartheta_i = \vartheta\left(\frac{\partial}{\partial x_i}\right).$$

Assegnata una 1-forma ϑ su M , restano definite le sue rappresentative locali, dette anche *covettori*,⁵¹ che (sempre con piccolo abuso) denoteremo

$$\underline{\vartheta}(\underline{x}) = (\vartheta_1(\underline{x}), \dots, \vartheta_n(\underline{x})).$$

La funzione $\vartheta(X) : M \rightarrow \mathbb{R}$ ha come rappresentativa locale il prodotto di vettore e covettore:

$$\underline{\vartheta}(\underline{x}) \cdot \underline{X}(\underline{x}) = \sum_{i=1}^n \vartheta_i(\underline{x}) X_i(\underline{x}).$$

Se $\underline{\vartheta}$ e $\tilde{\vartheta}$ sono rappresentative locali di una stessa 1-forma in due diverse carte, e $\psi : \tilde{V} \rightarrow V$ è la funzione di transizione dall'una all'altra, allora si ha

$$\tilde{\vartheta}(\tilde{\underline{x}}) = J^T(\tilde{\underline{x}}) \underline{\vartheta}(\psi(\tilde{\underline{x}})) \tag{A.5}$$

⁵¹I covettori, con riferimento alla legge di trasformazione locale (A.5), sono anche chiamati *vettori covarianti*; corrispondentemente i consueti vettori, che invece trasformano secondo la (A.2), sono detti *vettori controvarianti*. La terminologia, tipica del cosiddetto *calcolo tensoriale* o *calcolo differenziale assoluto* (G. Ricci Curbastro, T. Levi-Civita) si estende opportunamente a matrici e in generale a oggetti a più indici, con semplici regole che garantiscono la compatibilità del calcolo in coordinate al variare della carta e corrispondentemente la buona definizione geometrica di ogni oggetto e ogni procedimento.

(attenzione al verso: il confronto con (A.2) mostra che se i vettori vanno con J , le forme vanno con J^{-T} , inversa della trasposta, come si è visto nel paragrafo 2.1 per i momenti p), e corrispondentemente risulta

$$\tilde{\vartheta} \cdot \tilde{X} = \vartheta \cdot X, \quad \text{precisamente} \quad \tilde{\vartheta}(\tilde{x}) \cdot \tilde{X}(\tilde{x}) = \vartheta(\psi(\tilde{x})) \cdot X(\psi(\tilde{x})); \quad (\text{A.6})$$

viceversa se una famiglia di 1-forme locali, una per ogni carta dell'atlante, soddisfa la relazione di compatibilità (A.5), o equivalentemente (A.6), allora resta definita una 1-forma ϑ su M . La 1-forma $\tilde{\vartheta}$ è anche detta *pull-back* di ϑ sotto ψ .

Un caso particolare di 1-forma su M è il differenziale $d\mathcal{F}$ di una 0-forma o funzione $M \rightarrow \mathbb{R}$. La definizione intrinseca di $d\mathcal{F}$ è che per ogni campo vettoriale X su M si ha $d\mathcal{F}(X) = L_X\mathcal{F}$; in coordinate $d\mathcal{F}$ è rappresentata dall' n -pla delle derivate della rappresentativa f di \mathcal{F} , $(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})$, e si ha

$$d\mathcal{F}(X) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i.$$

Una 1-forma che sia il differenziale di una funzione si dice *esatta*.

2-forme su M .

Ricordiamo infine la nozione di 2-forma. Una *2-forma esterna* ω su M , o brevemente *2-forma*, è una mappa bilineare *antisimmetrica* che manda coppie di campi vettoriali su M in funzioni $M \rightarrow \mathbb{R}$: precisamente, a ogni $x \in M$ è associata una mappa bilineare antisimmetrica $\omega(x) : T_x M \times T_x M \rightarrow \mathbb{R}$, e dunque a ogni coppia X, Y di campi vettoriali su M resta associata una funzione $\omega(X, Y) : M \rightarrow \mathbb{R}$, $\omega(Y, X) = -\omega(X, Y)$. La nozione ricorda la nozione di metrica riemanniana, da cui però differisce per il fatto (fondamentale) che l'antisimmetria sostituisce la simmetria. In coordinate una 2-forma è rappresentata da una matrice (forse si dovrebbe dire co-matrice) antisimmetrica $\underline{\omega} = (\omega_{ij})$, precisamente si ha

$$(\omega(X, Y))(x) = \sum_{i,j=1}^n X_i(x) \omega_{ij}(x) Y_j(x), \quad \omega_{ij} = \omega\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right).$$

Per l'antisimmetria si scrive anche (omettiamo l'argomento x)

$$\omega(X, Y) = \sum_{0 \leq i < j \leq n} \omega_{ij}(X_i Y_j - X_j Y_i).$$

Se $\underline{\omega}$ e $\tilde{\underline{\omega}}$ sono rappresentative locali di una stessa 2-forma ω , allora si ha

$$\tilde{\underline{\omega}} = J^T \underline{\omega} J, \quad \text{più precisamente} \quad \tilde{\underline{\omega}}(\tilde{x}) = J^T(\tilde{x}) \underline{\omega}(\psi(\tilde{x})) J(\tilde{x}), \quad (\text{A.7})$$

e viceversa se una famiglia di 2-forme locali, una per ogni carta dell'atlante, soddisfa la relazione di compatibilità (A.7), allora resta definita una 2-forma su M . La 2-forma $\tilde{\underline{\omega}}$ è detta *pull-back* di $\underline{\omega}$ sotto ψ .

Si definisce il *prodotto esterno* di due 1-forme, denotato " \wedge ", il cui risultato è una 2-forma. La definizione è

$$(\vartheta \wedge \vartheta')(X, X') = \vartheta(X)\vartheta'(X') - \vartheta(X')\vartheta'(X); \quad (\text{A.8})$$

si ha evidentemente $\vartheta' \wedge \vartheta = -\vartheta \wedge \vartheta'$ e dunque $\vartheta \wedge \vartheta = 0$. Per ogni carta si possono considerare, in particolare, i prodotti esterni dei differenziali delle coordinate:

$$dx_i \wedge dx_j, \quad 1 \leq i, j \leq n.$$

Di questi, $\frac{n(n-1)}{2}$ sono indipendenti e costituiscono una base per le 2-forme antisimmetriche, ovvero ogni 2-forma esterna si può scrivere come loro combinazione lineare:

$$\omega(x) = \sum_{1 \leq i < j \leq n} \omega_{ij}(x) dx_i \wedge dx_j = \frac{1}{2} \sum_{i,j=1}^n \omega_{ij}(x) dx_i \wedge dx_j .$$

Un altro modo di generare 2-forme esterne da 1-forme è la *differenziazione esterna*: se, con riferimento a una carta qualsiasi, $\vartheta = \sum \vartheta_i(x) dx_i$ è una 1-forma su M , il suo differenziale $d\vartheta$ è per definizione la 2-forma esterna definita da

$$d\vartheta = \sum_{j=1}^n d\vartheta_j \wedge dx_j = \sum_{1 \leq i < j \leq n} \left(\frac{\partial \vartheta_j}{\partial x_i} - \frac{\partial \vartheta_i}{\partial x_j} \right) dx_i \wedge dx_j ;$$

il risultato non dipende dalla carta. Se $d\vartheta = 0$, la forma ϑ si dice *chiusa*. E' chiusa ogni forma esatta, $d(d\mathcal{F}) = 0$, ma vi sono forme chiuse che non sono esatte (non sono il differenziale di una funzione sulla varietà, ma solo di una funzione locale).

Una 2-forma ω si dice *non degenerare* se

$$\omega(X, Y) = 0 \quad \forall Y \quad \Rightarrow \quad X = 0 ;$$

equivalentemente, in ogni carta si ha $\det \underline{\omega}(x) \neq 0$ per ogni x . Data una 2-forma non degenerare ω resta stabilita una corrispondenza biunivoca tra 1-forme e campi vettoriali: infatti per ogni campo vettoriale Y è ben definita la 1-forma $\vartheta = \omega(\cdot, Y)$, ovvero la 1-forma ϑ tale che $\vartheta(X) = \omega(X, Y)$ per ogni X , e viceversa, grazie alla non degenerazione, assegnata ϑ esiste unico Y tale che per ogni X risulti $\vartheta(X) = \omega(X, Y)$ (è immediato in coordinate, usando $\det \underline{\omega} \neq 0$). Se ω è antisimmetrica e non degenerare, necessariamente la dimensione n della varietà è pari: si ha infatti $\det \underline{\omega} = \det \underline{\omega}^T = \det(-\underline{\omega}) = (-1)^n \det \underline{\omega}$.

k-forme esterne (*cenno*).

Accenniamo qui per completezza alle forme di grado più elevato. Una k -forma esterna α è una mappa multilineare, antisimmetrica in tutti i suoi argomenti (muta di segno per ogni permutazione dispari), che manda k campi vettoriali $X^{(1)}, \dots, X^{(k)}$ su M in funzioni $M \rightarrow \mathbb{R}$. Precisamente, a ogni $x \in M$ è associata una mappa multilineare antisimmetrica $\alpha(x)$ che manda $(T_x M)^k$ in \mathbb{R} .

Il prodotto esterno di forme di grado qualsiasi si definisce generalizzando la (A.8): precisamente, date una k -forma α e una k' -forma α' , il loro prodotto esterno $\alpha \wedge \alpha'$ è la $(k + k')$ -forma β definita da

$$\beta(X^{(1)}, \dots, X^{(k+k')}) = \sum_{\substack{j_1 < \dots < j_k \\ j_{k+1} < \dots < j_{k+k'}}} (\pm 1)^p \alpha(X^{(j_1)}, \dots, X^{(j_k)}) \alpha'(X^{(j_{k+1})}, \dots, X^{(j_{k+k'})}) ,$$

ove $(j_1, \dots, j_{k+k'})$ è una permutazione di $(1, \dots, k + k')$, e p è la sua parità. Resta definito in particolare il prodotto esterno di k 1-forme $\vartheta^{(1)}, \dots, \vartheta^{(k)}$,

$$\alpha = \vartheta^{(1)} \wedge \dots \wedge \vartheta^{(k)} ,$$

e precisamente è la k -forma

$$\alpha(X^{(1)}, \dots, X^{(k)}) = \sum_{j_1, \dots, j_k} (-1)^p \vartheta^{(1)}(X^{(j_1)}) \dots \vartheta^{(k)}(X^{(j_k)}) .$$

Una base per le k -forme è data dai prodotti $dx_{j_1} \wedge \cdots \wedge dx_{j_k}$, $j_1 < \cdots < j_k$, e dunque si potrà scrivere

$$\alpha = \sum_{0 \leq j_1 < \cdots < j_k \leq n} \alpha_{j_1, \dots, j_k}(x) dx_{j_1} \wedge \cdots \wedge dx_{j_k} .$$

Infine il differenziale $d\alpha$ della k -forma esterna α è per definizione la $(k+1)$ -forma esterna

$$\beta = \sum_{j_1 < \cdots < j_k} d\alpha_{j_1, \dots, j_k} \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_k} .$$

Una k -forma α si dice *chiusa* se il suo differenziale si annulla. Il differenziale di una k -forma esterna è una $(k+1)$ -forma chiusa per ogni $k \geq 0$. Il viceversa non vale: se la k -forma esterna α è chiusa, in generale essa non è il differenziale di una $(k-1)$ -forma (lo è solo localmente). Se una k forma α è il differenziale di una $(k-1)$ forma, allora essa si dice esatta.

A.2 Sistemi hamiltoniani su varietà simplettiche.

Sistemi lagrangiani e hamiltoniani “naturali”.

Nella costruzione della meccanica lagrangiana (per i cosiddetti “sistemi naturali”, cioè a partire dalla meccanica Newtoniana) si parte da una *varietà vincolare* o spazio delle configurazioni Q , $\dim Q = n$, e in ciascun punto $q \in Q$ si considera lo spazio tangente $T_q Q$ ove vivono le velocità; lo *spazio degli stati* (o spazio degli atti di moto), nel quale vivono i sistemi lagrangiani, è il fibrato tangente TQ , e la lagrangiana è una funzione $\mathcal{L} : TQ \rightarrow \mathbb{R}$. Le carte di Q inducono naturalmente carte su TQ “adattate alla struttura di fibrato”, precisamente le coordinate locali $(\underline{q}, \underline{v})$ sono tali che $\underline{q} = (q_1, \dots, q_n)$ corre su Q , mentre $\underline{v} = (v_1, \dots, v_n)$ corre sulla fibra $T_q Q$, e inoltre le coordinate v_1, \dots, v_n sulla fibra sono proprio le componenti del vettore tangente $v \in T_q Q$ nella base $\frac{\partial}{\partial q_1}, \dots, \frac{\partial}{\partial q_n}$. Le funzioni di transizione tra le carte di questo atlante naturale sono le *trasformazioni puntuali* estese naturalmente alle velocità:

$$\underline{q} = \beta(\tilde{q}) , \quad \underline{v} = B(\tilde{q})\tilde{v} , \quad B = \left(\frac{\partial \beta_i}{\partial \tilde{q}_j} \right) , \quad (\text{A.9})$$

ove β denota la funzione di transizione tra carte di Q .

Detta $l = l(\underline{q}, \underline{v})$ la rappresentativa locale di \mathcal{L} in una qualsiasi carta, si ha⁵² $l = k(\underline{q}, \underline{v}) + u(\underline{q})$, con

$$k(\underline{q}, \underline{v}) = \sum_{i,j=1}^n g_{ij}(\underline{q}) v_i v_j ,$$

$\underline{g} = (g_{ij})$ simmetrica positiva. Le equazioni di Lagrange hanno allora la forma (implicita, ma esplicitabile grazie a $\det \underline{g} \neq 0$)

$$\dot{q}_i = v_i , \quad \frac{d}{dt} \frac{\partial l}{\partial v_i} - \frac{\partial l}{\partial q_i} = 0 , \quad i = 1, \dots, n . \quad (\text{A.10})$$

Resta così definito per ogni carta un campo vettoriale locale $\underline{X} = (X_{q_1}, \dots, X_{q_n}, X_{v_1}, \dots, X_{v_n})$. Come è noto (ed è questa la principale ragione per introdurre il formalismo lagrangiano) tutto è coerente con i cambi di carte entro l’atlante naturale: i campi \underline{X} e $\tilde{\underline{X}}$, ottenuti da l e \tilde{l} tramite le

⁵²Non prendiamo qui in considerazione il caso di vincoli mobili, di coordinate dipendenti dal tempo, o di potenziali dipendenti dalla velocità.

(A.10), si incollano bene, $\tilde{\underline{X}} = B^{-1}\underline{X}$; ciò vuol dire che a $\mathcal{L} : TQ \rightarrow \mathbb{R}$ resta associato un campo vettoriale $X : TQ \rightarrow T(TQ)$. Si osservi che la matrice metrica trasforma secondo la regola delle forme,

$$\tilde{\underline{g}} = B^T \underline{g} B \quad (\text{A.11})$$

(tutte le quantità sono calcolate in \tilde{q}), e ciò garantisce $\tilde{k}(\tilde{q}, \tilde{v}) = k(q, v)$. In linea di principio nulla vieta di arricchire l'atlante di TQ di ulteriori carte “non adattate”, con funzioni di transizione diverse dalle (A.9): a prezzo però di perdere la forma (A.10) delle equazioni. Questo è un motivo di debolezza del formalismo lagrangiano: a meno di rinunciare alla forma lagrangiana delle equazioni, si resta confinati all'atlante elementare naturale, con carte obbligatoriamente connesse da trasformazioni puntuali estese; in particolare non si possono usare coordinate come i momenti angolari o le energie (le tipiche quantità conservate della meccanica), che mescolano tra loro posizioni e velocità.

I sistemi hamiltoniani naturali sono comunemente introdotti a partire dai sistemi lagrangiani, tramite la trasformata di Legendre. Operando localmente carta per carta, a ogni vettore \underline{v} si associa il covettore

$$\underline{p} = \frac{\partial k}{\partial v} = g\underline{v} , \quad (\text{A.12})$$

detto momento, che in conseguenza alla (A.11) trasforma secondo la regola dei covettori

$$\tilde{\underline{p}} = B^T \underline{p} . \quad (\text{A.13})$$

Alla lagrangiana l si associa h , detta hamiltoniana locale, definita da

$$h(\underline{p}, \underline{q}) = \underline{p} \cdot \underline{g}^{-1} \underline{p} - l(\underline{q}, \underline{g}^{-1} \underline{p}) ,$$

e si osserva che le equazioni locali per $(\underline{p}, \underline{q})$, equivalenti alle equazioni di Lagrange (A.10), sono le equazioni di Hamilton, o equazioni canoniche:

$$\dot{p}_i = -\frac{\partial h}{\partial q_i} , \quad \dot{q}_i = \frac{\partial h}{\partial p_i} , \quad i = 1, \dots, n . \quad (\text{A.14})$$

A h è dunque associato il campo vettoriale locale \underline{X}_h , detto campo vettoriale hamiltoniano (locale) di hamiltoniana h ,

$$\underline{X}_h = (X_{h\underline{p}}, X_{h\underline{q}}) = \left(-\frac{\partial h}{\partial \underline{q}}, \frac{\partial h}{\partial \underline{p}} \right) .$$

Dal quadro locale si passa facilmente al livello geometrico. Per questo accanto al fibrato tangente TQ si considera il fibrato cotangente T^*Q , che al pari di TQ è dotato naturalmente di un atlante adattato alla struttura di fibrato: le coordinate \underline{q} corrono su Q , mentre le coordinate \underline{p} corrono sulla fibra T_q^*Q . Le funzioni di transizione dell'atlante sono ancora trasformazioni puntuali, estese questa volta ai momenti: precisamente hanno la forma, analoga alle (A.9),

$$\underline{p} = B^{-T}(q)\tilde{\underline{p}} , \quad \underline{q} = \beta(\tilde{q}) . \quad (\text{A.15})$$

La corrispondenza tra vettori e covettori stabilita dalla (A.12) è intrinseca; ne segue che le diverse funzioni di Hamilton locali sono rappresentative di una funzione⁵³ $\mathcal{H} : T^*Q \rightarrow \mathbb{R}$. Infine, i diversi

⁵³La trasformata di Legendre (della quale si può dare una definizione geometrica intrinseca) si estende a generiche funzioni convesse (anche non quadratiche) di \underline{v} ; la funzione trasformata risulta essa stessa convessa, e la trasformazione è involutoria.

campi vettoriali hamiltoniani locali si incollano bene: la matrice jacobiana del cambio di carta è infatti

$$J(\underline{\tilde{p}}, \underline{\tilde{q}}) = \begin{pmatrix} B^{-T} & \mathbb{O} \\ \mathbb{O} & B \end{pmatrix}, \quad (\text{A.16})$$

e si verifica subito che $\underline{\tilde{X}} = J^{-1}\underline{X}$. Restano così definiti un campo vettoriale hamiltoniano $X_{\mathcal{H}}$ su T^*Q , $X_{\mathcal{H}} : T^*Q \rightarrow T(T^*Q)$ e un'equazione differenziale su T^*Q , di cui le equazioni di Hamilton (A.10) sono rappresentative locali nelle carte dell'atlante adattato.

Come si è sottolineato più volte in queste note, uno dei maggiori motivi di interesse del formalismo hamiltoniano, che lo distingue da quello lagrangiano, è il fatto che l'atlante adattato di T^*Q , formato da trasformazioni puntuali estese, può essere arricchito in modo significativo senza che per questo si perda la forma delle equazioni di Hamilton. Per capire, e allo stesso tempo per preparare la strada a una nozione intrinseca di sistema hamiltoniano (in un quadro più ampio di quello dei sistemi naturali), vale la pena di fare un passo indietro, e porre la questione della compatibilità fra equazioni di Hamilton in carte diverse in termini più generali. Sia allora M una qualunque varietà di dimensione pari, $\dim M = 2n$; sia \mathcal{A} il suo atlante, e per ogni carta di \mathcal{A} denotiamo $\underline{x} = (\underline{p}, \underline{q})$. Ci chiediamo quali requisiti debba avere \mathcal{A} perché, presa una qualsiasi funzione $\mathcal{H} : M \rightarrow \mathbb{R}$, e scritte in ciascuna carta le equazioni di Hamilton (A.14), queste abbiano carattere geometrico, cioè siano rappresentative locali di un'equazione differenziale sulla varietà. È la domanda di base che ci siamo posti all'inizio di queste note, sulla quale abbiamo appoggiato la nozione di trasformazione canonica; qui semplicemente ripercorriamo la stessa idea da una visuale geometrica, giungendo in modo più naturale alla caratterizzazione delle trasformazioni canoniche data attraverso la proposizione 1.

Sia allora h la rappresentativa locale di \mathcal{H} in una carta, e \underline{X}_h il suo campo vettoriale hamiltoniano locale, $\underline{X}_h = (-\frac{\partial h}{\partial \underline{q}}, \frac{\partial h}{\partial \underline{p}})$. Introduciamo la matrice $2n \times 2n$

$$\mathbb{E} = \begin{pmatrix} \mathbb{O} & -\mathbb{I} \\ \mathbb{I} & \mathbb{O} \end{pmatrix}, \quad \mathbb{E}^T = -\mathbb{E} = \mathbb{E}^{-1},$$

ove \mathbb{I} denota la matrice identità $n \times n$. La relazione tra h e \underline{X}_h si scrive allora nella forma sintetica

$$\mathbb{E}^T \underline{X}_h = \left(\frac{\partial h}{\partial x_1}, \dots, \frac{\partial h}{\partial x_{2n}} \right), \quad (\text{A.17})$$

adatta a discuterne il carattere geometrico (ben al di là del pretesto, usato nel paragrafo 2.1, della compattezza della notazione.) Infatti a destra dell'uguale appare l' n -pla delle derivate di h , che è rappresentativa locale di una 1-forma su M , precisamente il differenziale $d\mathcal{H}$. Si capisce allora che *i diversi campi locali \underline{X}_h sono rappresentativi di un campo vettoriale $X_{\mathcal{H}}$ su M , qualunque sia \mathcal{H} , se⁵⁴ la medesima matrice⁵⁵ \mathbb{E}^T antisimmetrica (e non singolare) è rappresentativa locale, in ogni carta, di una 2-forma ω (non degenera) su M . In tal caso infatti la (A.17) diventa globale, e precisamente si innalza a*

$$\omega(\cdot, X_{\mathcal{H}}) = d\mathcal{H}. \quad (\text{A.18})$$

Sia $\psi : \underline{\tilde{x}} \mapsto \underline{x}$ la funzione di transizione tra due carte di \mathcal{A} , e sia J la corrispondente matrice jacobiana. Dalla regola (A.7) sul *pull-back* di forme si vede che *perché una 2-forma ω su M sia rappresentata da \mathbb{E}^T in entrambe le carte, bisogna e basta che risulti*

$$J^T \mathbb{E} J = \mathbb{E}. \quad (\text{A.19})$$

⁵⁴La condizione è ovviamente sufficiente, ma è anche necessaria se vogliamo che la proprietà sia vera per ogni \mathcal{H} .

⁵⁵L'uso di \mathbb{E}^T al posto di \mathbb{E} qui è un po' forzato. Nei testi più orientati geometricamente è denotata \mathbb{E} quella che qui è \mathbb{E}^T ; la scelta più infelice fatta qui è unicamente motivata dalla compatibilità con la (1.2).

Matrici J che soddisfino la (A.19) si dicono *matrici simplettiche*. Si verifica facilmente che esse formano un gruppo, detto *gruppo simplettico* (in particolare dunque J^{-1} è simplettica), e inoltre che se J è simplettica lo è anche la trasposta J^T ; \mathbb{E} stessa è simplettica, e anzi è detta *identità simplettica*.

DEFINIZIONE 14 *Si dice che due carte di M , $\dim M = 2n$, sono canonicamente compatibili, ovvero che la funzione di transizione dall'una all'altra è una trasformazione canonica, se la matrice jacobiana della funzione di transizione è simplettica. Un atlante di carte canonicamente compatibili si dice atlante simplettico, o anche canonico.*

E' la stessa nozione di canonicità (stretta) introdotta in coordinate nel primo capitolo. La compatibilità tra le equazioni di Hamilton locali è così assicurata ogni volta che si ha una varietà M di dimensione pari, munita di un atlante simplettico.

Tornando ai sistemi naturali su T^*Q , da cui questo discorso è partito, si vede bene che la matrice jacobiana (A.16) è simplettica, e dunque l'atlante naturale di T^*Q considerato sopra è canonico (non poteva essere altrimenti). Si comprende anche che tale atlante si può allargare, mantenendo in ogni carta la forma hamiltoniana delle equazioni, fintantoché le funzioni di transizione restano trasformazioni canoniche (nulla vieta poi l'introduzione di carte qualsiasi, a prezzo di perdere la forma hamiltoniana delle equazioni locali).

Varietà simplettiche.

Riprendiamo il caso generale di una varietà M di dimensione pari munita di un atlante simplettico. La 2-forma ω su M , di cui la matrice \mathbb{E}^T è rappresentativa locale in ogni carta, è con evidenza

$$\omega = dp \wedge dq = \sum_{i=1}^n dp_i \wedge dq_i \quad (\text{A.20})$$

(ovvero $\underline{\omega} = \mathbb{E}^T$); la forma, oltre che ben definita su M , è non degenera e anche chiusa, ovvia conseguenza del fatto che ha coefficienti costanti. Essa è detta *2-forma canonica*. Un modo assai frequente di scrivere la relazione di compatibilità canonica (A.19) è

$$dp \wedge dq = d\tilde{p} \wedge d\tilde{q} ; \quad (\text{A.21})$$

la relazione si legge comunemente dicendo che *le trasformazioni canoniche preservano la 2-forma canonica* (A.20). La (A.21) è adoperata spesso come definizione di trasformazione canonica. In generale la 2-forma canonica è chiusa ma non è esatta. Localmente essa appare come il differenziale della 1-forma

$$\vartheta = p \cdot dq = \sum_{i=1}^n p_i dq_i ,$$

detta *1-forma canonica locale* o 1-forma di Liouville:

$$\omega = d\vartheta .$$

Ma in generale la 1-forma ϑ non esiste globalmente su M , e la differenza $p \cdot dq - \tilde{p} \cdot d\tilde{q}$ è chiusa ma non nulla; equivalente alla (A.21) è dunque la preservazione debole, ovvero a meno del differenziale di una funzione f definita localmente, della 1-forma di Liouville:

$$p \cdot dq = \tilde{p} \cdot d\tilde{q} + df .$$

Si ritrova per questa via una fondamentale caratterizzazione delle trasformazioni canoniche. Come si è visto nel testo, la 1-forma ϑ è preservata da tutte le trasformazioni puntuali estese.

E' spontanea ora la seguente definizione:

DEFINIZIONE 15 *Una varietà differenziabile di dimensione pari munita di un atlante simplettico si dice varietà simplettica.*

E poi:

DEFINIZIONE 16 *Sia M una varietà simplettica, e \mathcal{H} una funzione $M \rightarrow \mathbb{R}$. Si dice campo vettoriale hamiltoniano associato a \mathcal{H} , e si denota $X_{\mathcal{H}}$, il campo vettoriale definito dalla (A.18), ove ω è la 2-forma canonica su M . L'equazione differenziale sulla varietà*

$$\dot{x} = X_{\mathcal{H}} , \tag{A.22}$$

di cui la (A.14) è rappresentativa locale, è detta sistema hamiltoniano su M .

Grazie alla presenza di un atlante simplettico su M , o come si suol dire di una struttura simplettica, la nozione di sistema hamiltoniano è ben data⁵⁶

Accenniamo per completezza a una formulazione più profonda della nozione di varietà simplettica, e attraverso essa di sistema hamiltoniano. La via che abbiamo seguito sopra è stata quella di partire da equazioni locali in forma hamiltoniana e di chiederci se la definizione locale può essere estesa globalmente; in questo modo ci siamo imbattuti nella 2-forma canonica $\omega = dp \wedge dq$, e abbiamo preteso che ω fosse ben definita su M grazie alla (A.19) o equivalentemente alla (A.21). Per questa via ci siamo imbattuti nella nozione di atlante canonico, di varietà simplettica, e infine di sistema hamiltoniano. Della 2-forma $\omega = dp \wedge dq$ abbiamo osservato che essa è antisimmetrica, non degenera e anche chiusa.

E' spontaneo chiedersi se non ci si possa liberare del tutto dalle coordinate, e *definire varietà simplettica una coppia (M, ω) , con M varietà differenziabile di dimensione pari e ω una qualsiasi 2-forma chiusa antisimmetrica non degenera su M (senza cioè chiedere a priori che M sia munito di un atlante particolare, nel quale ω abbia la forma particolare $dp \wedge dq$)*. La definizione (A.18) di $X_{\mathcal{H}}$ resterebbe poi immutata.

La risposta è che la cosa si può fare, ma che la nuova nozione apparentemente più generale di varietà simplettica è di fatto identica alla precedente. Vale infatti la seguente proposizione:⁵⁷

PROPOSIZIONE 30 (Teorema di Darboux) *Sia M una varietà differenziabile di dimensione pari, e ω una 2-forma chiusa antisimmetrica non degenera su M . Allora esiste un atlante di M tale che per tutte le sue carte la 2-forma ω assume la forma canonica (A.20).*

⁵⁶Una volta chiarito quale sia la struttura simplettica di riferimento, resta sempre la libertà di aggiungere all'atlante carte arbitrarie non simplettiche: le equazioni del moto perdono la forma hamiltoniana locale (A.14), ma il sistema — che non è cambiato — resta hamiltoniano.

⁵⁷Il teorema è conseguenza di un lemma di algebra lineare, secondo il quale *se la matrice Ω di ordine $2n$ è antisimmetrica e non degenera, allora esiste J tale che*

$$J^T \Omega J = \mathbb{E} .$$

Il risultato ricorda il più noto risultato secondo cui, se Ω è simmetrica e definita positiva, allora esiste J tale che $J^T \Omega J = I$.

Parentesi di Poisson.

Già nel paragrafo 1.2 si è sottolineato il fatto che la parentesi di Poisson tra funzioni, pur definita in coordinate tramite la (1.7) o equivalentemente la (1.8), è in realtà indipendente dalle coordinate purchè il cambio di coordinate sia canonico. Anzi: l'invarianza dell'operazione di parentesi di Poisson si è vista essere una proprietà caratteristica delle trasformazioni canoniche (proposizione 4). Nel linguaggio geometrico questo vuol semplicemente dire che, se \mathcal{F} e \mathcal{G} sono funzioni: $M \rightarrow \mathbb{R}$, allora l'operazione in coordinate tra rappresentative locali in una qualunque carta dell'atlante симпlettico definisce sulla varietà l'operazione binaria di parentesi di Poisson $\{\mathcal{F}, \mathcal{G}\}$. In modo del tutto equivalente si può definire la parentesi di Poisson in modo intrinseco geometrico, precisamente

$$\{\mathcal{F}, \mathcal{G}\} = \omega(X_{\mathcal{G}}, X_{\mathcal{F}})$$

(la verifica è immediata), o ancora

$$\{\mathcal{F}, \mathcal{G}\} = F_{\mathcal{G}}\mathcal{F} ,$$

con la notazione agile $F_{\mathcal{G}} = F_{X_{\mathcal{G}}}$ consueta nel testo.

La parentesi di Poisson risulta in realtà essere la particolarizzazione hamiltoniana dell'operazione di parentesi di Lie. Precisamente, per campi vettoriali hamiltoniani avviene che la loro parentesi di Lie è ancora un campo hamiltoniano, e si ha

$$[X_{\mathcal{F}}, X_{\mathcal{G}}] = X_{\{\mathcal{G}, \mathcal{F}\}} ;$$

equivalentemente

$$[L_{\mathcal{F}}, L_{\mathcal{G}}] = L_{\{\mathcal{G}, \mathcal{F}\}} .$$

Si riconosce in questa espressione la (1.10). Come già commentato nel testo, tale espressione altro non è che l'identità di Jacobi

$$\{\{\mathcal{F}, \mathcal{G}\}, \mathcal{H}\} + \{\{\mathcal{H}, \mathcal{F}\}, \mathcal{G}\} + \{\{\mathcal{G}, \mathcal{H}\}, \mathcal{F}\} = 0 ,$$

versione hamiltoniana della (A.4).

B Equazioni differenziali che preservano il volume

Sia data l'equazione differenziale in \mathbb{R}^n

$$\dot{x} = X(x) ;$$

denotiamo con $\nabla \cdot X$ la divergenza di X , $\nabla \cdot X = \sum_i \frac{\partial X_i}{\partial x_i}$, e con $\Phi^t(x)$ il flusso di X . A un generico insieme A è naturalmente associato l'insieme evoluto $\Phi^t(A)$ all'istante t (figura 36).

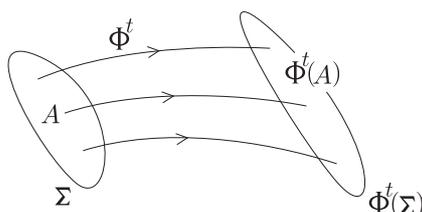


Figura 36: L'insieme A evolve sotto il flusso Φ^t .

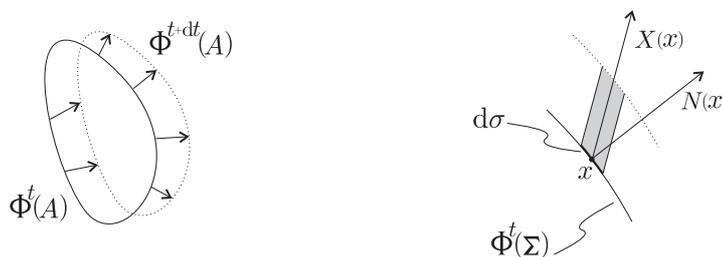


Figura 37: Un'illustrazione della dimostrazione 1.

PROPOSIZIONE 31 *Sia A misurabile. Allora il volume $\text{Vol}(\Phi^t(A))$ varia nel tempo secondo la legge*

$$\frac{d}{dt} \text{Vol}(\Phi^t(A)) = \int_{\Phi^t(A)} \nabla \cdot X \, dV . \quad (\text{B.1})$$

COROLLARIO 32 *Il flusso hamiltoniano preserva il volume in \mathbb{R}^n .*

DIMOSTRAZIONE DEL COROLLARIO. Il campo vettoriale hamiltoniano ha divergenza nulla: infatti $X_{p_i} = -\frac{\partial H}{\partial q_i}$ e $X_{q_i} = \frac{\partial H}{\partial p_i}$, e dunque

$$\nabla \cdot X = \sum_{i=1}^n \left(-\frac{\partial^2 H}{\partial p_i \partial q_i} + \frac{\partial^2 H}{\partial q_i \partial p_i} \right) = 0 .$$

□

Si può dimostrare la proposizione 31 in poche righe nel caso in cui A abbia un bordo regolare Σ , se si accetta di servirsi del linguaggio un po' euristico in cui si trattano i differenziali come quantità piccole.

DIMOSTRAZIONE 1. Nell'intervallo tra t e $t + dt$ l'evoluto di A passa da $\Phi^t(A)$ a $\Phi^{t+dt}(A)$, e in questo movimento l'elemento di superficie $d\sigma$ spazza il volume $X \cdot N \, d\sigma \, dt$, ove N è il versore normale alla superficie $\Phi^t(\Sigma)$ di $\Phi^t(A)$, rivolto verso l'esterno; si veda la figura 37. Segue

$$\frac{d}{dt} \text{Vol}(\Phi^t(A)) = \int_{\Phi^t(\Sigma_A)} X \cdot N \, d\sigma .$$

Per il teorema della divergenza l'integrale di superficie si trasforma in integrale di volume e si ottiene la (B.1). □

Una dimostrazione più generale, e senza nulla di euristico, si ottiene sulla base del seguente

LEMMA 33 *Sia J^t la matrice jacobiana di Φ^t ,*

$$J_{ij}^t(x) = \frac{\partial \Phi_i^t(x)}{\partial x_j} .$$

Risulta

$$\frac{d}{dt} \det J^t(x) = \det J^t(x) \nabla \cdot X(\Phi^t(x)) .$$

DEFINIZIONE 17 Sia A un insieme misurabile. Un punto $x \in A$ si dice *ricorrente in A* , se per ogni $T > 0$ esiste $t \geq T$ tale che $\Phi^t(x) \in A$. Un punto non ricorrente in A si dice invece *vagante*.

L'arbitrarietà di T implica che i punti ricorrenti rientrano infinite volte in A (ma la ricorrenza non implica alcuna regolarità nei tempi di ritorno in A). I punti vaganti invece, da un qualche T in poi, non tornano più in A . Ad esempio per il pendolo, se prendiamo come insieme A una palla a cavallo di una separatrice, i punti sulla separatrice sono vaganti (una volta usciti da A non rientrano), tutti gli altri sono periodici e dunque ricorrenti. Per un sistema dissipativo come l'oscillatore armonico smorzato, qualunque insieme A che non contenga il punto di equilibrio è composto interamente di punti vaganti; qualunque insieme contenente il punto di equilibrio è invece composto interamente di punti ricorrenti. Denoteremo con V_A l'insieme dei punti vaganti di A :

$$V_A = \{x \in A : \exists T > 0 : \Phi^t(x) \notin A \forall t \geq T\} .$$

Il teorema della ricorrenza riguarda equazioni differenziali che preservano il volume (in particolare dunque i sistemi hamiltoniani) e sono definite in un dominio $U \subset \mathbb{R}^n$ limitato invariante, cioè tale che $\Phi^t(U) \subset U$ per ogni $t > 0$ (ad esempio un dominio racchiuso da una superficie di energia costante limitata, o compreso tra due di queste superfici; si pensi a un gas contenuto in una scatola, con energia potenziale limitata inferiormente).

PROPOSIZIONE 34 Sia data l'equazione differenziale $\dot{x} = X(x)$ in un dominio $U \subset \mathbb{R}^n$ limitato invariante. Se il flusso Φ^t preserva il volume, allora per ogni A misurabile l'insieme V_A è misurabile e ha volume nullo.

DIMOSTRAZIONE. Innanzitutto ci restringiamo a t intero, e definiamo

$$\tilde{V}_A = \{x \in A : \exists T > 0 : \Phi^t(x) \notin A \forall t \in \mathbb{N}^+, t \geq T\} .$$

La condizione che definisce l'insieme è indebolita, perciò $V_A \subset \tilde{V}_A$, e dunque sarà sufficiente mostrare che \tilde{V}_A è misurabile e ha volume nullo. Poniamo poi

$$V_{A,T} = \{x \in A : \Phi^t(x) \notin A \forall t \in \mathbb{N}^+, t \geq T\}$$

(punti che vagano da un certo T in poi). Si vede subito che $V_{A,T} \subset V_{A,T'}$ per $T' > T$ e che

$$\tilde{V}_A = \bigcup_{T \in \mathbb{N}^+} V_{A,T} .$$

L'unione è numerabile, perciò basta mostrare che ciascuno dei $V_{A,T}$ è misurabile e ha volume nullo.

La misurabilità di $V_{A,T}$ segue dall'espressione, di facile verifica,

$$V_{A,T} = A \cap \left[\bigcap_{t \in \mathbb{N}^+, t \geq T} \Phi^{-t}(U \setminus A) \right]$$

(verifica: $x \in V_{A,T}$ se e solo se $x \in A$ e $\Phi^t(x) \in U \setminus A$ per ogni $t \in \mathbb{N}^+$, $t \geq T$). L'espressione mostra che $V_{A,T}$ è intersezione numerabile di insiemi misurabili, dunque è misurabile. Per mostrare che $V_{A,T}$ ha volume nullo, consideriamo gli insiemi

$$V_{A,T} , \quad \Phi^T(V_{A,T}) , \quad \Phi^{2T}(V_{A,T}) , \quad \dots , \quad \Phi^{kT}(V_{A,T}) , \quad \dots$$

Essi sono necessariamente a due a due disgiunti: infatti, nessuno di essi può intersecare $V_{A,T}$ (se $x \in \Phi^{kT}(V_{A,T})$ sta anche in $V_{A,T} \subset A$, si viola la definizione di $V_{A,T}$). Ma non può neanche esistere

$$x \in \Phi^{kT}(V_{A,T}) \cap \Phi^{lT}(V_{A,T})$$

per $k \neq l$, ad esempio $k > l$: altrimenti $y = \Phi^{-lT}(x)$ starebbe sia in $\Phi^{(k-l)T}(V_{A,T})$ che in $V_{A,T}$, e si ricade nel caso precedente. Ma poiché il flusso preserva il volume, tutti questi infiniti insiemi disgiunti hanno lo stesso volume, e poiché la loro unione è contenuta in U di volume finito, ciascuno ha necessariamente volume nullo. \square

COROLLARIO 35 *Sia preso ad arbitrio $\varepsilon > 0$. Per tutti i punti x di U , a meno di un insieme di misura nulla, l'evoluto $\Phi^t(x)$ ritorna infinite volte a distanza non superiore a ε da x .*

DIMOSTRAZIONE. Basta ricoprire U con palle di diametro ε , e applicare il teorema a ciascuna di esse. \square

C Commutazione di flussi hamiltoniani

Sia f una funzione: $U \rightarrow \mathbb{R}$. La “mancata commutazione” dei flussi Φ_H^s e Φ_K^t di due hamiltoniane H e K , vista attraverso f , è ben rappresentata dalla quantità

$$\Delta_f(s, t) = f \circ \Phi_K^t \circ \Phi_H^s - f \circ \Phi_H^s \circ \Phi_K^t .$$

E' chiaro che se i flussi commutano allora Δ_f è nullo per ogni f , e viceversa. Si dimostra allora la seguente proposizione:

PROPOSIZIONE 36 *Con riferimento alla notazione sopra introdotta,*

- i) *per ogni funzione regolare $f : U \rightarrow \mathbb{R}$ si ha $\Delta_f(s, t) = st \{ \{H, K\}, f \} + \mathcal{O}_3(s, t)$;*
- ii) *se $\{H, K\} = 0$ allora i flussi commutano (e viceversa, per il punto precedente).*

DIMOSTRAZIONE. La dimostrazione del primo punto è un semplice calcolo: si ha con evidenza $\Delta_f(0, t) = \Delta_f(s, 0) = 0$, e dunque con uno sviluppo di Taylor

$$\Delta_f(s, t) = st \left. \frac{\partial^2 \Delta_f}{\partial s \partial t} \right|_{s=t=0} + \mathcal{O}_3(s, t) .$$

Si calcola poi facilmente, posto $x = (p, q)$:

$$\left. \frac{\partial f}{\partial t} (\Phi_K^t(\Phi_H^s(x))) \right|_{t=0} = (\mathcal{L}_K f)(\Phi_H^s(x)) , \quad \left. \frac{\partial^2 f}{\partial s \partial t} (\Phi_K^t(\Phi_H^s(x))) \right|_{s=t=0} = (\mathcal{L}_H \mathcal{L}_K f)(x) ,$$

e analogamente $\left. \frac{\partial^2 f}{\partial s \partial t} (\Phi_H^s(\Phi_K^t(x))) \right|_{s=t=0} = (\mathcal{L}_K \mathcal{L}_H f)(x)$. Segue, come richiesto,

$$\left. \frac{\partial^2 \Delta_f}{\partial s \partial t} \right|_{s=t=0} = [\mathcal{L}_H, \mathcal{L}_K] f = -\mathcal{L}_{\{H, K\}} f = \{ \{H, K\}, f \}$$

(si è usata la 1.10).

Per quanto riguarda il secondo punto, in un verso la prova è banale (se $\Delta_f = 0$ per ogni f certamente $\{H, K\} = 0$), mentre per l'altro verso ci vuole un po' di lavoro. Si procede così: si prende $k > 1$ qualsiasi, e si pone $\sigma = s/k$, $\tau = t/k$, cosicché

$$\Phi_K^t \circ \Phi_H^s = \Phi_K^\tau \circ \cdots \circ \Phi_K^\tau \circ \Phi_H^\sigma \circ \cdots \circ \Phi_H^\sigma .$$

Per quanto già dimostrato, prendendo successivamente $f(p, q) = p_1, p_2, \dots, q_n$, si trova

$$\Phi_K^\tau \circ \Phi_H^\sigma = \Phi_H^\sigma \circ \Phi_K^\tau + \mathcal{O}(k^{-3}) ,$$

e se λ denota la costante di Lipshitz delle equazioni di Hamilton di K segue

$$\Phi_K^{(k-1)\tau} \circ \Phi_K^\tau \circ \Phi_H^\sigma \circ \Phi_H^{(k-1)\sigma} = \Phi_K^{(k-1)\tau} \circ \Phi_H^\sigma \circ \Phi_K^\tau \circ \Phi_H^{(k-1)\sigma} + \mathcal{O}(e^{\lambda(k-1)\sigma} k^{-3}) .$$

Iterando, con un numero di scambi pari a k si porta Φ_H^σ in testa, e con $\mathcal{O}(k^2)$ scambi si portano a uno a uno in testa tutti i Φ_H^σ . L'errore ogni volta è di ordine k^{-3} per un coefficiente che non supera

$$c = e^{\lambda(l\sigma + (k-l-1)\tau)} \leq e^{\lambda \max(s,t)} ,$$

avendo qui supposto che λ sia costante di Lipshitz anche per H (l è il numero di fattori Φ_K^τ rimasti a sinistra, $k - l - 1$ il numero di fattori Φ_H^σ già portati in testa). Segue

$$\Phi_K^t \circ \Phi_H^s - \Phi_H^s \circ \Phi_K^t = \mathcal{O}(k^2 k^{-3}) = \mathcal{O}(k^{-1}) ,$$

e per l'arbitrarietà di k si conclude che i flussi commutano. □

Il problema della commutazione dei flussi si pone anche al di fuori dell'ambito hamiltoniano, e si risolve sostanzialmente allo stesso modo. In breve: siano X e Y due campi vettoriali (ovvero due sistemi di equazioni differenziali in \mathbb{R}^n), e siano $\mathcal{L}_X, \mathcal{L}_Y$ le corrispondenti derivate di Lie. Si definisce il commutatore $Z = [X, Y]$ di X e Y ponendo

$$Z_i = \mathcal{L}_X Y_i - \mathcal{L}_Y X_i .$$

Non è difficile verificare che se X e Y sono campi vettoriali hamiltoniani corrispondenti alle hamiltoniane H e K , allora anche Z è hamiltoniano e la sua funzione di Hamilton è $-\{H, K\}$. Il risultato visto sopra nel caso hamiltoniano si generalizza a campi generici, precisamente si dimostra (con identica dimostrazione) che i flussi associati a X e Y commutano se e solo se i campi commutano:

$$[X, Y] = 0 \iff \Phi_X^s \circ \Phi_Y^t = \Phi_Y^t \circ \Phi_X^s \quad \forall s, t \in \mathbb{R} .$$

D Alcune dimostrazioni riguardanti le trasformazioni canoniche

- *Dimostrazione della Proposizione 7*

La dimostrazione si basa sul seguente lemma:

LEMMA 37 *Se la matrice $n \times n$ $A(x)$, $x \in \mathbb{R}^n$, è tale che invia il gradiente $\partial_x f$ di una qualunque funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ nel gradiente di una qualche funzione $g : \mathbb{R}^n \rightarrow \mathbb{R}$, allora $A = cI$ con c costante.*

DIMOSTRAZIONE DEL LEMMA. L'ipotesi è che per ogni f il campo vettoriale $A\partial_x f$ sia un gradiente e dunque che per ogni coppia di indici i, j risulti, qualunque sia f ,

$$\frac{\partial}{\partial x_j}(A\partial_x f)_i = \frac{\partial}{\partial x_i}(A\partial_x f)_j . \quad (\text{D.1})$$

Si giunge facilmente alla conclusione imponendo questa relazione, per ogni coppia i, j , per $f(x) = x_j$, $f(x) = x_j^2$ e $f(x) = x_i x_j$. Infatti, prendendo $f(x) = x_j$ si trova subito

$$\frac{\partial A_{ij}}{\partial x_j} = \frac{\partial A_{jj}}{\partial x_i} . \quad (\text{D.2})$$

Prendendo poi $f(x) = x_j^2$, e tenendo conto della (D.2), si trova

$$A_{ij} = A_{jj}\delta_{ij} ,$$

ovvero la matrice è diagonale. A questo punto la (D.2) dà

$$\frac{\partial A_{jj}}{\partial x_i} = 0 \quad \text{per } i \neq j , \quad (\text{D.3})$$

ovvero l'elemento diagonale A_{jj} è funzione solo di x_j . Grazie a quest'ultimo fatto e al fatto che A è diagonale, la (D.1) si riduce a

$$A_{ii} \frac{\partial^2 f}{\partial x_j \partial x_i} = A_{jj} \frac{\partial^2 f}{\partial x_i \partial x_j} ,$$

e per $f(x) = x_i x_j$ si trova $A_{ii} = A_{jj}$, ovvero la matrice è della forma $c(x)I$. Ma ciascun elemento A_{jj} dipende solo da x_j , perciò necessariamente c è costante. \square

DIMOSTRAZIONE DELLA PROPOSIZIONE. Per ipotesi i due campi vettoriali

$$X = \mathbb{E} \partial_x H , \quad \tilde{X} = \mathbb{E} \partial_{\tilde{x}} K$$

sono legati da $X \circ w = J(\tilde{x})\tilde{X}$ (equivalenza delle equazioni di Hamilton), ovvero si ha

$$\mathbb{E} \partial_x H \circ w = J(\tilde{x}) \mathbb{E} \partial_{\tilde{x}} K .$$

Posto $\tilde{H} = H \circ w$, tenendo conto della (2.11) si ottiene

$$\mathbb{E} J^{-T} \partial_{\tilde{x}} \tilde{H} = J(\tilde{x}) \mathbb{E} \partial_{\tilde{x}} K$$

e dunque

$$\partial_{\tilde{x}} K = - \mathbb{E} J^{-1} \mathbb{E} J^{-T} \partial_{\tilde{x}} \tilde{H} .$$

Si vede così che la matrice $A = - \mathbb{E} J^{-1} \mathbb{E} J^{-T}$ manda un qualunque gradiente in un gradiente; dal lemma segue allora $\partial_{\tilde{x}} K = c \partial_{\tilde{x}} \tilde{H}$, ovvero (a meno di una costante additiva) $K = c\tilde{H} = cH \circ w$, come preteso. Si osservi che da $A = cI$ si trova anche, come deve essere,

$$cJ \mathbb{E} J^T = \mathbb{E}$$

(J appartiene al gruppo simplettico allargato). \square

- *Dimostrazione della Proposizione 10*

DIMOSTRAZIONE. Scriviamo la 2.24 nella forma

$$p = \tilde{p} + \varepsilon a(\tilde{p}, \tilde{q}, \varepsilon) , \quad q = \tilde{q} + \varepsilon b(\tilde{p}, \tilde{q}, \varepsilon) .$$

Per ε piccolo le equazioni certamente si invertono rispetto a \tilde{q} e si ottiene

$$p = \tilde{p} + \varepsilon \alpha(\tilde{p}, q, \varepsilon) , \quad \tilde{q} = q + \varepsilon \beta(\tilde{p}, q, \varepsilon) ,$$

con α, β opportune; dobbiamo allora mostrare che esiste $\mathcal{S}(\tilde{p}, q, \varepsilon)$ tale che

$$\alpha(\tilde{p}, q, \varepsilon) = \frac{\partial \mathcal{S}}{\partial q}(\tilde{p}, q, \varepsilon) , \quad \beta(\tilde{p}, q, \varepsilon) = \frac{\partial \mathcal{S}}{\partial \tilde{p}}(\tilde{p}, q, \varepsilon) . \quad (\text{D.4})$$

Per questo osserviamo che in base alla condizione di canonicità (2.21) (la preservazione debole della 1-forma di Liouville, riscritta) esiste g tale che

$$(\tilde{p} + \varepsilon \alpha(\tilde{p}, q, \varepsilon)) \cdot dq + (q + \varepsilon \beta(\tilde{p}, q, \varepsilon)) \cdot d\tilde{p} = dg ;$$

per $\varepsilon = 0$ si ha (a meno di una costante) $g = \tilde{p} \cdot q$, pertanto

$$g(p, q, \varepsilon) = \tilde{p} \cdot q + \varepsilon \mathcal{S}(p, q, \varepsilon)$$

con opportuna \mathcal{S} . Segue immediatamente

$$\alpha \cdot dq + \beta \cdot d\tilde{p} = d\mathcal{S} ,$$

equivalente alla (D.4). □

E Risoluzione di alcuni esercizi

- *Esercizio 19*

Si ha un punto di equilibrio quando tutte le derivate parziali di K si annullano. Derivando la (2.38) si ottiene subito per il punto di equilibrio $(\tilde{p}^*, \tilde{q}^*)$ la condizione

$$\tilde{p}^* - m\Omega \tilde{q}^* = 0 , \quad V'(\tilde{r}) - m\omega^2 \tilde{r} = 0 ;$$

in particolare si trova la soluzione di equilibrio cercata, precisamente

$$(\tilde{p}^*, \tilde{q}^*) = (0, m\omega a, a, 0) ,$$

purché a sia tale che $V'(a) = m\omega^2 a$. Per il potenziale Kepleriano $V(r) = -mk/r$ si ha $V'(a) = mk/a^2$ e dunque il legame tra a e ω all'equilibrio è

$$\omega^2 a^3 = k . \quad (\text{E.1})$$

Non deve stupire l'aver trovato all'equilibrio $p^* \neq 0$: per una massa in un sistema rotante, come per una carica in un campo magnetico, il momento p non è proporzionale alla velocità \dot{q} , invece si ha $p = m\dot{q} + m\Omega q$. (Nulla vieta, all'occorrenza, di introdurre una traslazione nei momenti per riportare p^* a zero.)

- *Esercizio 20.*

Introduciamo le coordinate del baricentro $\mathcal{Q} = (\mathcal{Q}_1, \mathcal{Q}_2)$ e le coordinate relative $\varrho = (\varrho_1, \varrho_2)$ come è abituale nel problema a due corpi:

$$\mathcal{Q} = \frac{M\mathcal{Q} + m\varrho}{M + m}, \quad \varrho = q - \mathcal{Q}, \quad (\text{E.2})$$

e denotiamo $\mathcal{M} = M + m$. Le relazioni inverse alle (E.2) sono

$$\mathcal{Q} = \mathcal{Q} - \frac{m}{\mathcal{M}} \varrho, \quad q = \mathcal{Q} + \frac{M}{\mathcal{M}} \varrho;$$

si tratta con evidenza di una trasformazione puntuale lineare della forma

$$\begin{pmatrix} \mathcal{Q}_i \\ \varrho_i \end{pmatrix} = A \begin{pmatrix} \mathcal{Q}_i \\ \varrho_i \end{pmatrix}, \quad A = \begin{pmatrix} 1 & -\frac{m}{\mathcal{M}} \\ 1 & \frac{M}{\mathcal{M}} \end{pmatrix}, \quad i = 1, 2,$$

che si estende ai momenti con

$$\begin{pmatrix} P_i \\ p_i \end{pmatrix} = A^{-T} \begin{pmatrix} \mathcal{P}_i \\ \pi_i \end{pmatrix},$$

avendo denotato con \mathcal{P} e π i momenti coniugati rispettivamente a \mathcal{Q} e ϱ . Si trova subito

$$P = \frac{M}{\mathcal{M}} \mathcal{P} - \pi, \quad p = \frac{m}{\mathcal{M}} \mathcal{P} + \pi,$$

e allora la nuova hamiltoniana \tilde{H} , ottenuta per sostituzione in H , è

$$\tilde{H}(\mathcal{P}, \pi, \mathcal{Q}, \varrho) = \frac{\mathcal{P}^2}{2\mathcal{M}} + \frac{\pi^2}{2\mu} + V(r),$$

ove $\mu = \frac{Mm}{M+m}$ è la cosiddetta *massa ridotta* del sistema, e naturalmente $r = \|\varrho\|$. Il problema è così disaccoppiato (a questo servono le coordinate baricentriche) e il grado di libertà interno descritto dalle coordinate canoniche (π, ϱ) è governato dall'hamiltoniana a due gradi di libertà

$$\mathcal{H}(\pi, \varrho) = \frac{\pi^2}{2\mu} + V(r).$$

Con le nuove notazioni si ha, nel caso kepleriano,

$$V(r) = -\frac{GMm}{r} = -\frac{GM\mu}{r}.$$

Siamo così ricondotti al precedente problema 19, con $k = GM$. Si ha pertanto una configurazione di equilibrio con

$$\varrho^* = (a, 0),$$

e corrispondentemente (assumendo $\mathcal{Q} = 0$)

$$\tilde{Q}_1^* = -\frac{m}{M+m} a, \quad \tilde{q}_1^* = \frac{M}{M+m} a, \quad Q_2^* = q_2^* = 0, \quad (\text{E.3})$$

se vale la (E.1), ovvero se

$$\omega^2 a^3 = GM \quad (\text{E.4})$$

- *Esercizio 21.*

Se all'hamiltoniana studiata nel problema precedente, che qui denotiamo H_0 , aggiungiamo un terzo corpo di massa ε descritto da coordinate canoniche (p', q') , otteniamo (nel sistema fisso) l'hamiltoniana complessiva

$$H = H_0 + H'$$

ove

$$H' = \frac{p'^2}{2\varepsilon} - \frac{\varepsilon GM}{\|q' - Q\|} - \frac{\varepsilon Gm}{\|q' - q\|} .$$

Nell'approssimazione in cui la massa ε si trascura ai fini del moto di M e m (*problema ristretto*), H' diviene l'hamiltoniana del terzo corpo, nella quale Q e q non sono più coordinate ma funzioni assegnate, soluzioni particolari del problema a due corpi; si ha così, propriamente, non un problema hamiltoniano per il terzo corpo, ma una famiglia di problemi, uno per ciascuna soluzione del problema a due corpi. M e m sono detti i due *corpi primari*, o semplicemente i primari. Mentre per essi il moto è piano, non è detto che lo sia il moto del terzo corpo, che per generalità e senza spesa descriviamo con $p', q' \in \mathbb{R}^3$.

Il valore ε della terza massa è completamente irrilevante (come in tutti i problemi in cui agiscono soltanto forze proporzionali alle masse) e lo si può porre uguale a uno; formalmente ε si elimina con il riscaldamento canonico (non stretto)

$$p' = \varepsilon p'' , \quad q' = q'' , \quad H'' = \varepsilon^{-1} H' ,$$

che dà, sottolineando il fatto che Q e q sono ora funzioni assegnate del tempo,

$$H'' = \frac{p''^2}{2} - \frac{GM}{\|q'' - Q(t)\|} - \frac{Gm}{\|q'' - q(t)\|} .$$

Se prendiamo per il problema a due corpi la soluzione circolare (E.3) studiata nel problema precedente 19 (*problema ristretto circolare*), diventa conveniente passare al sistema corotante con velocità ω data dalla (E.4), nel quale i primari appaiono fermi in

$$Q^* = a(-\eta, 0, 0) , \quad q^* = a(1 - \eta, 0, 0) , \quad \text{ove } \eta = m/M ;$$

si ottiene

$$K(\tilde{p}, \tilde{q}) = \frac{\tilde{p}^2}{2} - \omega(\tilde{q}_1 \tilde{p}_2 - \tilde{q}_2 \tilde{p}_1) - \frac{GM}{\|\tilde{q} - Q^*\|} - \frac{Gm}{\|\tilde{q} - q^*\|} .$$

Le costanti G, \mathcal{M} , ω e a non sono però indipendenti ma sono legate dalla (E.4); sostituendo ad esempio $G = \omega^2 a^3 / \mathcal{M}$ si trova

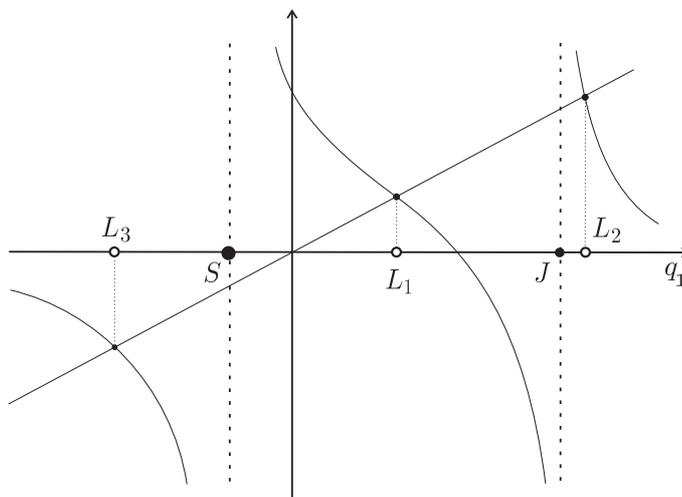
$$K(\tilde{p}, \tilde{q}) = \frac{\tilde{p}^2}{2} - \omega(\tilde{q}_1 \tilde{p}_2 - \tilde{q}_2 \tilde{p}_1) - \omega^2 a^3 \frac{1 - \eta}{\|\tilde{q} - Q^*\|} - \omega^2 a^3 \frac{\eta}{\|\tilde{q} - q^*\|} .$$

Nell'hamiltoniana compaiono ora la costante adimensionale η e due costanti dimensionali a e ω , che in quanto tali possiamo eliminare prendendole come unità di misura (si passa così a una formulazione completamente adimensionale del problema).⁵⁹ Riprendendo la notazione piana senza orpelli H , p e q si ottiene in questo modo l'hamiltoniana

$$H(p, q) = \frac{p^2}{2} - (q_1 p_2 - q_2 p_1) - \frac{1 - \eta}{R} - \frac{\eta}{r} \tag{E.5}$$

⁵⁹Formalmente si effettua innanzitutto il riscaldamento canonico (non stretto)

$$\tilde{q} = a\hat{q} , \quad \tilde{p} = \omega a\hat{p} , \quad \hat{K} = \frac{1}{\omega a^2} K ,$$

Figura 38: La costruzione grafica per determinare L_1, L_2, L_3 .

ove si è posto per brevità

$$R = \|q - Q^*\| = \sqrt{(q_1 - \eta)^2 + q_2^2 + q_3^2}, \quad r = \|q - \hat{q}^*\| = \sqrt{(q_1 - 1 + \eta)^2 + q_2^2 + q_3^2}.$$

Per la ricerca dei punti di equilibrio dobbiamo annullare tutte le derivate parziali di H . Derivando rispetto a p_3 e q_3 si vede immediatamente che (come è del tutto ovvio) deve essere

$$p_3 = 0, \quad q_3 = 0,$$

mentre per p_1, p_2, q_1, q_2 c'è un po' da lavorare. Derivando rispetto a p_1 e p_2 si trova

$$p_1 + q_2 = 0, \quad p_2 - q_1 = 0; \quad (\text{E.6})$$

derivando poi rispetto a q_1 e q_2 , e sostituendo già p_1 e p_2 con le (E.6), si trova poi

$$-q_1 + \frac{(1-\eta)(q_1+\eta)}{R^{3/2}} + \frac{\eta(q_1-1+\eta)}{r^{3/2}} = 0, \quad -q_2 + \frac{(1-\eta)q_2}{R^{3/2}} + \frac{\eta q_2}{r^{3/2}} = 0. \quad (\text{E.7})$$

La seconda di queste è risolta innanzitutto da $q_2 = 0$; la prima diviene allora un'equazione per q_1 , precisamente

$$q_1 - \left[\frac{1-\eta}{(q_1+\eta)^2} \frac{q_1+\eta}{|q_1+\eta|} + \frac{\eta}{(q_1-1+\eta)^2} \frac{q_1-1+\eta}{|q_1-1+\eta|} \right] = 0, \quad (\text{E.8})$$

che non si risolve esattamente, tuttavia in modo grafico si vede bene che ci sono tre soluzioni reali disposte come in figura 3, punti L_1, L_2 e L_3 . Infatti la quantità entro parentesi nella (E.8) dà qualitativamente la curva riportata in figura 38 (sono tre tratti distinti per la presenza dei moduli a denominatore: la curva diverge in $-\eta$ e $1-\eta$ quando i denominatori si annullano, con divergenza opposta a sinistra e a destra), e l'intersezione con la retta q_1 fornisce esattamente tre soluzioni reali disposte come in figura 3. L_1, L_2, L_3 sono detti punti lagrangiani "collineari".

che dà

$$\hat{K}(\tilde{p}, \tilde{q}) = \omega \left(\frac{\tilde{p}^2}{2} - (\hat{q}_1 \hat{p}_2 - \hat{q}_2 \hat{p}_1) - \frac{1-\eta}{\|\hat{q} - \hat{Q}^*\|} - \frac{\eta}{\|\hat{q} - \hat{q}^*\|} \right) \quad \text{ove} \quad \hat{Q}^* = (-\eta, 0, 0), \quad \hat{q}^* = (1-\eta, 0, 0);$$

poi si passa al tempo $\tilde{t} = \omega t$, così da rimuovere la costante ω a fattore dell'hamiltoniana.

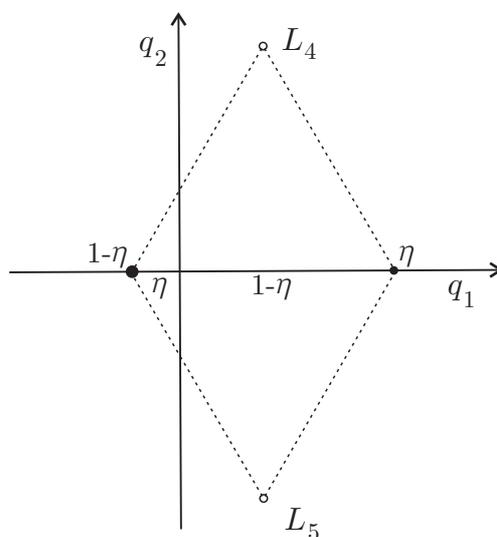


Figura 39: I punti lagrangiani triangolari L_4 e L_5 , nel sistema corotante.

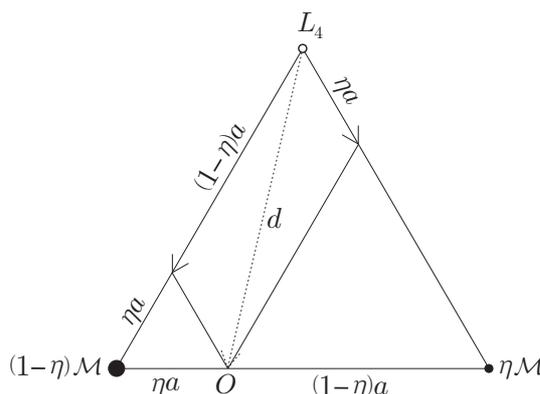


Figura 40: La costruzione geometrica per dimostrare in modo elementare l'esistenza di L_4, L_5 .

Due altre soluzioni delle (E.7), simmetriche in q_2 , si ottengono per $q_2 \neq 0$. Qui non occorrono calcoli, bisogna solo accorgersi che tutto si sbrogia se si prende $R = r$: la seconda equazione infatti è soddisfatta a vista, e anche la prima si risolve subito precisando ulteriormente $R = r = 1$. *Corrispondentemente si ottengono due punti di equilibrio L_4 e L_5 da parti opposte dell'asse q_1 , tali da formare con i due primari due triangoli equilateri*; si veda la figura 39. L_4 e L_5 sono detti punti lagrangiani “triangolari”.

Vale la pena di osservare che, al solo fine di determinare le posizioni di equilibrio del terzo corpo nel sistema rotante, tutto l'apparato della meccanica hamiltoniana (o delle meccanica lagrangiana) non è realmente necessario: *le equazioni per l'equilibrio sono esattamente quelle che avremmo scritto direttamente sulla base della meccanica newtoniana nel sistema rotante*, imponendo che attrazione gravitazionale e forza centrifuga si facciano equilibrio (la forza di Coriolis, nulla per velocità nulle, è sempre irrilevante per la ricerca dei punti di equilibrio).⁶⁰ La formulazione hamiltoniana diventa invece *preziosa* per lo studio della stabilità dell'equilibrio, che è risultato particolarmente difficile per

⁶⁰Per una dimostrazione del tutto elementare dell'esistenza dei punti di equilibrio L_4 e L_5 basta osservare la figura 40. Le due forze gravitazionali sono dirette lungo i lati del triangolo e sono proporzionali alle masse, precisamente

L_4 e L_5 (nonostante significativi progressi attorno al 1998–2000, lo studio non è ancora concluso).

• *Esercizio 22.* L'esercizio è un po' lungo e complesso, perciò ci limitiamo a proporne una traccia; come punto di partenza prendiamo l'esercizio precedente, in particolare l'hamiltoniana (E.5).

i) Si trasla l'origine dello spazio delle fasi nel punto di equilibrio

$$L_4 = (p_1^*, p_2^*, p_3^*, q_1^*, q_2^*, q_3^*) = (-q_2^*, q_1^*, 0, q_1^*, q_2^*, 0) ,$$

ove

$$q_1^* = \frac{1}{2} - \eta , \quad q_2^* = \frac{\sqrt{3}}{2} ,$$

e si effettua uno sviluppo di Taylor attorno a questo punto, troncato al secondo ordine. Posto $\tilde{p} = p - p^*$, $\tilde{q} = q - q^*$, il risultato si trova essere

$$\tilde{H}(\tilde{p}, \tilde{q}) = \tilde{H}_2(\tilde{p}, \tilde{q}) + \mathcal{O}(\tilde{q}^3) ,$$

con (abolendo ovunque per semplicità la tilde)

$$\begin{aligned} H_2(p, q) &= \mathcal{H}(p_1, p_2, q_1, q_2) + \frac{1}{2}(p_3^2 + q_3^2) \\ \mathcal{H}(p_1, p_2, q_1, q_2) &= \frac{1}{2}(p_1^2 + p_2^2) - q_1 p_2 + q_2 p_1 + \frac{1}{2} \left(\frac{1}{4} q_1^2 - \frac{3\sqrt{3}}{2} (1 - 2\eta) q_1 q_2 - \frac{5}{4} q_2^2 \right) . \end{aligned}$$

ii) Si studia l'hamiltoniana troncata H_2 (cioè il sistema linearizzato). Il grado di libertà (p_3, q_3) è disaccoppiato: rappresenta un oscillatore armonico di pulsazione unitaria, cioè con periodo uguale al periodo di rotazione dei primari. Gli altri due gradi di libertà sono retti da \mathcal{H} , e posto $x = (p_1, p_2, q_1, q_2)$ le equazioni del moto, lineari, hanno la forma

$$\dot{x} = Ax , \quad A = \begin{pmatrix} 0 & 1 & a & b \\ -1 & 0 & b & d \\ 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix} ,$$

ove

$$a = -\frac{1}{4} , \quad b = \frac{3\sqrt{3}}{4} (1 - 2\eta) , \quad d = \frac{5}{4} .$$

L'analisi dello spettro di A ci fornisce la stabilità del problema linearizzato. L'equazione agli autovalori risulta biquadratica, precisamente, a conti fatti,

$$\lambda^4 + \lambda^2 + c = 0 , \quad c = \frac{27}{4} (\eta - \eta^2) > 0 .$$

valgono

$$F_M = (1 - \eta) \frac{GM}{a^2} , \quad F_m = \eta \frac{GM}{a^2} ;$$

sono perciò proporzionali alle frecce di lunghezza $(1 - \eta)a$ e ηa tracciate in figura, con costante di proporzionalità GM/a^3 . La risultante allora, per la regola del parallelogramma, è rivolta verso l'origine e ha intensità F proporzionale, secondo la medesima costante, alla distanza d di L_4 dall'origine,

$$F = d \frac{GM}{a^3} .$$

A sua volta la forza centrifuga ha la stessa direzione, verso opposto e intensità $\omega^2 d$. Poiché $\omega^2 = GM/a^3$, risultante e forza centrifuga si compensano esattamente e dunque L_4 è punto di equilibrio.

Per ogni soluzione λ c'è dunque anche $-\lambda$, e naturalmente per la realtà dell'equazione c'è anche il coniugato $\bar{\lambda}$. Possiamo così avere, a priori, coppie di autovalori reali opposti, coppie di autovalori immaginari puri opposti, oppure un quartetto complesso del tipo $\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}$.⁶¹ Nel caso di due coppie di autovalori immaginari puri, il problema linearizzato è stabile (il punto si dice *ellittico*). Negli altri due casi, essendovi almeno un autovalore con parte reale negativa, il punto di equilibrio è instabile (punto *iperbolico* se tutti e quattro gli autovalori hanno parte reale diversa da zero, parzialmente iperbolico se c'è una coppia immaginaria coniugata e una coppia di reali opposti).

Per $z = \lambda^2$ si trovano le soluzioni

$$z_{\pm} = \frac{1}{2}(-1 \pm \sqrt{1 - 27(\eta - \eta^2)}) ,$$

e dobbiamo distinguere due casi:

- Se $1 - 27(\eta - \eta^2) = \alpha^2 \geq 0$, allora gli z_{\pm} sono entrambi reali negativi. Corrispondentemente gli autovalori di A sono tutti immaginari puri, precisamente

$$\lambda = \pm i \sqrt{\frac{1 \pm \alpha}{2}}$$

(con scelta indipendente dei segni), e il punto di equilibrio è ellittico.

- Se invece $1 - 27(\eta - \eta^2) = -\alpha^2 < 0$, allora si ha

$$z_{\pm} = \frac{1}{2}(-1 \pm i\alpha) , \quad \alpha \neq 0 ;$$

le quattro radici quadrate, ovvero i quattro valori di λ , hanno parti reali a due a due opposte, siamo nella situazione di quartetto, il punto di equilibrio è iperbolico. (E' parzialmente iperbolico, se si tiene conto del terzo grado di libertà.)

Risolvendo la disequazione si trova che il punto di equilibrio è ellittico per $\eta \leq \eta_R$ oppure $(1 - \eta) \leq \eta_R$, ove

$$\eta_R = \frac{1}{2} \left(1 - \sqrt{\frac{23}{27}} \right) \simeq 0.0385 ;$$

ci si restringe comunemente a $\eta \leq \eta_R$, supponendo come non restrittivo $m \leq M$ e dunque $\eta \leq \frac{1}{2}$. La soglia η_R è detta *limite di Rout*. Per il problema Sole–Giove si ha $\eta \simeq 10^{-3}$, per il sistema Terra–Luna $\eta \simeq 0.012$.

⁶¹E' obbligatoriamente così quando si linearizza un sistema hamiltoniano attorno a un punto di equilibrio, per un numero qualunque di gradi di libertà. Infatti l'hamiltoniana troncata ha obbligatoriamente la forma $H_2 = \frac{1}{2}x \cdot Bx$, con B hessiana di H calcolata nel punto di equilibrio (i termini lineari dello sviluppo non ci sono, trattandosi di un punto di equilibrio). Le equazioni di Hamilton sono allora della forma $\dot{x} = \mathbb{E}\partial_x H = Ax$, con $A = \mathbb{E}B$. Dall'espressione di A segue subito che il polinomio caratteristico $P(\lambda) = \det(A - \lambda I)$ è pari: infatti

$$P(\lambda) = \det(\mathbb{E}B - \lambda I) = \det(\mathbb{E}B - \lambda I)^T = \det(-B\mathbb{E} - \lambda I) = \det \mathbb{E}(-B\mathbb{E} - \lambda I)\mathbb{E} = \det(\mathbb{E}B + \lambda I) = P(-\lambda) .$$

Del tutto in generale dunque gli autovalori di A compaiono in coppie reali opposte, coppie immaginarie pure opposte, quartetti con parti reali e immaginarie indipendentemente opposte.

- iii) Per $\eta < \eta_R$, con una trasformazione canonica lineare si passa alle coordinate normali, che denotiamo (riprendendo anche il terzo grado di libertà) P_1, P_2, \dots, Q_3 . Il sistema appare disaccoppiato in tre oscillatori armonici, ma l'hamiltoniana ha la forma

$$K(P, Q) = \frac{\omega_1}{2}(Q_1^2 + P_1^2) - \frac{\omega_2}{2}(Q_2^2 + P_2^2) + \frac{\omega_3}{2}(Q_3^2 + P_3^2),$$

con

$$\omega_1 = \sqrt{\frac{1+\alpha}{2}}, \quad \omega_2 = \sqrt{\frac{1-\alpha}{2}}, \quad \omega_3 = 1.$$

La presenza del segno ‘-’ nel secondo termine di K , sostanzialmente irrilevante se ci si limita a considerare il problema linearizzato, è di importanza essenziale si considera il problema completo, cioè se si aggiungono i termini $\mathcal{O}(q^3)$ precedentemente trascurati. A causa di quel segno meno, infatti, l'origine *non* è un punto di minimo di K , ma un punto di sella, e *non* si può concludere, come si farebbe se il segno fosse positivo, che l'aggiunta di termini cubici non può pregiudicare la stabilità. Il problema della stabilità di L_4 e L_5 (si veda il testo dell'esercizio 23) si è rivelato molto difficile, tanto che a tutt'oggi, oltre due secoli dopo la sua formulazione, non è ancora del tutto risolto. Di fatto, attorno ai punti lagrangiani del sistema Sole–Giove, vi sono molti asteroidi (i cosiddetti “Greci” e “Troiani”, con i nomi degli eroi omerici), la cui presenza ancora non è completamente spiegata.

- *Esercizio 29.*

Consideriamo la traslazione sul circolo $S_{2\pi}^1 = \mathbb{R}/2\pi\mathbb{Z}$ definita da

$$\varphi \mapsto \Phi(\varphi) = \varphi + 2\pi\alpha \quad \text{mod } 2\pi, \quad 0 \leq \alpha < 1.$$

Il cuore dell'esercizio 29 è la dimostrazione che

PROPOSIZIONE 38 Per un qualunque $\varphi \in S_{2\pi}^1$ l'orbita $\Phi^t(\varphi)$, $t \in \mathbb{Z}$, è densa su $S_{2\pi}^1$ se e solo se α è irrazionale.

DIMOSTRAZIONE. Se α è razionale, $\alpha = k/l$, allora per ogni $\varphi \in S_{2\pi}^1$ si ha $\Phi^l(\varphi) = \varphi$. Pertanto l'orbita è periodica e non può essere densa. Viceversa è anche evidente che se un'orbita è periodica, allora α è necessariamente razionale (se è periodica di periodo l , allora necessariamente $\alpha = k/l$ con $k \in \mathbb{Z}$). Preso allora α irrazionale, i punti

$$\varphi, \quad \Phi(\varphi), \quad \dots, \quad \Phi^k(\varphi), \quad \dots$$

sono necessariamente distinti. Dire che l'orbita è densa vuol dire che preso ad arbitrio un intervallo di lunghezza ε sul circolo, esiste k tale che $\Phi^k(\varphi)$ vi appartiene. Si può ragionare così: per k abbastanza grande (basta $k > 2\pi/\varepsilon$) vi sono certamente almeno due punti $\Phi^{k'}(\varphi)$, $\Phi^{k''}(\varphi)$ a distanza più piccola di ε :

$$\text{dist}(\Phi^{k''}(\varphi), \Phi^{k'}(\varphi)) < \varepsilon.$$

Ma Φ è una traslazione rigida su $S_{2\pi}^1$ (Φ preserva la distanza), perciò tornando indietro di k' passi, e ponendo $l = k'' - k'$, si trova

$$\text{dist}(\Phi^l(\varphi), \varphi) < \varepsilon.$$

Ma allora $\varphi, \Phi^l(\varphi), \Phi^{2l}(\varphi), \dots$, procedono a passi inferiori a ε , dunque entrano in qualunque intervallo di lunghezza ε . □

Si può dimostrare in realtà qualche cosa di più, precisamente che per α irrazionale ogni orbita si distribuisce in modo uniforme sul circolo: precisamente, preso un qualunque intervallo A di ampiezza $|A|$ (più in generale: preso un qualunque sottoinsieme misurabile A di $S_{2\pi}^1$), per ogni φ si ha che

$$\lim_{t \rightarrow \infty} \frac{1}{t} N_A(t, \varphi) = \frac{|A|}{2\pi}$$

(più in generale: $= \text{mes}(A)/2\pi$), ove $N_A(t, \varphi)$ è il numero di punti $\Phi^s(\varphi)$ che cadono in A per $0 \leq s < t$.

• *Esercizio 31.*

Per $\omega_1 = \omega_2$ l'hamiltoniana è

$$H = \omega(I_1 + I_2) ;$$

una trasformazione canonica con la proprietà richiesta è, ad esempio,

$$I_1 = J_1 - J_2, \quad I_2 = J_2, \quad \varphi_1 = \psi_1, \quad \varphi_2 = \psi_1 + \psi_2 .$$

Nel caso di ω_2/ω_1 razionale, esistono $k_1, k_2 \in \mathbb{Z}$ primi tra loro tali che $k_1\omega_1 + k_2\omega_2 = 0$. Dimostriamo allora che

PROPOSIZIONE 39 *Se $k_1, k_2 \in \mathbb{Z}$ sono primi tra loro, esistono sempre $l_1, l_2 \in \mathbb{Z}$ tali che la matrice*

$$\mathcal{A} = \begin{pmatrix} l_1 & l_2 \\ k_1 & k_2 \end{pmatrix}$$

abbia determinante uno.

Ciò risolve l'esercizio perché allora la trasformazione canonica (3.12), con $A = \mathcal{A}^{-1}$, dà, come è facile verificare,

$$I_1 = l_1 J_1 + k_1 J_2, \quad I_2 = l_2 J_1 + k_2 J_2, \quad H = \omega_1 I_1 + \omega_2 I_2 = (l \cdot \omega) J_1 + (k \cdot \omega) J_2 = (l \cdot \omega) J_1 .$$

DIMOSTRAZIONE. Per ogni scelta di l_1 , sia l_2 tale che

$$0 \leq \det \mathcal{A} = l_1 k_2 - l_2 k_1 < k_1$$

($l_2 =$ quoziente della divisione intera $l_1 k_2 / k_1$; il determinante è il resto). Si vede facilmente che al variare di l_1 tra 0 e $k_1 - 1$, i valori che si ottengono per il determinante $l_1 k_2 - l_2 k_1$ sono tutti diversi. Infatti, se l_1 e l'_1 fossero tali che

$$l_1 k_2 - l_2 k_1 = l'_1 k_2 - l'_2 k_1 ,$$

risulterebbe

$$\frac{k_2}{k_1} = \frac{l_2 - l'_2}{l_1 - l'_1} ,$$

contro l'ipotesi che k_1, k_2 siano primi tra loro. Tra i k_1 valori diversi compresi tra 0 e $k_1 - 1$ c'è sicuramente il numero 1. \square

La proprietà messa in evidenza nell'esercizio ha una naturale continuazione per sistemi della forma $H(I, \varphi) = \omega \cdot I$ in dimensione qualsiasi. Precisamente, si dice che l' n -pla delle frequenze

$\omega = (\omega_1, \dots, \omega_n)$ soddisfa r relazioni di risonanza, $r = 1, \dots, n-1$, se esistono r vettori linearmente indipendenti $k^{(1)}, \dots, k^{(r)} \in \mathbb{Z}^n$ tali che

$$k^{(s)} \cdot \omega = 0, \quad s = 1, \dots, r.$$

Si dimostra che se ω soddisfa r relazioni risonanza, allora esiste una trasformazione canonica del tipo (3.12) tale che la nuova hamiltoniana ha la forma

$$K = \Omega_1 J_1 + \dots + \Omega_{n-r} J_{n-r}.$$

Il sistema è così r volte degenere e il moto è confinato a un toro di dimensione $n-r$ (e lo invade densamente, in assenza di ulteriori relazioni di risonanza).

• *Esercizio 32.*

La trasformazione (3.14) che cerchiamo è puntuale nei momenti, estesa canonicamente alle coordinate. Alla luce dell’esercizio 13, con coordinate e momenti scambiati, si ha necessariamente $\varphi = B^{-T}(J)[\psi + a(J)]$, con B jacobiana di u e a tale che $\frac{\partial a_i}{\partial J_j} = \frac{\partial a_j}{\partial J_i}$; denotando $A = B^{-T}$ si riscrive

$$\varphi = A(J)[\psi + a(J)].$$

Ma come insegna l’esercizio 28, per ogni fissato J , se si vuole avere una mappa regolare biunivoca su \mathbb{T}^2 , $A(J)$ deve essere intera con $|\det A(J)| = 1$. Obbligatoriamente allora A è costante.

F Il teorema di Liouville–Arnol’d, punti i. e ii.

La condizione (3.11) garantisce innanzitutto che Σ_{f^*} stessa è una varietà n -dimensionale contenuta in U . Per l’ipotesi (3.10), ciascuna delle F_1, \dots, F_n è conservata da ciascun sistema di hamiltoniana F_i , $i = 1, \dots, n$, e questo basta per il primo punto.

Denotiamo ora con

$$X_i = \left(-\frac{\partial F_i}{\partial q}, \frac{\partial F_i}{\partial p} \right)$$

il campo vettoriale associato all’hamiltoniana F_i e con Φ_i^t il corrispondente flusso; per l’ipotesi (3.11) gli n campi vettoriali X_1, \dots, X_n sono linearmente indipendenti in Σ_{f^*} , inoltre essi sono tangenti a Σ_{f^*} , come è evidente per il fatto che ciascuno dei flussi Φ_i si svolge su Σ_{f^*} ; per la (3.10) poi i flussi Φ_1, \dots, Φ_n (appendice C) commutano a due a due.

Il punto *ii.* del teorema allora si riporta interamente al seguente lemma, ove in luogo di Σ_{f^*} si discute di una generica varietà n -dimensionale Σ :

LEMMA 40 *Se su una varietà n -dimensionale connessa Σ esistono n campi vettoriali tangenti X_1, \dots, X_n linearmente indipendenti in ogni suo punto, i cui flussi $\Phi_1^t, \dots, \Phi_n^t$ sono prolungabili per $t \in \mathbb{R}$ e commutano a due a due, allora esiste k intero, $0 \leq k \leq n$, tale che Σ stessa è diffeomorfa al cilindro*

$$\mathcal{C}^{n,k} = \mathbb{T}^k \times \mathbb{R}^{n-k}.$$

DIMOSTRAZIONE. Preso $\tau = (\tau_1, \dots, \tau_n) \in \mathbb{R}^n$ qualsiasi, sia $\Phi^\tau : \Sigma \rightarrow \Sigma$ la mappa definita da

$$\Phi^\tau = \Phi_n^{\tau_n} \circ \dots \circ \Phi_1^{\tau_1};$$

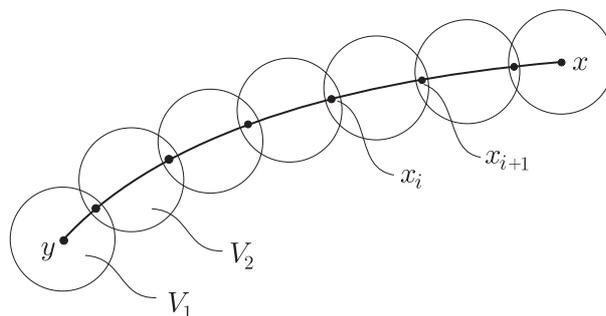


Figura 41: Gli aperti V_1, \dots, V_l ricoprono γ .

per la supposta commutatività dei flussi l'ordine in cui essi appaiono non è rilevante, e se ne deduce immediatamente che

$$\Phi^0 = \text{identità}, \quad (\Phi^\tau)^{-1} = \Phi^{-\tau}, \quad \Phi^{\tau+\tau'} = \Phi^\tau \circ \Phi^{\tau'},$$

ovvero $\{\Phi^\tau, \tau \in \mathbb{R}^n\}$ è un gruppo commutativo a n parametri di diffeomorfismi: $\Sigma \rightarrow \Sigma$; si osservi che questa è una naturale generalizzazione della nozione di flusso.

Fissiamo ora un qualsiasi punto y (dato iniziale, origine) in Σ , e definiamo $\Psi_y : \mathbb{R}^n \rightarrow \Sigma$ ponendo

$$\Psi_y(\tau) = \Phi^\tau(y)$$

(in altre parole: ora pensiamo il punto y fisso, e τ variabile; ma è solo un cambiamento di notazione). La mappa Ψ_y ha alcune proprietà rilevanti, di facile dimostrazione:

- a) Ψ_y è un diffeomorfismo locale, che in particolare mappa un conveniente intorno U_y di $\tau = 0$ in un conveniente intorno V_y di y . Che sia diffeomorfismo locale segue dal fatto che per ogni τ si ha, come è immediato verificare,

$$\frac{\partial \Psi_y}{\partial \tau_j}(\tau) = X_j(\Psi_y(\tau)), \quad \frac{\partial \Psi_y}{\partial \tau_j}(0) = X_j(y),$$

e a loro volta i campi vettoriali si sono supposti ovunque indipendenti su Σ . Un'ulteriore ovvia conseguenza è la seguente: presi ad arbitrio $x', x'' \in V_y$, esiste $\tau \in \mathbb{R}^n$ tale che $\Phi^\tau(x') = x''$ (infatti se $x' = \Phi^{\tau'}(y)$ e $x'' = \Phi^{\tau''}(y)$, allora $x'' = \Phi^{\tau''-\tau'}(x')$).

- b) L'applicazione $\Psi_y : \mathbb{R}^n \rightarrow \Sigma$ è suriettiva. Infatti, preso $x \in \Sigma$, sia γ una curva congiungente y con x . Per ogni $y' \in \gamma$ esiste l'aperto $V_{y'}$ cui si è fatto riferimento sopra, e per la compattezza di γ , ne basta un numero finito V_1, \dots, V_l a ricoprire γ stessa. Si costruisce allora facilmente (si veda la figura 41) una catena finita

$$y = x_0, x_1, \dots, x_l = x,$$

con x_{j-1} e x_j nello stesso aperto V_j ; per il punto precedente si trovano allora τ^1, \dots, τ^l , tali che $x_j = \Phi^{\tau^j}(x_{j-1})$, e dunque, posto $\tau = \tau^1 + \dots + \tau^l$, si ha $\Phi^\tau(y) = x$, ovvero $x = \Psi_y(\tau)$.

- c) Se Ψ_y non è biettiva (altrimenti il lemma è già dimostrato, con Σ diffeomorfa a \mathbb{R}^n ovvero con $k = 0$), allora l'insieme $\mathcal{T} \subset \mathbb{R}^n$ dei periodi di y , ovvero l'insieme

$$\mathcal{T} = \{\tau \in \mathbb{R}^n : \Phi^\tau(y) = y\},$$

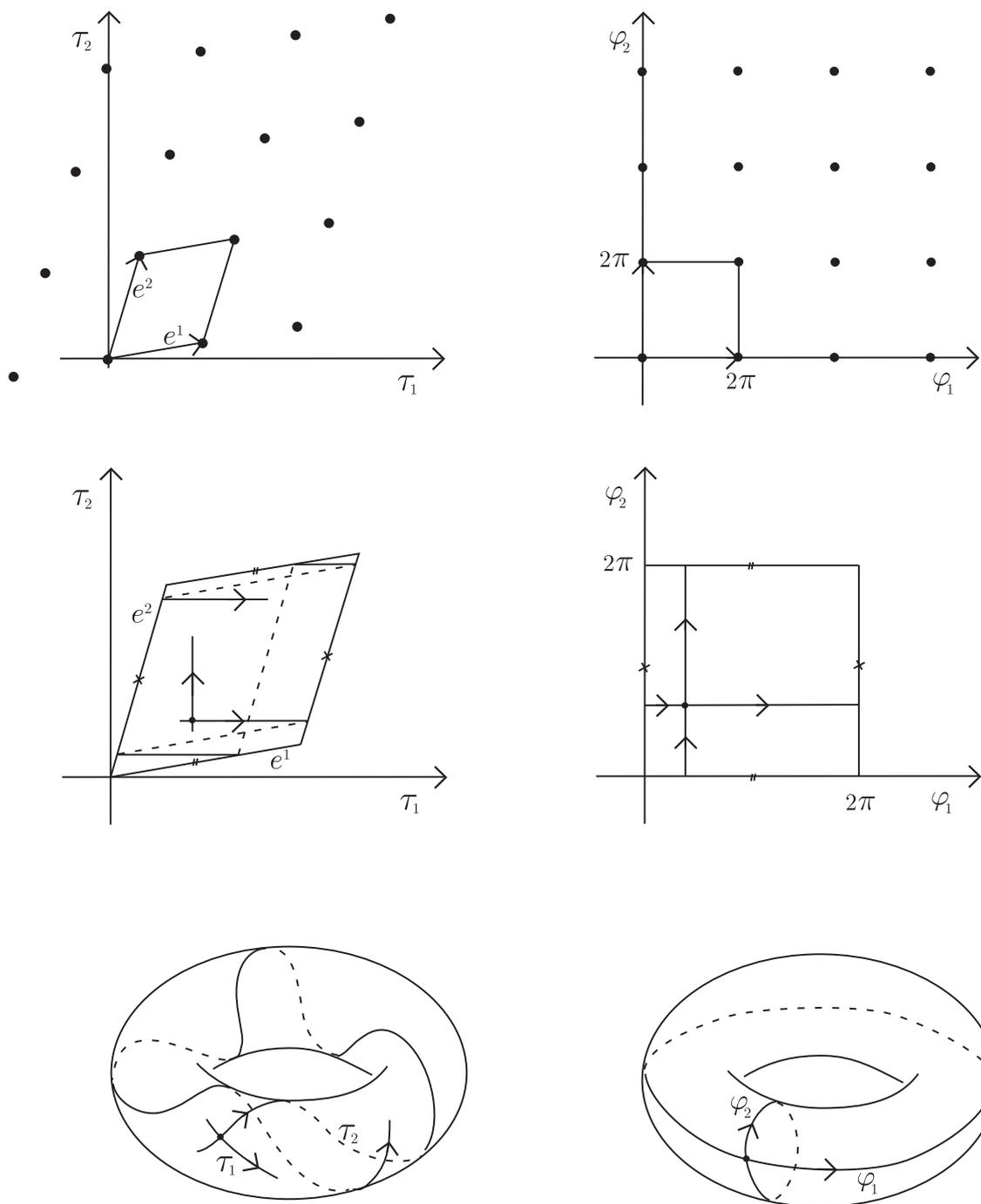


Figura 42: *Le variabili τ e le variabili φ .*

è un sottogruppo discreto di \mathbb{R}^n , che non dipende da y . Che l'insieme dei periodi sia un gruppo è del tutto evidente. Che sia discreto segue immediatamente dal fatto che Ψ_y è un diffeomorfismo locale (vicino a $\tau = 0$ non possono esservi altri periodi; ma se il gruppo dei periodi avesse un'accumulazione, allora si troverebbero elementi τ', τ'' arbitrariamente vicini, e la loro differenza $\tau = \tau' - \tau''$ sarebbe un periodo vicino a zero). Per dimostrare che il gruppo non dipende da y , facciamo vedere che presi y e x qualsiasi su Σ , se τ è periodo di y allora è anche periodo di x (e di conseguenza anche viceversa). Infatti, se $x = \Phi^\sigma(y)$, allora

$$\Phi^\tau(x) = \Phi^\tau(\Phi^\sigma(y)) = \Phi^\sigma(\Phi^\tau(y)) = \Phi^\sigma(y) = x .$$

Identifichiamo ora valori diversi di τ che differiscano per un periodo, ovvero prendiamo il quoziente $\tilde{\Sigma} = \mathbb{R}^n/\mathcal{T}$. Come mostra il diagramma, resta allora definita l'applicazione quoziente $\tilde{\Psi}_y = \Psi_y/\mathcal{T} : \tilde{\Sigma} \rightarrow \Sigma$, che per quanto sopra detto risulta essere un diffeomorfismo.

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\Psi_y} & \Sigma \\ \downarrow/\mathcal{T} & \nearrow \tilde{\Psi}_y \text{ (difeo)} & \\ \tilde{\Sigma} & & \end{array}$$

Studiamo allora $\tilde{\Sigma}$. Per questo ci viene in aiuto un lemma di natura algebrica, la cui dimostrazione, poco interessante, è qui omessa:

LEMMA 41 *Se \mathcal{T} è un sottogruppo discreto di \mathbb{R}^n , allora esistono $e^1, \dots, e^k \in \mathcal{T}$ linearmente indipendenti, $k \leq n$, tali che \mathcal{T} è costituito da tutte e sole le loro combinazioni lineari intere:*

$$\mathcal{T} = \left\{ \sum_{i=1}^k m_i e^i, m_1, \dots, m_k \in \mathbb{Z} \right\} .$$

AmMESSO questo lemma, la conclusione del lemma 40 segue facilmente: se $k < n$, prendiamo (a piacere) e^{k+1}, \dots, e^n in modo da ottenere una base e^1, \dots, e^n , ed eseguiamo il cambiamento di coordinate lineare invertibile dalle variabili $\tau \in \mathbb{R}^n$ a nuove variabili $\varphi \in \mathbb{R}^n$, definito da

$$\tau = \frac{1}{2\pi} \sum_{j=1}^n \varphi_j e^j ,$$

ovvero poniamo

$$\tau = \frac{1}{2\pi} T \varphi ,$$

ove T è la matrice $n \times n$ di elementi $T_{ij} = (e^j)_i$. Si ottengono evidentemente i periodi quando $\varphi_1, \dots, \varphi_k$ sono multipli interi di 2π , mentre $\varphi_{k+1}, \dots, \varphi_n$ sono nulli; l'elemento di matrice T_{ij} ci dice precisamente di quanto avanza τ_i , quando φ_j , $j \leq k$, avanza di 2π . La matrice T è detta *matrice dei periodi*. Il gruppo dei periodi, nelle nuove variabili, è $(2\pi\mathbb{Z})^k$.

E' chiaro che l'operazione di quoziente $\mathbb{R}^n/(2\pi\mathbb{Z})^k$, che per definizione dà il cilindro standard $\mathcal{C}^{n,k}$, corrisponde all'operazione di quoziente \mathbb{R}^n/\mathcal{T} , che dà $\tilde{\Sigma}$. Resta allora ben definito il diffeomorfismo \tilde{T} che appare nel diagramma

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{\frac{1}{2\pi}T} & \mathbb{R}^n & \xrightarrow{\Psi_y} & \Sigma \\ \downarrow/(2\pi\mathbb{Z})^k & & \downarrow/\mathcal{T} & \nearrow \tilde{\Psi}_y & \\ \mathcal{C}^{n,k} & \xrightarrow{\tilde{T}} & \tilde{\Sigma} & & \end{array}$$

e con esso il diffeomorfismo $\tilde{\Psi}_y \circ \tilde{T}$ che, come preteso nel lemma 40, manda il cilindro standard $\mathbb{C}^{n,k}$ in Σ . La figura 42 illustra la situazione per $k = n = 2$. \square

Se Σ è compatta si ha evidentemente $k = n$, e il cilindro altro non è che il toro \mathbb{T}^n ; questo è il caso preso in considerazione nell'enunciato 16 del teorema di Liouville–Arnold. Ciò conclude la dimostrazione del punto ii. del teorema. Per una dimostrazione completa (comprendente anche la dimostrazione del lemma 41, qui omessa) si veda ad esempio G. Benettin *Sistemi Dinamici*, capitolo III (www.math.unipd.it/~benettin).

G Nozioni di base sulla dinamica del corpo rigido

Ci limiteremo qui a corpi rigidi composti da un numero finito di punti materiali P_1, \dots, P_N ; tutto si estende in realtà al caso di un corpo continuo, ma non ce ne occuperemo.

G.1 Cinematica

DEFINIZIONE 18 *Si dice che gli $N \geq 2$ punti materiali P_1, \dots, P_N sono soggetti al vincolo di rigidità, ovvero che costituiscono un corpo rigido, se le mutue distanze $d_{ij} = \|P_i - P_j\|$ sono fissate.*

Il corpo rigido più semplice è dunque costituito da due punti materiali vincolati a mantenere distanza fissata. Ci restringeremo tuttavia nel seguito al caso interessante di corpi rigidi costituiti da almeno tre punti materiali non allineati.

A ogni corpo rigido si può sempre associare un sistema di riferimento $Oe_1e_2e_3$ ortonormale (destrorso) *fisso nel corpo*, o *solidale*, cioè tale che in esso le coordinate dei punti P_1, \dots, P_N , che denoteremo $x_k^{(i)}$, $i = 1, \dots, N$, $k = 1, 2, 3$, si mantengano costanti.⁶² Un vettore u si dice *solidale*, se le sue componenti $u_k = u \cdot e_k$, $k = 1, 2, 3$, nel sistema solidale si mantengono costanti; sono con evidenza solidali i vettori $P_j - O$, o $P_i - P_j$, $i, j = 1, \dots, N$.

Sia invece $O'e_xe_ye_z$ un sistema di riferimento ortonormale (destrorso) *fisso nello spazio*, o brevemente *fisso*, diciamo il sistema inerziale rispetto al quale vogliamo descrivere il moto del corpo.

Il caso cui dedicheremo maggiore attenzione è quello di un corpo rigido con un punto fisso.

DEFINIZIONE 19 *Diremo che il corpo rigido ha un punto fisso, se esso è soggetto all'ulteriore vincolo che un particolare punto del sistema solidale mantenga coordinate costanti nel sistema fisso.*

Se c'è un punto fisso, allora non è restrittivo, ed è ovviamente conveniente, prendere tale punto come origine sia per la base fissa che per la base solidale. Si vede allora che la cinematica del corpo rigido con un punto fisso si riduce a descrivere la collocazione della terna solidale $e_1e_2e_3$ rispetto alla terna fissa $e_xe_ye_z$. *Ogni configurazione del corpo rigido con un punto fisso resta così definita in modo univoco da una matrice $R \in SO(3)$, precisamente la matrice che manda la terna fissa nella terna solidale.*⁶³

$$e_1 = Re_x, \quad e_2 = Re_y, \quad e_3 = Re_z. \quad (\text{G.1})$$

⁶²Se il corpo è composto, come abbiamo supposto, da almeno tre punti non allineati, si può ad esempio prendere P_1 come origine, il versore e_1 parallelo al segmento P_1P_2 , il versore e_2 perpendicolare a e_1 nel piano contenente P_1, P_2, P_3 , infine il versore e_3 formante con e_1, e_2 una terna ortonormale (destrorsa); elementari considerazioni geometriche consentono di concludere che un sistema siffatto è solidale. La scelta del sistema di riferimento solidale è con evidenza largamente arbitraria. L'esistenza di un sistema solidale si potrebbe usare come definizione alternativa di corpo rigido.

⁶³In letteratura si trova indifferentemente questa convenzione o quella opposta, $e_x = Re_1$ eccetera.

Ciò vuol dire che lo spazio delle configurazioni, o varietà vincolare, del corpo rigido con un punto fisso è il gruppo⁶⁴ $SO(3)$. Ogni matrice $R \in SO(3)$ va pensata come punto della varietà vincolare; ogni moto del corpo rigido si potrà scrivere nella forma $t \mapsto R(t)$, $t \in \mathbb{R}$.

Il gruppo $SO(3)$ è una varietà tridimensionale; lo si vede bene ad esempio perché le matrici 3×3 hanno nove elementi, e la relazione di ortogonalità $RR^T = I$ fornisce sei relazioni indipendenti che li legano. Il corpo rigido con un punto fisso (composto da almeno tre punti non allineati) ha dunque tre gradi di libertà. Una scelta possibile di tre coordinate libere su $SO(3)$, interessante in diverse applicazioni, è quella degli *angoli di Eulero*, descritti sotto nel paragrafo (G.5).

Quando non vi sia un punto fisso invece lo spazio delle configurazioni del corpo è formato dalle coppie (R, r) , con $R \in SO(3)$ ancora definita dalle (G.1) e $r = O - O' \in \mathbb{R}^3$. Lo spazio delle configurazioni in questo caso è $SO(3) \times \mathbb{R}^3$ e ha sei dimensioni.

Ogni vettore u si può rappresentare nelle due basi, solidale e fissa:

$$u = u_1 e_1 + u_2 e_2 + u_3 e_3 = u_x e_x + u_y e_y + u_z e_z ;$$

denotate con

$$u^c = (u_1, u_2, u_3) , \quad u^s = (u_x, u_y, u_z)$$

le terne di componenti nelle due basi, fisse rispettivamente nel corpo e nello spazio, si ha corrispondentemente

$$u^c = R^T u^s , \quad \text{ovvero} \quad u^s = R u^c .$$

La proprietà più rilevante della cinematica rigida è l'esistenza — non ovvia, quando non vi sia un asse fisso — della velocità angolare.

PROPOSIZIONE 42 *Per ogni moto rigido, ad ogni istante di tempo, esiste unico un vettore ω tale che per ogni vettore u solidale risulta*

$$\dot{u} = \omega \times u . \tag{G.2}$$

Il vettore ω è detto *velocità angolare* istantanea.

La (G.2) si può scrivere in coordinate indifferentemente nelle due basi solidale e fissa, nelle quali ω è rappresentato rispettivamente dalle terne $\omega^c = (\omega_1, \omega_2, \omega_3)$ e $\omega^s = (\omega_x, \omega_y, \omega_z)$; in ciascuna base l'operazione $\omega \times \cdot$ è rappresentata dalla moltiplicazione per una matrice antisimmetrica,

$$A^s = \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix} , \quad A^c = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} . \tag{G.3}$$

DIMOSTRAZIONE. Sia $t \mapsto R(t)$ un qualunque moto rigido; per ogni vettore solidale si può scrivere $u(t) = R(t)v$ con v fisso ($v = u_1 e_x + u_2 e_y + u_3 e_z$), dunque $\dot{u} = \dot{R}v$.

Si vede immediatamente che esiste unica A^s antisimmetrica, tale che risulti

$$\dot{R} = A^s R . \tag{G.4}$$

Infatti, derivando la relazione di ortogonalità $RR^T = I$ si ottiene $\dot{R}R^T + R\dot{R}^T = 0$, ovvero $\dot{R}R^T + (\dot{R}R^T)^T = 0$; la matrice

$$A^s = \dot{R}R^T$$

⁶⁴ $SO(3)$, ricordiamo, è il gruppo delle matrici 3×3 ortogonali (cioè tali che $RR^T = I$) con determinante uno (l'ortogonalità implica solo $\det R = \pm 1$).

pertanto è antisimmetrica, e soddisfa la (G.4). Per il solo fatto di essere antisimmetrica, A^s si può scrivere come la prima delle (G.3). La (G.2) segue da $\dot{u} = \dot{R}v = A^s Rv = Au$. \square

Dalla proposizione segue subito

COROLLARIO 43 *Per ogni coppia di punti P e Q fissi nel sistema solidale, si ha*

$$v_P = v_Q + \omega \times (P - Q), \quad (\text{G.5})$$

ove ω è il vettore velocità angolare definito sopra.

(Basta prendere $u = P - Q$.) La (G.5) è nota come *formula fondamentale della cinematica rigida*.

Se il vettore $u = u_1 e_1 + u_2 e_2 + u_3 e_3$ non è solidale, la (G.2) si generalizza con evidenza in

$$\dot{u} = \omega \times u + \dot{u}', \quad (\text{G.6})$$

con

$$\omega \times u = u_1 \dot{e}_1 + u_2 \dot{e}_2 + u_3 \dot{e}_3, \quad \dot{u}' = \dot{u}_1 e_1 + \dot{u}_2 e_2 + \dot{u}_3 e_3.$$

Il termine \dot{u}' , detto *velocità relativa*, rappresenta la velocità di u “percepita da un osservatore solidale”; il termine $\omega \times u$, detto anche *velocità di trascinamento*, è invece la velocità del vettore solidale che istantaneamente coincide con u .

Nel caso di un corpo rigido con un punto fisso O si ha $v_O = 0$ e la (G.5) dà

$$v_P = \omega \times (P - O); \quad (\text{G.7})$$

in particolare si vede che l’asse per O parallelo a ω è asse istantaneo di rotazione: tutti i suoi punti hanno velocità istantanea nulla.

G.2 Cinematica delle masse, operatore di inerzia, base principale

Siano M_1, \dots, M_N le masse di P_1, \dots, P_N . Una prima elementare informazione sulla distribuzione delle masse nel corpo è data dalla posizione del suo baricentro B :

$$B - O = \frac{1}{M} \sum_{i=1}^N M_i (P_i - O), \quad M = \sum_{i=1}^N M_i.$$

Per ogni asse passante per O , che possiamo identificare con un versore e , si definisce poi il *momento di inerzia* I_e relativo all’asse, precisamente

$$I_e = \sum_{i=1}^N M_i d_i^2,$$

ove d_i denota la distanza di P_i dall’asse considerato, $d_i = \|e \times (P_i - O)\|$; con evidenza, O può essere qui sostituito da un qualunque altro punto dell’asse. Tale quantità interviene in particolare nell’espressione dell’energia cinetica del sistema: ad esempio, restringendosi al caso semplice che poi riprenderemo in cui O sia fisso, detto $e_\omega = \omega/\|\omega\|$ il versore di ω , si ha una ben nota proposizione:

PROPOSIZIONE 44 *Per un corpo rigido con un punto fisso, l'energia cinetica $K = \frac{1}{2} \sum_i M_i v_i^2$ è data da*

$$K = \frac{1}{2} I_{e_\omega} \omega^2 . \quad (\text{G.8})$$

DIMOSTRAZIONE. Dalla (G.7) segue subito

$$K = \frac{\omega^2}{2} \sum_{i=1}^N M_i (e_\omega \times (P_i - O))^2 ,$$

e questa è precisamente la (G.8). □

Un'informazione più completa sulla distribuzione della massa del sistema, più precisamente su come la massa è distribuita attorno a un suo punto O , è contenuta in un operatore lineare $\mathcal{J} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, detto operatore di inerzia — una matrice 3×3 , in coordinate — così definito:

$$\mathcal{J}x = \sum_{i=1}^N M_i (P_i - O) \times (x \times (P_i - O)) , \quad x \in \mathbb{R}^3$$

(l'operatore \mathcal{J} dipende dalla scelta di O anche se, per avere una notazione agile, tale dipendenza non appare esplicitamente).

PROPOSIZIONE 45 *L'operatore \mathcal{J} è simmetrico, ovvero per ogni scelta di $x, y \in \mathbb{R}^3$ si ha*

$$y \cdot \mathcal{J}x = x \cdot \mathcal{J}y , \quad (\text{G.9})$$

e (per un corpo composto da almeno tre punti non allineati) definito positivo, ovvero per ogni $x \in \mathbb{R}^3$, $x \neq 0$, si ha

$$x \cdot \mathcal{J}x > 0 . \quad (\text{G.10})$$

Per ogni versore e si ha infine

$$I_e = e \cdot \mathcal{J}e . \quad (\text{G.11})$$

DIMOSTRAZIONE. Poiché

$$y \cdot (P_i - O) \times (x \times (P_i - O)) = (y \times (P_i - O)) \cdot (x \times (P_i - O)) ,$$

si trova

$$y \cdot \mathcal{J}x = \sum_{i=1}^N M_i (y \times (P_i - O)) \cdot (x \times (P_i - O)) ,$$

che è, come richiesto, un'espressione simmetrica in x e y . Per $y = x$ si ha in particolare

$$x \cdot \mathcal{J}x = \sum_{i=1}^N (x \times (P_i - O))^2 ;$$

questa espressione non è mai negativa e si annulla solo se $(P_i - O)$ è parallelo a x per ogni i , cosa che richiederebbe che tutti i punti P_i fossero allineati. Se x è il versore e si ha $(x \times (P_i - O))^2 = d_i^2$, e segue anche la (G.11). □

Non è difficile verificare (esercizio) che nella base solidale $e_1e_2e_3$ l'operatore di inerzia è rappresentato dalla matrice, simmetrica e definita positiva,

$$I = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix},$$

ove I_{kk} è il momento di inerzia relativo all'asse e_k , mentre per $k \neq l$ si è posto

$$I_{kl} = - \sum_{i=1}^N M_i x_k^{(i)} x_l^{(i)}.$$

L'operatore di inerzia interviene in particolare nell'espressione del momento angolare del corpo rigido con un punto fisso,

$$m = \sum_{i=1}^N M_i (P_i - O) \times v_i,$$

e della sua energia cinetica.⁶⁵ Precisamente

PROPOSIZIONE 46 *Per un corpo rigido con un punto fisso il momento angolare m e l'energia cinetica K sono dati da*

$$m = \mathcal{J}\omega, \quad K = \frac{1}{2}\omega \cdot m = \frac{1}{2}\omega \cdot \mathcal{J}\omega = \frac{1}{2}m \cdot \mathcal{J}^{-1}m. \quad (\text{G.12})$$

La (G.12) mostra in particolare che l'angolo tra m e ω è sempre acuto.

DIMOSTRAZIONE. La prima delle (G.12) segue immediatamente esprimendo v_i tramite la (G.7). L'altra è una riscrittura della (G.8), tenendo conto della (G.11). \square

La prima delle (G.12) mostra che il momento angolare m e la velocità angolare ω in generale *non* sono paralleli: lo sono però se ω è autovettore dell'operatore di inerzia. Dal momento che \mathcal{J} è simmetrico, esso ammette una base ortonormale di autovettori, che è evidentemente conveniente prendere come base solidale. Una base solidale in cui e_1, e_2, e_3 siano autovettori dell'operatore di inerzia è detta *base propria*. In una base propria la matrice di inerzia è diagonale,

$$I = \begin{pmatrix} A_1 & & \\ & A_2 & \\ & & A_3 \end{pmatrix},$$

e gli elementi diagonali A_k , $k = 1, 2, 3$, sono contemporaneamente i momenti di inerzia relativi agli assi coordinati della base propria e gli autovalori dell'operatore di inerzia; essi sono detti *momenti principali* di inerzia. Con riferimento alla base propria si hanno le facili e chiare relazioni

$$m = m_1e_1 + m_2e_2 + m_3e_3 \quad \text{con} \quad m_k = A_k\omega_k, \quad k = 1, 2, 3,$$

⁶⁵Nel caso in cui non vi sia un punto fisso, è tuttavia possibile prendere l'origine O nel baricentro e scrivere (teorema di König)

$$K = \frac{1}{2}Mv_O^2 + K', \quad m = M(O - O') \times v_O + m',$$

ove O' è un punto fisso nello spazio, con K' , m' dati dalle (G.12).

e poi

$$K = \frac{1}{2}(A_1\omega_1^2 + A_2\omega_2^2 + A_3\omega_3^2) = \frac{m_1^2}{2A_1} + \frac{m_2^2}{2A_2} + \frac{m_3^2}{2A_3} . \quad (\text{G.13})$$

L'operatore di inerzia, tramite l'equazione

$$x \cdot Jx = 1 ,$$

definisce una quadrica, precisamente un ellissoide, che prende il nome di *ellissoide d'inerzia*. Con evidenza gli assi dell'ellissoide d'inerzia coincidono con gli assi principali di inerzia, inoltre dette a_1, a_2, a_3 le lunghezze dei suoi semiassi si ha $a_i = 1/\sqrt{A_i}$; più in generale, come risulta dalla (G.11), se a_e denota il raggio dell'ellissoide nella direzione del versore e , risulta $a_e = \sqrt{I_e}$. L'ellissoide d'inerzia non contiene informazioni ulteriori rispetto all'operatore di inerzia, ma interviene in modo naturale nella dinamica del corpo rigido.

DEFINIZIONE 20 *Un corpo rigido, con riferimento a un suo punto O , si dice a simmetria cilindrica, o semplicemente simmetrico, se l'ellissoide di inerzia è di rotazione, ovvero se due dei momenti principali di inerzia coincidono; si dice a simmetria sferica se l'ellissoide d'inerzia è una sfera, ovvero se tutti e tre i momenti principali di inerzia coincidono; si dice triassiale se non è simmetrico.*

Ad esempio, $N \geq 3$ masse uguali ai vertici di un poligono regolare di N lati, prendendo come punto di riferimento O il centro del poligono, costituiscono un corpo rigido a simmetria cilindrica; N masse uguali ai vertici di un solido regolare, prendendo ancora come punto di riferimento O il centro del solido, costituiscono un corpo rigido a simmetria sferica.

G.3 Il corpo rigido di Eulero.

Per un corpo rigido con un punto fisso, se il vincolo al punto fisso si suppone ideale e le forze interne che realizzano il vincolo di rigidità si suppongono newtoniane (e dunque ideali), vale la seconda equazione cardinale della dinamica

$$\dot{m} = N , \quad (\text{G.14})$$

ove m e N denotano rispettivamente il momento angolare del corpo e il momento delle forze attive agenti su di esso, riferiti al punto fisso. La (G.14) equivale a tre equazioni scalari, che si dimostrano essere indipendenti e determinano il moto del corpo.⁶⁶

Con l'espressione *corpo rigido di Eulero*, o di *Eulero-Poinsot*, ci si riferisce al corpo rigido con un punto fisso in assenza di forze esterne, o con modesta generalizzazione, in presenza di forze con momento N nullo; la (G.14) diviene allora

$$\dot{m} = 0 .$$

Il corpo rigido di Eulero ha con evidenza quattro costanti del moto: le tre componenti di m nel riferimento fisso, $m_x, m_y,$ e m_z , e l'energia cinetica (G.13). Invece le componenti m_1, m_2, m_3 nella terna solidale in generale non si conservano. Tenendo presente la (G.6), si scrivono facilmente tre equazioni per m_1, m_2, m_3 : dalla (G.6) infatti si ha

$$\dot{m}_1 e_1 + \dot{m}_2 e_2 + \dot{m}_3 e_3 = -\omega \times m = m \times J^{-1} m , \quad (\text{G.15})$$

⁶⁶La questione è un po' delicata. Un modo di vederlo è osservare che la (G.14), attraverso il principio di d'Alembert, è equivalente alle equazioni di Lagrange, che a loro volta sono indipendenti e determinano il moto.

da cui segue immediatamente

$$\dot{m}_1 = -\left(\frac{1}{A_2} - \frac{1}{A_3}\right)m_2m_3, \quad \dot{m}_2 = -\left(\frac{1}{A_3} - \frac{1}{A_1}\right)m_3m_1, \quad \dot{m}_3 = -\left(\frac{1}{A_1} - \frac{1}{A_2}\right)m_1m_2. \quad (\text{G.16})$$

Le (G.16) sono dette *equazioni di Eulero*;⁶⁷ è interessante osservare che esse riguardano solo m (non coinvolgono la configurazione del corpo nello spazio) e sono chiuse. Saltano agli occhi tre famiglie di soluzioni particolari costanti, che si hanno quando una sola delle m_k è diversa da zero, e dunque m si mantiene parallelo a uno degli assi principali di inerzia. Tali moti, di ovvia importanza, sono detti *rotazioni proprie*. Con evidenza, nel caso triassiale le (G.16) non ammettono altre soluzioni costanti.

Nel caso simmetrico (con $A_1 = A_2$) le equazioni di Eulero assumono la forma particolare

$$\dot{m}_1 = -\alpha m_3 m_2, \quad \dot{m}_2 = \alpha m_3 m_1, \quad \dot{m}_3 = 0, \quad \text{con } \alpha = \frac{1}{A_1} - \frac{1}{A_3}. \quad (\text{G.17})$$

Pertanto *nel caso simmetrico* m_3 è costante. Non si tratta però di una quinta costante del moto: dall'espressione dell'energia segue subito, per $A_1 = A_2$,

$$K = \frac{1}{2A_1}(m_1^2 + m_2^2) + \frac{m_3^2}{2A_3} = \frac{1}{2A_1}(\|m\|^2 + \eta m_3^2), \quad \text{ove } \eta = \frac{A_1 - A_3}{A_3},$$

e dunque le costanti indipendenti sono solo quattro. Il coefficiente η è un *fattore di forma* del corpo, che può variare nell'intervallo⁶⁸

$$-\frac{1}{2} \leq \eta < \infty;$$

si osservi che $\eta > 0$ se l'ellissoide di inerzia è *prolato* (allungato, $A_3 < A_1$), $\eta < 0$ se è *oblato* (schiacciato, $A_3 > A_1$), $\eta = 0$ per simmetria sferica. Nel caso simmetrico le equazioni di Eulero si risolvono in modo elementare anche per m_1 e m_2 : come mostrano le (G.17), il vettore

$$m_{\perp} = m_1 e_1 + m_2 e_2$$

ruota uniformemente nel piano equatoriale dell'ellissoide d'inerzia con velocità angolare $\alpha m_3 = -\frac{\eta}{A_1} m_3$.

Nel caso triassiale, una descrizione qualitativa del comportamento di m nel sistema solidale si ottiene dalle leggi di conservazione. Infatti la conservazione di m implica, in particolare, la conservazione di $\|m\| = \|m^s\| = \|m^c\|$, ove si è denotato, coerentemente con la notazione sopra introdotta, $m^s = (m_x, m_y, m_z)$, $m^c = (m_1, m_2, m_3)$; pertanto m_1, m_2, m_3 soddisfano a ogni istante la coppia di equazioni

$$m_1^2 + m_2^2 + m_3^2 = \|m\|^2, \quad \frac{m_1^2}{2A_1} + \frac{m_2^2}{2A_2} + \frac{m_3^2}{2A_3} = E.$$

⁶⁷Strettamente parlando, il termine “equazioni di Eulero” si riferisce di solito alle analoghe equazioni per le componenti ω_k di ω nella base propria, che risultano essere

$$A_1 \dot{\omega}_1 = (A_2 - A_3)\omega_2\omega_3, \quad A_2 \dot{\omega}_2 = (A_3 - A_1)\omega_3\omega_1, \quad A_3 \dot{\omega}_3 = (A_1 - A_2)\omega_1\omega_2;$$

l'equivalenza tuttavia è totale (si passa banalmente dalle une alle altre tramite $m_k = A_k \omega_k$).

⁶⁸Si ricordi che $I_3 \leq I_1 + I_2$, e dunque nel caso simmetrico $I_3 \leq 2I_1$, con uguaglianza che vale solo per un corpo piatto; si ha dunque $\eta = -1/2$ per un corpo piatto, $\eta > -1/2$ in ogni altro caso. Sarebbe $\eta = \infty$ per un corpo con tutti i punti allineati, ma è una situazione singolare che fin dall'inizio abbiamo escluso.

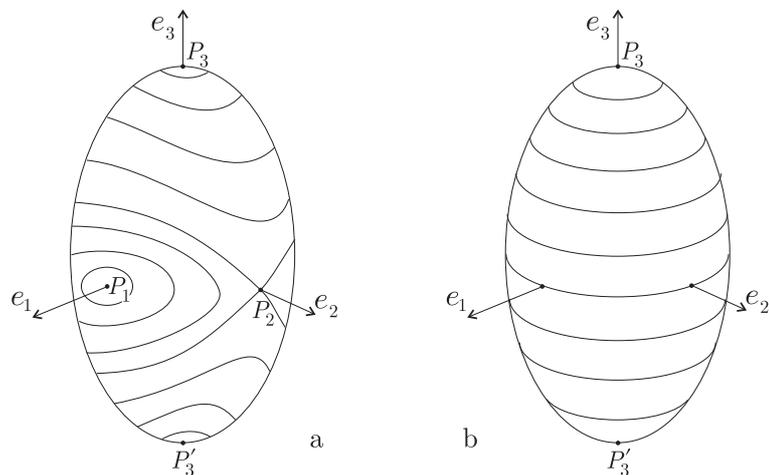


Figura 43: (a) Un ellissoide di energia costante è intersecato da alcune sfere di $\|m\|$ costante. (b) Il caso simmetrico, con e_1, e_2 arbitrari sul piano equatoriale.

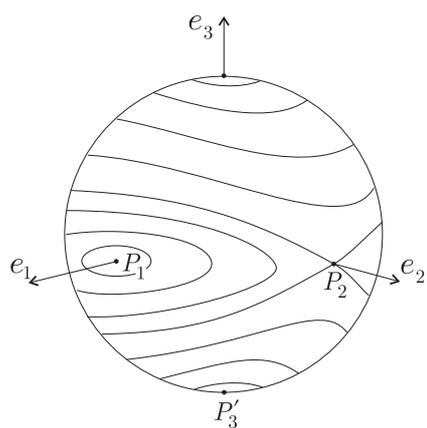


Figura 44: Una sfera di $\|m\|$ costante è intersecata da alcuni ellipsoidi di energia costante (caso triassiale).

Queste definiscono rispettivamente una sfera e un ellissoide; m^c , nel corso del tempo, segue evidentemente l'intersezione tra queste due figure, che genericamente è una curva in \mathbb{R}^3 . La figura 43a rappresenta le curve che si ottengono per E fissata (ellissoide fissato) al variare di $\|m\|$ (del raggio della sfera), e si riferisce al caso triassiale con, per fissare le idee, $A_1 < A_2 < A_3$. Pensiamo ad esempio di far crescere il raggio della sfera. La prima intersezione non vuota si ha per $\|m\|^2 = 2A_1E$, ed è una coppia di punti P_1, P'_1 diametralmente opposti sull'asse e_1 ; i moti corrispondenti sono le rotazioni proprie, nei due versi, attorno a e_1 . Per $2A_1E < \|m\| < 2A_2E$ si hanno curve chiuse attorno a P_1 e P'_1 ; nel corso del moto m^c precede lungo tali curve, senza potersi arrestare (al di fuori delle rotazioni proprie, come si è visto sopra, non possono esservi punti di arresto). Per $\|m\| = 2A_2E$ l'intersezione è complessa: comprende una coppia di punti P_2, P'_2 , corrispondenti alle rotazioni proprie attorno a e_2 , ma anche due *separatrici* che congiungono tali punti; lungo tali curve m precede, senza potersi arrestare, convergendo, per $t \rightarrow \pm\infty$, a P_2 o P'_2 . Per $2A_2E < \|m\| < 2A_3E$ si hanno ancora curve chiuse attorno a punti diametralmente opposti P_3, P'_3 , infine per $\|m\| = 2A_3E$ le curve degenerano in P_3, P'_3 , e si hanno le rotazioni proprie attorno a e_3 . In modo complementare si può invece tenere fisso $\|m\|$, cioè la sfera, e pensare di variare E , cioè l'ellissoide; si veda la figura 44. Nel caso eccezionale di corpo simmetrico, con $A_1 = A_2 \neq A_3$, le intersezioni tra la sfera e l'ellissoide sono cerchi o eccezionalmente punti; si veda la figura 43b. Si osservi che sono scomparse le separatrici, mentre ogni punto dell'equatore corrisponde a una rotazione propria.

L'analisi qualitativa mostra con chiarezza che⁶⁹

PROPOSIZIONE 47 *Per il corpo rigido triassiale, con $A_1 < A_2 < A_3$, le rotazioni proprie attorno agli assi e_1 e e_3 sono stabili, mentre le rotazioni proprie attorno all'asse e_2 sono instabili.*

ESERCIZIO 47 *Si scriva la matrice $\left. \frac{\partial(\dot{m}_1, \dot{m}_2, \dot{m}_3)}{\partial(m_1, m_2, m_3)} \right|_{(0, m_2, 0)}$, e si verifichi che essa ha un autovalore negativo (un modo analitico di dedurre l'instabilità delle rotazioni proprie attorno a e_2). [Risposta: l'equazione agli autovalori è $\lambda^3 - \alpha_1 \alpha_3 \|m\| \lambda = 0$, con $\alpha_1 = (\frac{1}{A_2} - \frac{1}{A_3})$ e simile per α_3 . Le soluzioni sono $\lambda = 0$ (autovettore $(0, 1, 0)$, cioè la direzione $\dim m$) e $\lambda = \pm \|m\| \sqrt{\alpha_1 \alpha_3}$, reali discordi.]*

Nel caso simmetrico, la soluzione esplicita sopra riportata mostra direttamente che le rotazioni proprie attorno all'asse e_3 sono stabili, mentre le rotazioni proprie attorno a un asse equatoriale sono instabili (se m_3 non è esattamente nullo, m_\perp comincia lentamente a precedere lungo l'equatore).

G.4 La descrizione di Poinsot (caso simmetrico)

Avendo compreso il comportamento di m nel corpo, ci resta da descrivere nel suo complesso il moto del corpo nello spazio. La descrizione è particolarmente semplice e chiara nel caso simmetrico, e ci limiteremo a questa, rinviando per il caso generale triassiale a un qualunque testo classico di Meccanica Razionale. Per fissare le idee assumiamo come al solito $A_1 = A_2 \neq A_3$.

PROPOSIZIONE 48 *Si consideri un corpo rigido simmetrico, e si supponga m non parallelo a e_3 .*

i) *I tre vettori m , ω e e_3 sono sempre complanari e formano tra loro angoli costanti; si ha inoltre*

$$\omega = \omega' e_m + \omega'' e_3 \quad \text{con} \quad \omega' = \frac{\|m\|}{A_1}, \quad \omega'' = \eta \frac{m_3}{A_1}. \quad (\text{G.18})$$

⁶⁹Verifica sperimentale: si lanci in aria un libro, cercando di dargli come condizione iniziale una rotazione propria attorno ai tre assi principali, e si osservi la differenza.

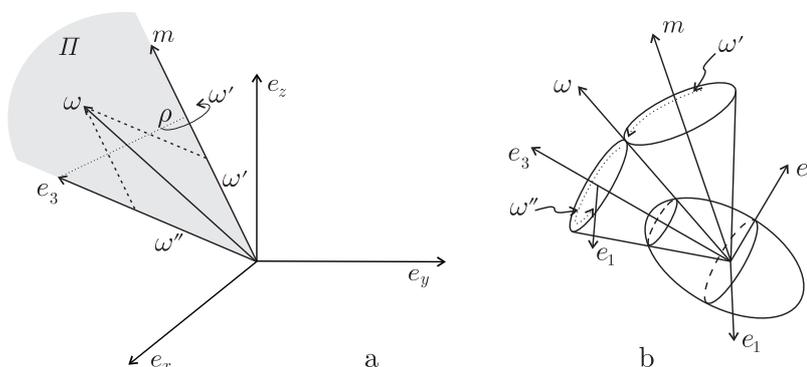


Figura 45: Caso simmetrico: (a) i vettori m , ω e e_3 sono complanari; il loro piano ruota con velocità angolare ω' attorno a m fisso. (b) I coni di Poinsot.

- ii) Il piano Π contenente m , ω e e_3 ruota con velocità angolare costante ω' attorno a m fisso.
- iii) In un sistema di riferimento attaccato a tale piano, il corpo ruota con velocità angolare costante ω'' attorno a e_3 .

La situazione è illustrata in figura 45a.

DIMOSTRAZIONE. Si ha

$$m \times \omega \cdot e_3 = m_1 \omega_2 - m_2 \omega_1 = (A_1 - A_2) \omega_1 \omega_2 = 0,$$

e questo prova la complanarità. Già si è detto che m_3 resta costante, dunque è costante l'angolo tra m e e_3 ; parimenti costante è $\omega_3 = A_3^{-1} m_3$, e dunque anche l'angolo tra ω e e_3 è costante. La decomposizione (G.18) di ω viene da

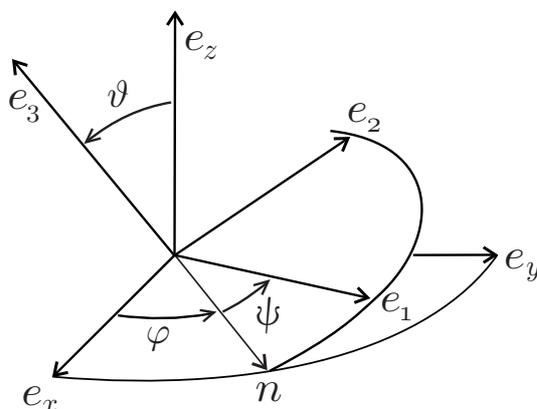
$$\omega = \frac{1}{A_1} (m_1 e_1 + m_2 e_2) + \frac{1}{A_3} m_3 e_3 = \frac{m}{A_1} + \left(\frac{1}{A_3} - \frac{1}{A_1} \right) m_3 e_3,$$

che conduce immediatamente alla (G.18). La velocità angolare di precessione di e_3 , e dunque del piano Π , attorno a m è

$$\frac{\|\dot{e}_3\|}{\varrho}, \quad \text{ove} \quad \varrho = \|e_3 \times e_m\| = \frac{\|e_3 \times m\|}{\|m\|}$$

(ϱ è il raggio del circolo descritto dalla “punta” di e_3 attorno a m , si veda la figura); ma dalla (G.2) e dalla (G.18) si ha $\dot{e}_3 = A_3^{-1} m \times e_3$, pertanto la velocità di precessione viene a coincidere con ω' . Infine, la velocità angolare con cui il corpo ruota attorno a e_3 relativamente al piano Π è opposta alla velocità angolare con cui m precede nel riferimento solidale, e questa si è già vista essere $-\frac{\eta}{A_1} m_3 = -\omega''$. □

Una descrizione sintetica del moto del corpo si ottiene tracciando un cono attorno a m , di apertura tale che ω stia su una direttrice, e un secondo cono attorno a e_3 , tangente al precedente in ω (figura 45b); si pensi il primo fisso nello spazio, il secondo solidale al corpo. Da quanto si è detto, si vede che ω avanza lungo il cono fisso con velocità angolare ω' . Il cono solidale ruota col corpo, mantenendo l'asse e_3 complanare a m e ω : ma come sappiamo, i punti del sistema solidale lungo l'asse ω hanno velocità istantanea nulla, pertanto *il cono solidale rotola senza strisciare sul cono fisso*, come se vi fosse un ingranaggio. I due coni sono detti *coni di Poinsot*.

Figura 46: Gli angoli di Eulero ϑ , φ , ψ .

ESERCIZIO 48 Nella figura 45 la disposizione dei vettori m , ω e e_3 corrisponde al caso di un corpo prolato (ω'' e m_3 hanno lo stesso segno, dunque $\eta > 0$). Come possono essere disposti i vettori per un corpo oblato? E come sono disposti corrispondentemente i coni di Poinsot? (Non è restrittivo supporre $m_3 > 0$; si ricordi che m e ω formano sempre un angolo acuto e inoltre $\omega' > 0$.)

G.5 Gli angoli di Eulero

Le coordinate più comunemente impiegate nella descrizione lagrangiana del corpo rigido con un punto fisso sono gli *angoli di Eulero*. La figura 46 mostra le due terne fissa e solidale e con esse i tre angoli di Eulero⁷⁰ ϑ , φ , ψ . Precisamente, ϑ è l'angolo tra gli assi e_z e e_3 , che si suppone non siano paralleli; φ è l'angolo nel piano equatoriale fisso $e_x e_y$ tra l'asse e_x e la linea dei nodi $n = e_z \times e_3$; ψ è l'angolo nel piano equatoriale mobile $e_1 e_2$ tra la linea dei nodi n e l'asse e_1 .

Il dominio dei tre angoli è evidentemente

$$0 < \vartheta < \pi, \quad 0 \leq \varphi < 2\pi, \quad 0 \leq \psi < 2\pi;$$

il dominio non comprende le configurazioni in cui e_z e e_3 sono paralleli. Gli angoli di Eulero forniscono così soltanto una carta di $SO(3)$, che non lo ricopre; per coprire $SO(3)$ occorrono con evidenza due carte, corrispondenti a due diverse scelte del sistema fisso con assi e_z , e_z' non paralleli.

Non è difficile, anche se un po' laborioso, calcolare la matrice R come funzione degli angoli di Eulero. Un metodo è quello di immaginare tre successive rotazioni, ciascuna attorno a uno degli assi coordinati fissi, che portino la terna fissa in quella solidale.

- i) Una rotazione di ψ attorno a e_z : e_z resta fermo, e_x , e_y ruotano di ψ sul piano coordinato $e_x e_y$ portandosi in e'_x , e'_y . Si veda la figura 47, sinistra.
- ii) Una rotazione di ϑ attorno a e_x : e_z va in e''_z mantenendosi nel piano $e_y e_z$; il piano $e_x e_y$, contenente e'_x e e'_y , ruota come indicato in figura 47, destra; e'_x , e'_y si portano in e''_x , e''_y .
- iii) Una rotazione di φ attorno all'asse z . L'asse x si porta sulla linea dei nodi; l'asse e''_z precede attorno a e_z e va in e_3 ; e''_x e e''_y si portano in e_2 e e_3 . Il risultato è la figura 46.

⁷⁰La notazione con φ e ψ scambiati è altrettanto diffusa.

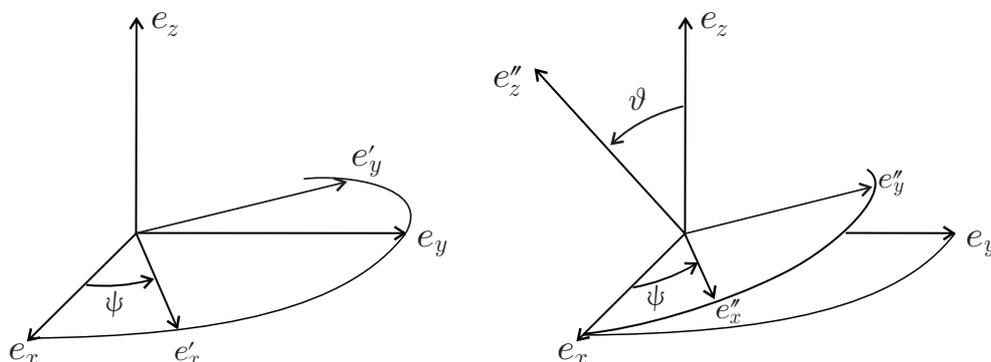


Figura 47: Sinistra: l'azione di R^ψ su e_x, e_y, e_z . Destra: la successiva azione di R^ϑ .

Le tre matrici corrispondenti alle rotazioni sopra descritte sono, nell'ordine,

$$R^\psi = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad R^\vartheta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \vartheta & -\sin \vartheta \\ 0 & \sin \vartheta & \cos \vartheta \end{pmatrix}, \quad R^\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

il loro prodotto $R = R^\varphi R^\vartheta R^\psi$, che con un po' di pazienza si trova essere

$$R = \begin{pmatrix} \cos \varphi \cos \psi - \cos \vartheta \sin \varphi \sin \psi & \sin \varphi \cos \psi + \cos \vartheta \cos \varphi \sin \psi & \sin \vartheta \sin \psi \\ -\cos \varphi \sin \psi - \cos \vartheta \sin \varphi \cos \psi & -\sin \varphi \sin \psi + \cos \vartheta \cos \varphi \cos \psi & \sin \vartheta \cos \psi \\ \sin \vartheta \sin \varphi & -\sin \vartheta \cos \varphi & \cos \vartheta \end{pmatrix},$$

manda la terna e_x, e_y, e_z in e_1, e_2, e_3 .

Quanto alla velocità angolare ω , ci interessa trovare la sua rappresentazione ω^c nella base solidale.

PROPOSIZIONE 49 *Si ha*

$$\omega = (\dot{\vartheta} \cos \psi + \dot{\varphi} \sin \vartheta \sin \psi) e_1 + (-\dot{\vartheta} \sin \psi + \dot{\varphi} \sin \vartheta \cos \psi) e_2 + (\dot{\varphi} \cos \vartheta + \dot{\psi}) e_3. \quad (\text{G.19})$$

DIMOSTRAZIONE. È un conto noioso. Per determinare ω^c basta determinare A^c che compare nelle (G.3). Come per ogni altra matrice, nel cambio di base si ha $A^c = R^T A^s R$, e dunque, dalla (G.4), $A^c = R^T \dot{R}$. Con pazienza si trova il risultato. Se si usa $R = R^\varphi R^\vartheta R^\psi$,

$$\dot{R} = \dot{R}^\varphi R^\vartheta R^\psi + R^\varphi \dot{R}^\vartheta R^\psi + R^\varphi R^\vartheta \dot{R}^\psi,$$

i conti si semplificano un po'. □

Nel seguito in realtà useremo l'espressione di ω^c solo per calcolare l'energia cinetica K , e lo faremo solo per il corpo rigido simmetrico. Per ovvie ragioni di simmetria, K non può dipendere né da φ né da ψ : pertanto basta calcolarla nel caso $\varphi = \psi = 0$. Il calcolo di ω , per $\varphi = \psi = 0$, si semplifica molto perché R_φ e R_ψ vanno a coincidere con l'identità.

PROPOSIZIONE 50 *Per il corpo rigido simmetrico si ha*

$$\begin{aligned} K(\vartheta, \varphi, \psi, \dot{\vartheta}, \dot{\varphi}, \dot{\psi}) &= \frac{1}{2} A_1 (\dot{\vartheta}^2 + \dot{\varphi}^2 \sin^2 \vartheta) + \frac{1}{2} A_3 (\dot{\psi} + \dot{\varphi} \cos \vartheta)^2 \\ &= \frac{1}{2} A_1 \dot{\vartheta}^2 + \frac{1}{2} (A_1 \sin^2 \vartheta + A_3 \cos^2 \vartheta) \dot{\varphi}^2 + \frac{1}{2} A_3 \dot{\psi}^2 + A_3 \cos \vartheta \dot{\varphi} \dot{\psi}. \end{aligned}$$

DIMOSTRAZIONE. Si usa la (G.13); il calcolo è lasciato come esercizio. \square

La lagrangiana del corpo rigido simmetrico è così della forma

$$L(\vartheta, \varphi, \psi, \dot{\vartheta}, \dot{\varphi}, \dot{\psi}) = K(\vartheta, \varphi, \psi, \dot{\vartheta}, \dot{\varphi}, \dot{\psi}) + V(\vartheta, \varphi, \psi),$$

con un qualche potenziale V . Se V , come avviene col potenziale gravitazionale $V = mgl \cos \vartheta$ (baricentro sull'asse e_3 a distanza l dal punto fisso), allora φ e ψ sono coordinate ignorabili. *L'introduzione degli angoli di Eulero come coordinate per $SO(3)$ è proprio adattata a questa situazione di simmetria.*

I momenti p_ϑ , p_φ e p_ψ coniugati agli angoli di Eulero si trovano facilmente essere, nel caso simmetrico che stiamo trattando,

$$p_\vartheta = A_1 \dot{\vartheta}, \quad p_\varphi = (A_1 \sin^2 \vartheta + A_3 \cos^2 \vartheta) \dot{\varphi} + A_3 \cos \vartheta \dot{\psi}, \quad p_\psi = A_3 \cos \vartheta \dot{\varphi} + A_3 \dot{\psi}. \quad (\text{G.20})$$

PROPOSIZIONE 51 *L'hamiltoniana del corpo rigido simmetrico si scrive*

$$H(p_\vartheta, p_\varphi, p_\psi, \vartheta, \varphi, \psi) = \frac{p_\vartheta^2}{2A_1} + \frac{(p_\varphi - \cos \vartheta p_\psi)^2}{2A_1 \sin^2 \vartheta} + \frac{p_\psi^2}{2A_3} + V(\vartheta, \varphi, \psi).$$

DIMOSTRAZIONE. Risolvendo le (G.20) in $\dot{\vartheta}, \dot{\varphi}, \dot{\psi}$, si trova

$$\dot{\vartheta} = \frac{p_\vartheta}{A_1}, \quad \dot{\varphi} = \frac{p_\varphi - \cos \vartheta p_\psi}{A_1 \sin^2 \vartheta}, \quad \dot{\psi} + \cos \vartheta \dot{\varphi} = \frac{p_\psi}{A_3} \quad (\text{G.21})$$

(si ricava $\dot{\psi}$ dalla terza delle (G.20) e si sostituisce nella seconda, ricavando $\dot{\varphi}$; l'ultima è la terza delle (G.20) rigirata, che non conviene sviluppare perché a sinistra dell'uguale c'è l'espressione che compare in K). L'espressione di H segue senza difficoltà. \square

ESERCIZIO 49 *Si verifichi che $p_\varphi = M_z$, $p_\psi = M_3$.*

G.6 La “trottola di Lagrange”.

Quanto visto sopra a proposito degli angoli di Eulero è orientato a studiare il caso di potenziali che condividono con K la simmetria per rotazione attorno a e_z e a e_3 , ovvero potenziali $V(\vartheta)$ indipendenti anch'essi dagli angoli φ e ψ . In questo caso il sistema è con evidenza integrabile, avendo come costanti del moto, oltre a H stesso, i momenti p_φ e p_ψ ; lo spazio delle fasi è foliato in tori \mathbb{T}^3 ovunque la matrice

$$\frac{\partial(H, p_\varphi, p_\psi)}{\partial(p_\vartheta, p_\varphi, p_\psi, \vartheta, \varphi, \psi)} \quad (\text{G.22})$$

abbia rango tre. Di fatto, siamo di fronte a un problema efficace a un solo grado di libertà per la coppia di coordinate canoniche p_ϑ, ϑ , in cui p_φ, p_ψ sono relegati al ruolo di parametri:

$$\mathcal{H}_{p_\varphi, p_\psi}(p_\vartheta, \vartheta) = \frac{p_\vartheta^2}{2A_1} + \mathcal{V}_{p_\varphi, p_\psi}(\vartheta), \quad \mathcal{V}_{p_\varphi, p_\psi}(\vartheta) = V(\vartheta) + \frac{(p_\varphi - \cos \vartheta p_\psi)^2}{2A_1 \sin^2 \vartheta}$$

(si è ommesso il termine costante $\frac{p_\psi^2}{2A_3}$).

ESERCIZIO 50 *Si determini dove la matrice (G.22) ha una caduta di rango. [Risposta: sugli equilibri relativi, ovvero sugli equilibri di \mathcal{H} .]*

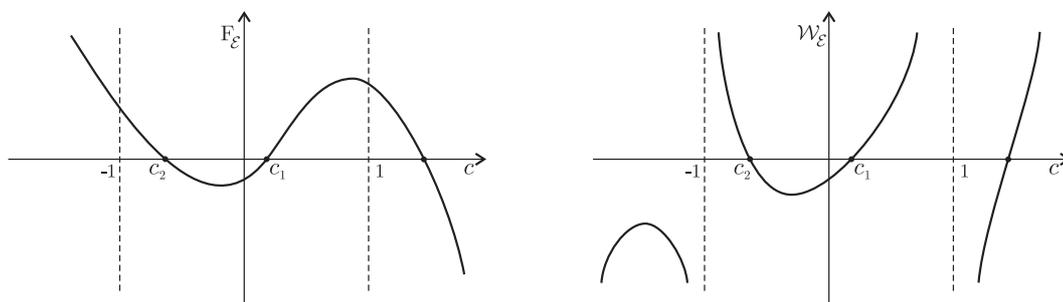


Figura 48: Una rappresentazione schematica di F_ε (sinistra) e di \mathcal{W}_ε (destra) come funzioni di $c = \cos \vartheta$, nelle condizioni in cui esistono zeri.

Proseguiamo considerando il caso della gravità, $V(\vartheta) = mgl \cos \vartheta$; per alleggerire la notazione ometteremo il pedice p_φ, p_ψ nelle funzioni che pur dipendono da questi parametri. L'hamiltoniana si scrive allora, raccogliendo il fattore $1/A_1$,

$$\mathcal{H}(p_\vartheta, \vartheta) = \frac{1}{A_1} \left[\frac{p_\vartheta^2}{2} + \mathcal{V}(p_\vartheta, \vartheta) \right],$$

con

$$\mathcal{V}(\vartheta) = \alpha \cos \vartheta + \frac{(p_\varphi - \cos \vartheta p_\psi)^2}{2 \sin^2 \vartheta}, \quad \alpha = A_1 mgl.$$

Le costanti dimensionali si eliminano con un elementare riscalamento,⁷¹

$$p_\vartheta = \sqrt{\alpha} p'_\vartheta, \quad p_\varphi = \sqrt{\alpha} p'_\varphi, \quad p_\psi = \sqrt{\alpha} p'_\psi, \quad t = A_1 / \sqrt{\alpha} t', \quad \mathcal{H}' = \mathcal{H} A_1 / \sqrt{\alpha};$$

togliendo gli apici si resta con

$$\mathcal{H}(p_\vartheta, \vartheta) = \frac{p_\vartheta^2}{2} + \mathcal{V}(\vartheta), \quad \mathcal{V}(\vartheta) = \cos \vartheta + \frac{(p_\varphi - \cos \vartheta p_\psi)^2}{2 \sin^2 \vartheta}.$$

Le traiettorie nel piano ϑ, p_ϑ si trovano a partire da \mathcal{V} con la consueta analisi qualitativa. Ci limitiamo qui al caso generico $p_\varphi \neq \pm p_\psi$. Innanzitutto, il potenziale \mathcal{V} diverge a $+\infty$ per $\vartheta \rightarrow 0, \pi$, perciò le traiettorie non incappano nelle singolarità delle coordinate di Eulero. L'analisi di \mathcal{V} mostra poi che \mathcal{V} ha un unico minimo in $0 < \vartheta < \pi$, e corrispondentemente le traiettorie, per ogni valore di energia \mathcal{E} superiore al minimo, sono singole curve chiuse comprese tra due punti di arresto ϑ_1, ϑ_2 .

In effetti è più facile dedurre direttamente questa seconda affermazione, ricavandone la conseguenza che \mathcal{V} ha un unico minimo, piuttosto che analizzare \mathcal{V} studiandone le derivate. Si vede infatti facilmente che *per ogni valore di \mathcal{E} , il grafico di \mathcal{V} è tagliato dalla retta $\mathcal{V} = \mathcal{E}$ in nessuno o due punti (eventualmente coincidenti)*. Per vederlo studiamo gli zeri di $\mathcal{V}(\vartheta) - \mathcal{E}$, utilizzando per comodità $c = \cos \vartheta$ come variabile. La funzione da studiare è allora

$$\mathcal{W}_\varepsilon(c) = \frac{F_\varepsilon(c)}{1 - c^2}, \quad F_\varepsilon(c) = (c - \varepsilon)(1 - c^2) + \frac{1}{2}(p_\varphi - cp_\psi)^2,$$

e ci si riduce a studiare gli zeri di $F_\varepsilon(c)$ in $-1 < c < 1$. Lo studio è elementare: infatti F_ε è una cubica, che diverge a $\mp\infty$ per $c \rightarrow \pm\infty$, ed è positiva in $c = \pm 1$; necessariamente uno zero reale è

⁷¹Il riscalamento per p_φ, p_ψ sdi legge indifferentemente come riscalamento dei momenti nel problema a tre gradi di libertà, come ridefinizione dei parametri nel problema a un grado di libertà.

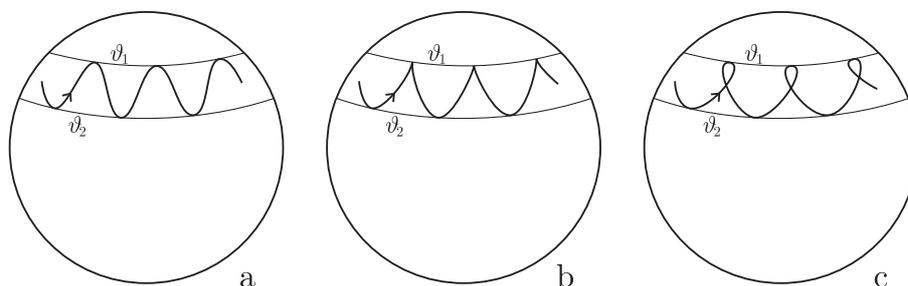


Figura 49: I possibili moti di e_3 sulla sfera unitaria.

tra 1 e $+\infty$, mentre nell'intervallo $-1 < c < 1$ o non ci sono zeri o ce n'è una coppia c_1, c_2 : si veda la figura 48, sinistra. Corrispondentemente $\mathcal{W}_\varepsilon(c)$ diverge in $c = \pm 1$, precisamente a $+\infty$ all'interno dell'intervallo $-1 < c < 1$ e a $-\infty$ all'esterno, e il suo grafico è grosso modo come in figura 48, destra. Tornando alla variabile ϑ , i valori di c al di fuori dell'intervallo $-1 \leq c \leq 1$ corrispondono a ϑ complesso; dentro all'intervallo invece $\mathcal{V}(\vartheta) - \mathcal{E}$ ha nessuno o una coppia di zeri, ovvero il grafico di \mathcal{V} , come sopra preteso, è tagliato dalle rette $\mathcal{V} = \mathcal{E}$ in due punti di arresto ϑ_1, ϑ_2 , oppure non vi sono intersezioni. Situazioni più complicate di un singolo minimo, come potrebbe essere una doppia buca, magari con biforcazioni, sono escluse.

Veniamo ora al moto dell'asse e_3 del corpo sulla sfera unitaria; per essere definiti supponiamo $p_\varphi, p_\psi > 0$, $0 < \vartheta_1 < \vartheta_2 < \pi/2$ (gli altri casi sono del tutto analoghi). Per quanto visto sopra l'angolo ϑ oscilla tra i due valori ϑ_1, ϑ_2 , corrispondentemente e_3 oscilla sulla sfera tra i due paralleli $\vartheta = \vartheta_1$ e $\vartheta = \vartheta_2$; tale moto è detto *nutazione*. L'espressione (G.21) di $\dot{\varphi}$ mostra poi che vi sono tre situazioni possibili: se il numeratore $p_\varphi - \cos \vartheta p_\psi$ non ha zeri tra ϑ_1 e ϑ_2 , allora $\dot{\varphi}$ ha segno definito e φ avanza monotonamente; si veda la figura 49a. Se invece il numeratore è negativo in ϑ_1 e positivo in ϑ_2 , allora il moto di e_3 forma occhielli; si veda la figura 49c. Infine, se il numeratore è nullo in ϑ_1 e positivo in ϑ_2 , si ha il grafico intermedio di figura 49b; la situazione è menzionata, benché ovviamente eccezionale, perchè corrisponde al caso interessante in cui l'asse e_3 è abbandonato inizialmente con velocità nulla: $\dot{\vartheta} = 0$, $\dot{\varphi} = 0$ (la gravità comincia a far cadere il corpo, che però poco dopo risale alle condizioni di partenza). Qualche considerazione in più sul caso comune di trottola che ruota velocemente si può trovare nei comuni testi di meccanica (in particolare in V.I. Arnol'd, *Metodi matematici della meccanica classica*, paragrafo 30).

H L'equazione di Van der Pol

L'equazione di Van der Pol, proposta negli anni '30 per far da modello alle oscillazioni che si instaurano in certe condizioni in un circuito elettrico amplificato, si scrive⁷²

$$\ddot{x} + \varepsilon(x^2 - 1)\dot{x} + x = 0, \quad \varepsilon > 0, \quad x \in \mathbb{R},$$

o equivalentemente

$$\dot{x} = v, \quad \dot{v} = -x - \varepsilon(x^2 - 1)v;$$

⁷²Questa su cui lavoreremo è la forma adimensionale dell'equazione. Proviene da un'equazione del tipo

$$\ddot{x} + \mu(x^2/a^2 - 1)\dot{x} + \omega^2 x = 0,$$

e la si deduce ponendo $\mu = \varepsilon\omega$ e usando a come unità per la variabile x e ω^{-1} come unità di tempo.

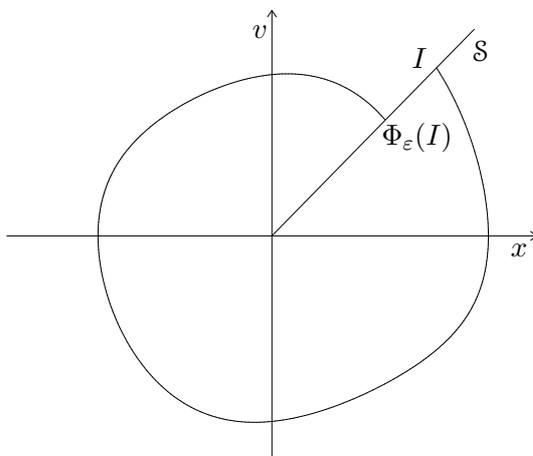


Figura 50: La mappa $\Phi_\varepsilon : \mathcal{S} \rightarrow \mathcal{S}$.

nella sostanza, è un oscillatore armonico smorzato di pulsazione uno, con coefficiente di smorzamento $\mu(x, \varepsilon) = \varepsilon(x^2 - 1)$ positivo per $|x| > 1$ (attrito ordinario), e negativo per $|x| < 1$. La forma semplice di μ consente di eseguire i calcoli con facilità, ma è sostituibile con una classe assai più vasta di funzioni.⁷³

Introduciamo le variabili di azione–angolo dell'oscillatore armonico che si ha per $\varepsilon = 0$:

$$v = \sqrt{2I} \cos \varphi, \quad x = \sqrt{2I} \sin \varphi,$$

e viceversa

$$I = \frac{1}{2}(x^2 + v^2), \quad \varphi = \arctan \frac{x}{v} \tag{H.1}$$

(si sottintende $\varphi = \arctan \frac{x}{v} + \pi$, se $v < 0$; poiché la pulsazione è uno, I coincide con l'energia. Le equazioni per I e φ si trovano subito derivando le (H.1):

$$\dot{I} = \varepsilon f(I, \varphi), \quad \dot{\varphi} = 1 + \varepsilon g(I, \varphi),$$

con

$$f(I, \varphi) = -2I \cos^2 \varphi (2I \sin^2 \varphi - 1), \quad g = \sin \varphi \cos \varphi (2I \sin^2 \varphi - 1)$$

(si osservi che \dot{I} è la potenza delle forze dissipative, $\dot{I} = -\varepsilon(x^2 - 1)v^2$). L'espressione di g , aggiunta qui per completezza, in realtà non giocherà alcun ruolo.

Per $\varepsilon = 0$ si ha un oscillatore armonico di pulsazione uno, dunque il ritratto in fase consiste di infinite curve chiuse, precisamente cerchi concentrici, che vengono percorse con velocità costante nel periodo $T_0 = 2\pi$. Non appena ε è positivo si ha invece la formazione di un *ciclo limite attrattivo*: una sola delle curve chiuse sopravvive, deformata di ε , alla perturbazione, ed è attrattiva per l'intero piano privato dell'origine (che resta un punto di equilibrio), precisamente per dati iniziali rispettivamente interni e esterni alla curva il moto spiraleggia verso l'esterno o verso l'interno,

⁷³Quanto diremo in effetti si estende in modo ovvio a equazioni del tipo

$$\ddot{x} + \varepsilon \mu(x) \dot{x} + x = 0$$

(equazioni di Lennard), con l'unico requisito che $\mu(x)$ sia negativo per $|x|$ piccolo e positivo per $|x|$ grande.

convergeno in ogni caso alla curva. Il sistema così non ha più infiniti moti periodici, ma uno solo, cui asintoticamente per $t \rightarrow \infty$ convergono tutti gli altri.⁷⁴

Diamo qui una dimostrazione di esistenza del ciclo limite di tipo perturbativo, e dunque valida per ε piccolo.

PROPOSIZIONE 52 *Per ε sufficientemente piccolo esiste un ciclo limite attrattivo prossimo al circolo di raggio 2, con periodo prossimo a 2π .*

DIMOSTRAZIONE. Fissiamo l'attenzione su una qualunque semiretta \mathcal{S} uscente dall'origine nel piano xv (figura 50), con $\varphi = \bar{\varphi}$ assegnato. Per $\varepsilon = 0$ tutte le traiettorie tagliano \mathcal{S} trasversalmente, e impiegano il tempo $T_0 = 2\pi$ a fare un giro completo (\mathcal{S} è una buona "sezione di Poincaré"). Per ε piccolo e dato iniziale $(I, \bar{\varphi})$ su \mathcal{S} , se denotiamo con $\hat{I}_\varepsilon(t, I)$, $\hat{\varphi}_\varepsilon(t, I)$ la corrispondente soluzione, si ha evidentemente⁷⁵

$$\hat{I}_\varepsilon(t, I) = I + \mathcal{O}(\varepsilon) , \quad \hat{\varphi}_\varepsilon(t, I) = \bar{\varphi} + t + \mathcal{O}(\varepsilon) .$$

Si vede che φ avanza monotonamente con velocità prossima a uno, pertanto dopo un tempo $T_\varepsilon(I) = T_0 + \mathcal{O}(\varepsilon)$ la traiettoria interseca \mathcal{S} , e il punto di intersezione $\Phi_\varepsilon(I)$ è prossimo a I :

$$\Phi_\varepsilon(I) = I + \varepsilon F_\varepsilon(I) .$$

Se $F_\varepsilon(I) > 0$ la traiettoria è un arco di spirale verso l'esterno, e da quel momento in poi (non potendo la traiettoria intersecare se stessa) il moto procede a spirale verso l'esterno. Viceversa se $F_\varepsilon(I) < 0$ il moto procede a spirale verso l'interno. Infine se $F_\varepsilon(I) = 0$ la traiettoria è chiusa e il moto è periodico. La ricerca di una traiettoria periodica si riduce così alla ricerca di uno zero di F_ε .

Si trova ora senza difficoltà una buona approssimazione di F_ε , precisamente

$$F_\varepsilon(I) = -I(aI - b) + \mathcal{O}(\varepsilon) \tag{H.2}$$

con a, b costanti positive:

$$a = 4 \int_0^{2\pi} \sin^2 \varphi \cos^2 \varphi \, d\varphi = \pi , \quad b = 2 \int_0^{2\pi} \cos^2 \varphi \, d\varphi = 2\pi .$$

Per vederlo, scriviamo innanzitutto

$$\hat{I}_\varepsilon(T_\varepsilon, I) = I + \varepsilon \int_0^{T_\varepsilon} f(\hat{I}_\varepsilon(t, I), \hat{\varphi}_\varepsilon(t, I)) \, dt ,$$

⁷⁴Il ciclo limite è un esempio di *attrattore*. Data un'equazione differenziale $\dot{x} = X(x)$ in \mathbb{R}^n (o su una varietà), un insieme chiuso $Z \subset \mathbb{R}^n$ si dice essere un'attrattore per l'equazione se esiste un aperto B che lo contiene, detto *bacino di attrazione* di Z , tale che per $t > 0$ risulti $\Phi_X^t(B) \subset B$ e

$$\bigcap_{t>0} \Phi_X^t(B) = Z .$$

Si vede facilmente (si usa la proprietà di gruppo di Φ_X^t) che per t crescente gli insiemi $\Phi_X^t(B)$ sono annidati e che Z necessariamente è invariante. Si ha un ciclo limite quando Z è una curva chiusa priva di punti critici e dunque una traiettoria. L'unione dei bacini di attrazione costituisce il bacino di attrazione massimale; per l'equazione di Van der Pol il bacino di attrazione massimale del ciclo limite è tutto \mathbb{R}^2 privato dell'origine. Tra gli anni '50 e '70 del '900 si è capito che (in dimensione più alta di due) è tipica, per sistemi dissipativi, la formazione di *attrattori strani*, che non hanno la struttura di varietà: non sono lisci e tagliati trasversalmente al moto danno insiemi totalmente disconnessi come l'insieme di Cantor.

⁷⁵Si sta qui usando la sola regolarità della soluzione al variare dei parametri dell'equazione, per t limitato.

che dà

$$F_\varepsilon(I) = \int_0^{T_\varepsilon} f(\hat{I}_\varepsilon(t, I), \hat{\varphi}_\varepsilon(t, I)) dt .$$

Osserviamo poi che F_ε è funzione regolare in ε , pertanto $F_\varepsilon(I) = F_0(I) + \mathcal{O}(\varepsilon)$. Ma il calcolo di F_0 è del tutto elementare: per $\varepsilon = 0$ infatti il moto incognito $(\hat{I}_\varepsilon(t, I), \hat{\varphi}_\varepsilon(t, I))$ è rimpiazzato dal moto imperturbato, noto, $(\hat{I}_0(t, I), \hat{\varphi}_0(t, I)) = (I, \bar{\varphi} + t)$, pertanto si ha

$$\begin{aligned} F_0(I) &= \int_0^{T_0} f(I, \bar{\varphi} + t) dt = \int_0^{2\pi} f(I, \varphi) d\varphi \\ &= - \int_0^{2\pi} 2I \cos^2 \varphi (2I \sin^2 \varphi - 1) d\varphi . \end{aligned}$$

L'integrale è elementare e la (H.2) è immediata. I valori esatti di a e b non sono importanti (se non per localizzare il ciclo limite): conta solo che siano costanti positive, e questo è ovvio anche senza fare il calcolo.

Per $\varepsilon = 0$ la curva F_ε si riduce dunque alla parabola $F_0(I) = -I(aI - b)$, che passa per l'origine e taglia l'asse delle ascisse in $I_0^* = b/a = 2$. La curva $F_\varepsilon(I)$ sicuramente passa anch'essa per l'origine (si veda l'equazione per I), e per ε piccolo, dovendo differire di poco da F_0 , sicuramente⁷⁶ taglia anch'essa l'asse delle ascisse in un unico punto I_ε^* prossimo a I_0 :

$$I_\varepsilon^* = I_0^* + \mathcal{O}(\varepsilon) .$$

Si ha così $\Phi_\varepsilon(I_\varepsilon^*) = I_\varepsilon^*$, e corrispondentemente il sistema ha una traiettoria periodica con dato iniziale $(I_\varepsilon^*, \bar{\varphi})$, che per $\varepsilon = 0$ si riduce al circolo di raggio $I_0^* = 2$.

Si dimostra facilmente che tale traiettoria è attrattiva: precisamente, dalla forma di F_ε si vede subito che per ogni fissato dato iniziale $I > 0$ diverso da I_ε^* , la successione I^0, I^1, I^2, \dots definita da $I^0 = I$ e $I^k = \Phi_\varepsilon(I^{k-1})$, è monotona crescente per $I^0 < I_\varepsilon^*$ e decrescente per $I^0 > I_\varepsilon^*$. Ma il limite della successione è necessariamente un punto fisso (per questo basta la continuità di Φ al variare del dato iniziale), e dunque coincide con I_ε^* . \square

L'andamento qualitativo di F_ε e Φ_ε , assieme alla costruzione grafica degli I_k , è rappresentato in figura 51. La figura 52, ottenuta integrando numericamente le equazioni del moto, rappresenta il ciclo limite e due moti che vi si avvicinano dall'interno e dall'esterno, per $\varepsilon = 0.1$.

Veniamo ora al confronto con l'equazione mediata, che si scrive

$$J = \varepsilon \mathcal{F}(J) , \quad \text{ove} \quad \mathcal{F}(J) = \frac{1}{2\pi} \int_0^{2\pi} f(J, \varphi) d\varphi ;$$

l'integrale è, significativamente, lo stesso che abbiamo incontrato nel corso della dimostrazione, e dà

$$\mathcal{F}(J) = -\frac{1}{2\pi} J(aJ - b) ,$$

ove a e b sono le costanti introdotte sopra. L'analisi dell'equazione mediata è semplicissima: con evidenza l'equazione ammette la soluzione stazionaria $J(t) = I_0^* = b/a$, e tale soluzione è attrattiva:

⁷⁶Formalmente lo si deduce dal teorema della funzione implicita: denotando per comodità $F(I, \varepsilon)$ anziché $F_\varepsilon(I)$, si ha che se $F(I^*, 0) = 0$ e $\frac{\partial F}{\partial I}(I^*, 0) \neq 0$, allora in un intorno di $\varepsilon = 0$ resta implicitamente definita una funzione $I^*(\varepsilon)$ tale che $F(I^*(\varepsilon), \varepsilon) = 0$. La persistenza degli zeri trasversi delle funzioni, o più in generale delle intersezioni trasverse tra curve o varietà, è un fatto molto generale, valido in dimensione qualsiasi, che è conseguenza del teorema della funzione implicita.

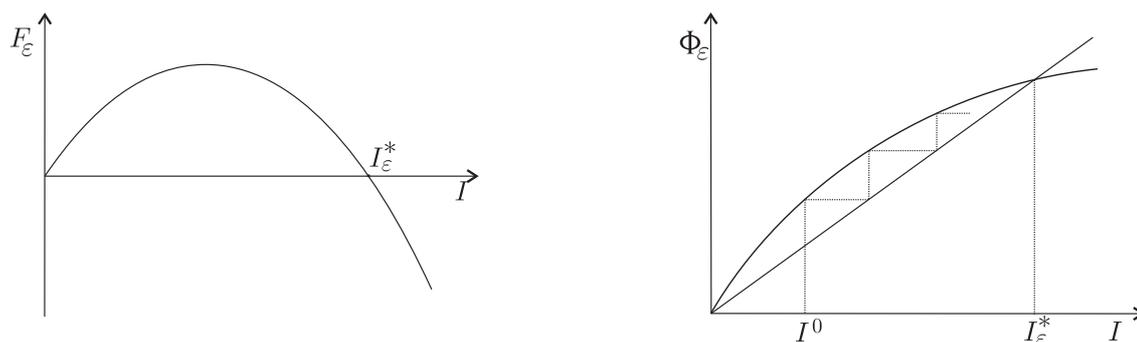


Figura 51: *L'andamento di F_ε e Φ_ε .*

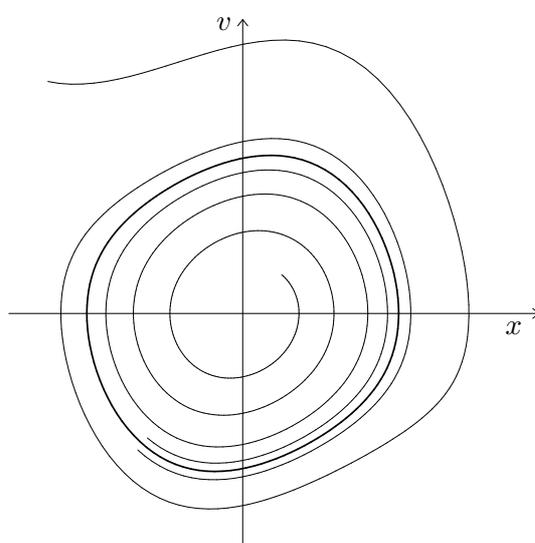


Figura 52: *Il ciclo limite per $\varepsilon = 0.1$, con due moti che vi convergono dall'interno e dall'esterno.*

si ha infatti $\dot{J} > 0$ per $0 < J < I_0^*$ e $\dot{J} < 0$ per $J > I_0^*$. Si vede così che il comportamento asintotico dell'azione vera e di quella mediata sono identiche entro $\mathcal{O}(\varepsilon)$.

Ma anche durante il transiente le soluzioni $J_\varepsilon(t)$ e $I_\varepsilon(t)$, a parità di dato iniziale, restano vicine entro $\mathcal{O}(\varepsilon)$. Infatti, innanzitutto basta confrontare I e J al passaggio attraverso \mathcal{S} (infatti tra un passaggio e l'altro entrambe si spostano solo di $\mathcal{O}(\varepsilon)$). Poi si osserva facilmente che la soluzione vera e quella mediata a ogni giro si separano di $\mathcal{O}(\varepsilon^2)$. Ne consegue che dopo un numero di passaggi $\mathcal{O}(1/\varepsilon)$, corrispondente a $t = 1/\varepsilon$, le soluzioni a parità di dato iniziale sono ancora ε -vicine. Il principio della media è così soddisfatto. Con poco lavoro in più si può estendere il confronto tra equazione vera e equazione mediata a tempi qualsiasi, al di là di quanto chiede il principio: per questo basta osservare che il termine $\mathcal{O}(\varepsilon^2)$ che separa a ogni giro le due soluzioni è anche proporzionale alla distanza di ciascuna dal valore limite, e di conseguenza la somma degli scarti converge.

L'equazione di Van der Pol ammette in realtà un ciclo limite attrattivo per ogni $\varepsilon > 0$, anche non piccolo, che risulta essere attrattore per l'intero piano xv privato dell'origine. Si veda ad esempio G. Benettin, *Appunti per il corso di Fisica Matematica*, capitolo I, appendice C (www.math.unipd.it/~benettin), ove si trova anche una giustificazione dell'equazione in connessione alle auto-oscillazioni di un circuito oscillante amplificato reazionato positivamente.

I Il modello classico per la precessione degli equinozi

(a cura di Massimiliano Guzzo)

Il fenomeno per cui l'asse di rotazione della Terra ruota attorno alla verticale dell'orbita con un periodo di circa 26000 anni viene detto precessione degli equinozi. Per poter studiare questo fenomeno è necessario innanzitutto definire un modello matematico che consente lo studio della rotazione della Terra. Purtroppo, lo studio di un modello realistico che tiene conto delle forze più importanti che agiscono sulla Terra è estremamente complicato. Tuttavia, esso è una piccola perturbazione di un sistema hamiltoniano integrabile, detto modello classico della precessione. Se si osserva la rotazione della Terra su tempi non troppo lunghi (diciamo dell'ordine di 10^5 anni), il modello classico di precessione descrive bene la dinamica: essa appare quasi-periodica, ed una delle frequenze è la frequenza di precessione dell'asse di simmetria della Terra rispetto alla verticale dell'orbita. Come vedremo, il modello classico di precessione è abbastanza semplice da studiare, ed in effetti in queste note riusciremo a darne una motivazione, a descriverne i moti e fornire una stima della frequenza di precessione degli equinozi.

Se invece si vuole descrivere la rotazione della Terra su tempi più lunghi (dal milione di anni in su) allora le perturbazioni al modello classico di precessione possono contribuire in modo non banale alla dinamica, e lo studio si complica notevolmente (in effetti per molti aspetti è argomento di ricerca attuale⁷⁷).

I.1 Il modello hamiltoniano

Per semplicità di esposizione, partiamo da un modello già semplificato rispetto al modello vero (che al solito è una sua perturbazione). Il modello di partenza consiste di un corpo rigido a simmetria assiale il cui centro di massa B descrive un'orbita circolare attorno al Sole O con frequenza

⁷⁷Fra le pubblicazioni più significative sulla dinamica della rotazione terrestre in modelli realistici si consiglia: Laskar, J.; Robutel, P.: "The chaotic obliquity of the planets" *Nature*, vol. 361, no. 6413, p. 608-612. Laskar, J.; Joutel, F.; Robutel, P. "Stabilization of the earth's obliquity by the moon", *Nature*, vol. 361, no. 6413, p. 615-617.

kepleriana Ω , ovvero

$$\begin{aligned} B(t) - O &= a(\cos(\Omega t), \sin(\Omega t), 0) \\ \Omega &= \sqrt{\frac{\gamma M}{a^3}} \end{aligned}$$

ove a è il raggio dell'orbita, M la massa del Sole e γ la costante gravitazionale. Si tratta di un sistema vincolato a tre gradi di libertà dipendente dal tempo e dunque la funzione di Hamilton $H = K + V$ sarà definita sull'insieme $\mathcal{A} \times \mathbb{R}$, ove \mathcal{A} è lo spazio delle fasi del corpo rigido, K è l'energia cinetica del sistema e V quella potenziale gravitazionale. Siccome l'energia cinetica del baricentro è costante, essa non contribuisce alla funzione di Hamilton, e dunque possiamo sostituire K con l'energia cinetica K' relativa al baricentro. L'energia potenziale sarà invece data da

$$V = - \sum_{i=1}^N \frac{\gamma m_i M}{\|P_i - O\|}$$

ove P_1, \dots, P_N sono i punti di massa m_1, \dots, m_N che costituiscono il corpo. Le funzioni K' e V vanno poi espresse come funzioni delle coordinate canoniche di \mathcal{A} . Ai fini del calcolo della frequenza di precessione è conveniente utilizzare le coordinate canoniche di Andoyer–Deprit G, L, J, g, l, j introdotte nel paragrafo 3.1. In queste coordinate la funzione K' diventa

$$K'(G, L) = \frac{G^2}{2I_1} + \frac{\eta L^2}{2I_1}$$

ove

$$\eta = \frac{I_1 - I_3}{I_3} < 0 .$$

Il calcolo di V nelle coordinate di Andoyer–Deprit è invece più complicato. Tuttavia possiamo facilmente dedurre qualche informazione preliminare dalle simmetrie del problema. Innanzitutto V non dipende da l , per la simmetria di rotazione del corpo attorno all'asse e_3 . Pertanto L è un integrale primo, e lungo il moto si ha

$$L(t) = L_0 .$$

Siccome V dipende solo dalla configurazione del corpo, e non dal modulo del momento angolare, V dipende dalle azioni solo attraverso i rapporti L/G e J/G . Inoltre, la simmetria di rotazione attorno ad e_z ci dice che V dipende da j e dal tempo solo attraverso la differenza $j - \Omega t$. Dunque, il potenziale è una funzione del tipo

$$V\left(\frac{L}{G}, \frac{J}{G}, g, j - \Omega t\right) .$$

È conveniente sviluppare preliminarmente V in serie di Fourier rispetto a g

$$V = V_0\left(\frac{L}{G}, \frac{J}{G}, j - \Omega t\right) + \sum_{k \in \mathbb{Z} \setminus 0} V_k\left(\frac{L}{G}, \frac{J}{G}, j - \Omega t\right) e^{ikg} .$$

Si potrebbe verificare che ciascuna armonica V_k è in realtà una funzione del tipo

$$V_k = \left(1 - \frac{L}{G}\right)^{\frac{|k|}{2}} \tilde{V}_k\left(\frac{L}{G}, \frac{J}{G}, j - \Omega t\right) ,$$

ove \tilde{V}_k è una funzione limitata. Volendo studiare le equazioni della precessione per condizioni iniziali in cui il momento angolare è quasi allineato all'asse di simmetria, cioè⁷⁸ $G \simeq L$, allora l'hamiltoniana è perturbazione di ordine $\sqrt{1 - \frac{L}{G}}$ della seguente funzione di Hamilton:

$$H_0 = \frac{G^2}{2I_1} + \frac{\eta L^2}{2I_1} + V_0\left(\frac{L}{G}, \frac{J}{G}, j - \Omega t\right).$$

L'Hamiltoniana H_0 definisce il modello di precessione classica, che ben descrive la precessione se non vi sono risonanze di spin-orbita. Lo studio del modello che include anche i termini di ordine $(1 - \frac{L}{G})^{|k|/2}$ sarebbe possibile con la teoria hamiltoniana delle perturbazioni ad un livello che è al di là di quello che si considera noto nella presente discussione. In queste note ci limitiamo a descrivere i moti del modello classico di precessione.

I.2 Il modello classico di precessione

Le soluzioni delle equazioni di Hamilton per H_0 sono tali che, oltre ad L , anche G è un integrale del moto. Inoltre H_0 dipende solo da un angolo, precisamente dalla combinazione $j - \Omega t$, dunque è integrabile. In linea di principio si potrebbero introdurre coordinate di angolo azione per H_0 e calcolare esattamente la frequenza di precessione. In realtà è più semplice considerare V_0 come perturbazione (vedremo che effettivamente ciò è possibile) ed eliminare dall'hamiltoniana l'angolo $j - \Omega t$ con una media. Vedremo che per effettuare questa operazione di media non dovremo preoccuparci troppo dell'unica possibile risonanza, $\frac{dj}{dt} - \Omega \simeq 0$: Ω infatti è la frequenza dell'anno mentre $\frac{dj}{dt} = \frac{\partial V_0}{\partial J}$ sarà dell'ordine di 10^4 anni.

Preliminarmente consideriamo l'estensione di H_0 ad un sistema autonomo, che si fa introducendo le variabili coniugate Λ, λ e l'hamiltoniana

$$\mathcal{H}_0 = \frac{G^2}{2I_1} + \frac{\eta L^2}{2I_1} + \Omega \Lambda + V_0\left(\frac{L}{G}, \frac{J}{G}, j - \lambda\right)$$

(in questo modo si ha $\dot{\lambda} = \frac{\partial \mathcal{H}_0}{\partial \Lambda} = \Omega$ ovvero $\lambda(t) = \Omega t$, mentre non ci interessa assolutamente l'evoluzione di Λ). Assumiamo poi (ma la cosa sarà verificata nella prossima sezione) che V_0 sia piccolo, e poniamo $\varepsilon = |V_0|^\infty$ (la sup-norma di V_0 nel dominio che stiamo considerando). Introduciamo lo sviluppo di Fourier di V_0 rispetto all'angolo $j - \lambda$

$$V_0 = \bar{V}\left(\frac{L}{G}, \frac{J}{G}\right) + \sum_{\nu \in \mathbb{Z} \setminus 0} \hat{V}_\nu\left(\frac{L}{G}, \frac{J}{G}\right) e^{i\nu(j-\lambda)},$$

e la funzione

$$\chi = - \sum_{\nu \in \mathbb{Z} \setminus 0} \frac{\hat{V}_\nu}{i\nu\Omega} e^{i\nu(j-\lambda)},$$

che è ben definita poiché $|\nu\Omega| \geq \Omega \neq 0$, e con una opportuna costante c soddisfa la stima

$$|\chi|^\infty \leq c |V_0|^\infty = c\varepsilon.$$

⁷⁸Si osservi che $(1 - L/G)^{|k|/2}$ non è analitica per $G = L$. In effetti questo problema è legato al fatto che le coordinate di Andoyer-Deprit non sono definite sulle rotazioni giroscopiche, cioè gli stati per cui $e_3 = m$ (gli angoli g ed l separatamente non sono definiti) e dunque non esiste una carta di coordinate di Andoyer-Deprit che sia anche un intorno di questi stati. In queste note ignoreremo il problema, che può essere risolto solo con un livello un po' più avanzato di conoscenza della geometria dell'insieme \mathcal{A} .

Consideriamo poi la trasformazione canonica

$$(G, L, J, \Lambda, g, l, j, \lambda) = \Phi_{\chi}^1(G', L', J', \Lambda', g', l', j', \lambda')$$

corrispondente al flusso al tempo 1 associato all' hamiltoniana χ . In particolare, siccome χ non dipende da g e l , si ha $G' = G$ e $L' = L$, mentre

$$|j' - j| = \left| \frac{\partial \chi}{\partial J} \right| = \mathcal{O}(\varepsilon), \quad |J' - J| = \left| \frac{\partial \chi}{\partial j} \right| = \mathcal{O}(\varepsilon),$$

e dunque la trasformazione è prossima all'identità.

La nuova funzione di Hamilton è ovviamente

$$\mathcal{H}'_0 = \mathcal{H}_0 \circ \Phi_{\chi}^1;$$

calcolando esplicitamente il primo ordine in ε si ottiene

$$\mathcal{H}'_0 = h + (V_0 + \{h, \chi\}) + \mathcal{O}(\varepsilon^2),$$

ove $h = \frac{G^2}{2I_1} + \frac{\eta L^2}{2I_1} + \Omega \Lambda$. Per la definizione stessa della χ risulta

$$V_0 + \{h, \chi\} = V_0 - \frac{\partial h}{\partial \Lambda} \frac{\partial \chi}{\partial \lambda} = V_0 - \Omega \frac{\partial \chi}{\partial \lambda} = \bar{V},$$

cosicché

$$\mathcal{H}'_0 = h(G', L', \Lambda') + \bar{V}\left(\frac{L'}{G'}, \frac{J'}{G'}\right) + \mathcal{O}(\varepsilon^2).$$

Il calcolo a questo punto è concluso: la frequenza associata a j' è

$$\dot{j}' = \frac{\partial \bar{V}}{\partial J} + \mathcal{O}(\varepsilon^2).$$

Con un minimo di riflessione si capisce che il contributo principale alla frequenza di precessione è

$$\alpha = \frac{\partial \bar{V}}{\partial J}\left(1, \frac{J_0}{G_0}\right) = \frac{1}{2\pi} \frac{\partial}{\partial J} \int_0^{2\pi} V\left(1, \frac{J_0}{G_0}, g, x\right) dx$$

(il termine $\mathcal{O}(\varepsilon^2)$, la sostituzione di J' con J_0 , di $j'(t)$ con $j(t)$ di L_0 con G_0 forniscono solo piccole correzioni). Per ottenere numeri confrontabili con le osservazioni sperimentali bisogna effettuare esplicitamente il calcolo di V in funzione delle coordinate di Andoyer–Deprit.

I.3 Il potenziale gravitazionale

Per semplificare l'espressione del potenziale sfrutteremo il fatto che il diametro equatoriale del pianeta R è piccolo rispetto alla distanza a dal Sole. Nel calcolo inoltre si potranno eliminare tutti i termini che sono costanti, cioè non dipendono dalle coordinate di Andoyer–Deprit. Dimostreremo innanzitutto la seguente formula:

$$-\sum_{i=1, N} \frac{\gamma M}{\|P_i - O\|} = \frac{3}{2} \frac{\gamma M}{a^3} \mathcal{J}_B \varrho \cdot \varrho + \mathcal{O}\left(\frac{R}{a}\right)^3 + \text{termini costanti},$$

ove $\mathcal{J}_B \varrho \cdot \varrho$ è il momento d'inerzia del corpo relativamente all'asse parallelo al vettore $\varrho = \frac{B-O}{a}$ e passante per B ; \mathcal{J}_B denota l'operatore di inerzia.

Dimostrazione di (I.3). Sia

$$x_i = \frac{P_i - O}{a} ,$$

cosicché $\|x_i\| \leq R/a$ per ogni i . Utilizzando lo sviluppo in serie di Mac Laurin della funzione $1/\sqrt{1+\beta}$ si deduce che

$$\begin{aligned} \frac{1}{\|P_i - O\|} &= \frac{1}{\|(P_i - B) + (B - O)\|} = \frac{1}{a\|\varrho + x_i\|} = \frac{1}{a\sqrt{1 + (\|x_i\|^2 + 2\varrho \cdot x_i)}} = \\ &= \frac{1}{a} \left(1 - \varrho \cdot x_i - \frac{1}{2}\|x_i\|^2 + \frac{3}{2}(\varrho \cdot x_i)^2 \right) + \mathcal{O}\left(\frac{R}{a}\right)^3 , \end{aligned}$$

che a meno dei termini costanti 1 e $\|x_i\|^2$ è

$$\frac{1}{a} \left(-\varrho \cdot x_i + \frac{3}{2}(\varrho \cdot x_i)^2 \right) + \mathcal{O}\left(\frac{R}{a}\right)^3 .$$

Moltiplicando per m_i e sommando su tutti i punti, il contributo dovuto a $\varrho \cdot x_i$ è nullo, perché per le proprietà del baricentro risulta

$$\varrho \cdot \left(\sum_i m_i x_i \right) = 0 .$$

Dunque, a meno di termini costanti e di termini di ordine $(R/a)^3$ il potenziale gravitazionale è dato dalla espressione

$$-\frac{3}{2} \frac{\gamma M}{a^3} \sum_i m_i (\varrho \cdot (P_i - B))^2 .$$

A questo punto è facile verificare che

$$\sum_i m_i (\varrho \cdot (P_i - B))^2 = \sum_i m_i \|P_i - B\|^2 - \mathcal{J}_B \varrho \cdot \varrho ,$$

ove $\sum_i m_i \|P_i - B\|^2$ non dipende dalla configurazione e si può trascurare, mentre $\mathcal{J}_B \varrho \cdot \varrho$ è il momento d'inerzia del corpo relativamente all'asse parallelo a ϱ e passante per B . Abbiamo così dimostrato la validità di (I.3).

Il conto poi prosegue nel modo seguente. Se indichiamo con $\varrho_1, \varrho_2, \varrho_3$ le componenti del vettore ϱ nel riferimento solidale principale, e ricordiamo che ϱ è un vettore, si ottiene

$$\mathcal{J}_B \varrho \cdot \varrho = I_1(\varrho_1^2 + \varrho_2^2) + I_3 \varrho_3^2 = I_1 - (I_1 - I_3) \varrho_3^2 = I_1 - (I_1 - I_3)(\varrho \cdot e_3)^2 ;$$

il termine I_1 può essere trascurato in quanto costante. Calcoliamo $(\varrho \cdot e_3)^2$ in funzione delle coordinate di Andoyer-Deprit valutandolo per $G = L$: in questo caso il momento angolare è allineato all'asse di simmetria, e le componenti di e_3 nel sistema di riferimento e_x, e_y, e_z sono

$$e_3 = \left(\sqrt{1 - \frac{J^2}{G^2}} \sin j, -\sqrt{1 - \frac{J^2}{G^2}} \cos j, \frac{J}{G} \right) .$$

Le componenti di ϱ sono invece

$$\varrho = (\cos \Omega t, \sin \Omega t, 0) ,$$

cosicché

$$(\varrho \cdot e_3)^2 = \left(1 - \frac{J^2}{G^2}\right) \sin^2(j - \Omega t)$$

e dunque⁷⁹

$$\bar{V} = -\frac{1}{2\pi} \frac{3\gamma M}{2a^3} \left(1 - \frac{J^2}{G^2}\right) (I_1 - I_3) \int_0^{2\pi} \sin^2(x) dx = -\frac{3\gamma M}{4a^3} (I_1 - I_3) \left(1 - \frac{J^2}{G^2}\right).$$

La frequenza della precessione è allora

$$\alpha = \frac{\partial \bar{V}}{\partial J} = \frac{3\gamma M}{2a^3} (I_1 - I_3) \frac{J}{G^2}.$$

L'espressione di α diviene trasparente grazie a tre sostituzioni: $\frac{\gamma M}{a^3} = \Omega^2$, ove Ω denota la frequenza del moto di rivoluzione; $J = G \cos(i)$, ove i è l'inclinazione dell'asse di simmetria rispetto alla verticale dell'orbita; e infine, data l'approssimazione $G = L$, $G = I_3 \omega_G$, ove ω_G è il modulo della velocità angolare del corpo rigido, ovvero la frequenza associata al giorno. Si ottiene così

$$\alpha = \frac{3}{2} \frac{\Omega}{\omega_G} \frac{I_1 - I_3}{I_3} \cos(i) \Omega,$$

nota come formula di precessione classica.

I.4 La precessione della Terra

Per calcolare il valore della precessione della Terra è sufficiente sostituire nella formula precedente i valori numerici, che per la Terra sono $i \simeq 23^\circ$ e $(I_1 - I_3)/I_3 = -1/298,25$; si ottiene $\alpha \simeq -1.26 \cdot 10^{-5} \Omega$, cui corrisponde un periodo di circa 79000 anni. Tale periodo è circa il triplo di quello osservato sperimentalmente. Questo perché esso è stato calcolato tenendo conto solamente del contributo del Sole alla precessione, mentre in realtà è importante considerare anche il contributo α' dovuto alla Luna. In questo caso ripetendo il ragionamento fatto per il calcolo di α si otterrebbe

$$\alpha' = \frac{3\gamma M'}{2a'^3} (I_1 - I_3) \frac{J}{G^2},$$

ove M' è la massa che causa la precessione, cioè la massa della Luna, ed a' è la distanza Terra-Luna. Indicando con m la massa della Terra e con ω_M la frequenza del moto lunare (2π /mese) si ha

$$\frac{\gamma M'}{a'^3} = \frac{\gamma m M'}{a'^3 m} = \frac{M'}{m} \omega_M^2,$$

cosicché

$$\alpha' = \frac{3}{2} \frac{M'}{m} \frac{\omega_M}{\omega_G} \frac{I_1 - I_3}{I_3} \cos(i') \omega_M;$$

sostituendo i valori numerici si ottiene $\alpha' \simeq -2.75 \cdot 10^{-5} \Omega$, e dunque la frequenza di precessione complessiva è

$$\alpha + \alpha' \simeq -4.01 \cdot 10^{-5} \Omega$$

che corrisponde ad un periodo di circa 25000 anni, corretto entro il 4% rispetto al valore sperimentale.

⁷⁹Si confronti questa funzione con l'approssimazione integrabile e si verifichi che effettivamente può essere considerata una perturbazione.

I.5 Oltre il modello classico della precessione

Il modello classico della precessione spiegato in queste note è sufficiente per descrivere anche in modo quantitativo il fenomeno della precessione degli equinozi, osservato fin dall'antichità. Diversi quesiti che riguardano la modalità di rotazione dei pianeti sono tuttavia ancora di attualità scientifica e sono correlati ai molti quesiti ancora aperti riguardanti la formazione e l'evoluzione del nostro sistema Solare (la rotazione retrograda di Venere, l'instabilità della rotazione di Marte, le risonanze spin-orbita dei satelliti). Come già osservato, per poter affrontare in modo corretto questi problemi bisogna descrivere la rotazione dei pianeti su tempi dell'ordine della vita stessa del nostro sistema solare, e per farlo bisogna andare oltre il modello classico di precessione. La modifica principale al modello consiste nel sostituire l'orbita circolare del pianeta con la sua vera orbita, che a causa delle interazioni mutue fra i pianeti, è suscettibile di variazioni quasi-periodiche su tempi scala di 10^4 - 10^5 anni, e anche di variazioni di tipo 'caotico' su tempi scala dell'ordine di 10^6 - 10^7 anni. Tali variazioni dell'orbita possono indurre non solo una precessione dell'asse di rotazione ma anche una lenta variazione dell'angolo che tale asse forma rispetto alla verticale. Ad esempio, per la stabilità della rotazione terrestre su questo lungo periodo è fondamentale la presenza della Luna: se la Luna non ci fosse, la frequenza di precessione sarebbe solo α (e non $\alpha + \alpha'$), e a ciò corrisponderebbe nel modello completo una grande instabilità dell'angolo dell'asse di rotazione rispetto alla verticale su tempi dell'ordine di qualche milione di anni. Una tale instabilità avrebbe influenzato in modo drammatico il clima del nostro pianeta, e quindi anche l'evoluzione delle specie viventi su di esso.

J Oltre i rotatori: un passo perturbativo per generici sistemi convessi

Scopo di questa appendice è di generalizzare lo studio dei sistemi anisocroni rimuovendo progressivamente le peculiarità e le particolari ipotesi che caratterizzano il modello dei rotatori. Precisamente, studieremo hamiltoniane della forma generica

$$H_\varepsilon(I, \varphi) = h(I) + \varepsilon f(I, \varphi) , \quad (I, \varphi) \in B_R \times \mathbb{T}^n ,$$

generalizzando progressivamente la specialissima f dei rotatori fino a una f (analitica) qualsiasi, mentre per quanto riguarda h la generalizzazione, più modesta, sarà ad h "quasi convessa". La quasi convessità è una piccola generalizzazione della convessità, che come vedremo risulta naturale nel nostro contesto; una funzione h si dice quasi convessa in B , se per ogni $I \in B$ si ha che

$$\begin{cases} h' \cdot v = 0 \\ h''(I)v \cdot v = 0 \end{cases} \Rightarrow v = 0 ,$$

ove si è denotato

$$h' = \frac{\partial h}{\partial I} = \omega , \quad h'' = \frac{\partial^2 h}{\partial I \partial I} ;$$

in sostanza si chiede la convessità di h (o di $-h$) ristretta al piano ortogonale a h' . Un esempio di funzione quasi convessa ma non convessa è⁸⁰ $h(I_1, I_2, I_3) = I_1^2 + I_2^2 + I_3$. In queste ipotesi riusciremo a garantire la stabilità di tutte le azioni fino a tempi $1/\varepsilon$.

⁸⁰Si osservi che con l'ipotesi di quasi convessità si prendono dentro anche alcune funzioni che non soddisfano la condizione (4.32) di stretta non degenerazione dell'hessiana.

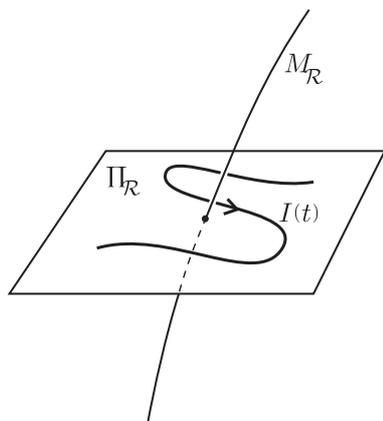


Figura 53: Il moto delle azioni è appiattito sul piano di scivolamento rapido $\Pi_{\mathcal{R}}$.

A. Perturbazioni con un numero finito di componenti di Fourier; il caso non risonante. Manteniamo qui provvisoriamente la restrizione che f sia un polinomio di Fourier, diciamo

$$f(I, \varphi) = \sum_{|k| \leq N} f_k(I) e^{ik \cdot \varphi} ,$$

ove N è una costante detta “taglio (*cut-off*) ultravioletto”; rimuoviamo invece le altre peculiarità di f (interazione a primi vicini ecc.), di cui abbiamo approfittato studiando il modello dei rotatori.

Consideriamo innanzitutto una regione dello spazio delle azioni lontana dalle risonanze, precisamente nella quale risulti

$$|k \cdot \omega(I)| \geq c > 0 \quad \text{per ogni } k \in \mathbb{Z}^n: 0 < |k| \leq N . \tag{J.1}$$

Con un passo perturbativo fatto con

$$\chi(I, \varphi) = \sum_{0 < |k| \leq N} \frac{\hat{f}_k(I)}{ik \cdot \omega(I)} e^{ik \cdot \varphi} \tag{J.2}$$

si possono eliminare tutte le armoniche di f , ottenendo

$$\tilde{H}(\tilde{I}, \tilde{\varphi}) = h(\tilde{I}) + \varepsilon g(\tilde{I}) + \varepsilon^2 \tilde{f}(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^3) \tag{J.3}$$

con

$$g = \langle f \rangle , \quad \tilde{f} = \frac{1}{2} \{f + \langle f \rangle, \chi\} .$$

La stabilità di tutte le nuove azioni (con oscillazioni $O(\varepsilon)$) per tempi $1/\varepsilon$ è chiaramente assicurata, e come sempre, la stabilità delle nuove azioni implica quella delle vecchie che restano ε -vicine alle nuove. Su h , per il momento, non abbiamo dovuto fare ipotesi.

Come si è visto per i rotatori, avvicinandosi in frequenza ai piani risonanti $k \cdot \omega = 0$, in azione alle varietà risonanti $k \cdot \omega(I) = 0$, $|k| \leq N$, la stabilità si indebolisce, fino a che, in un intorno $\mathcal{O}(\sqrt{\varepsilon})$ delle varietà risonanti, il passo perturbativo non diviene inutile: in tale intorno infatti uno o più divisori $k \cdot \omega(I)$ diventano di ordine $\sqrt{\varepsilon}$; eseguendo allora la parentesi di Poisson $\{f + \bar{f}, \chi\}$ che dà \tilde{f} , compaiono obbligatoriamente i quadrati di queste quantità, e esattamente come per i rotatori ciò rende illusorio il guadagno di un fattore ε di fronte a \tilde{f} .

B. La costruzione risonante; il ruolo della quasi-convessità di h . Veniamo allora alla costruzione risonante. Sia $\mathcal{R} \subset \mathbb{Z}^N$ un reticolo di dimensione d , $1 \leq d < n$, che ammetta una base $k^{(1)}, \dots, k^{(d)}$ con $|k^{(s)}| \leq N$ per ciascun s . La varietà di codimensione d

$$\mathcal{M}_{\mathcal{R}} = \{I \in B : k \cdot \omega(I) = 0 \quad \forall k \in \mathcal{R}\}$$

è detta *varietà risonante* associata a \mathcal{R} . In un intorno di raggio $\sqrt{\varepsilon}$ di $\mathcal{M}_{\mathcal{R}}$, se risulta $k \cdot \omega(I) \geq c > 0$ per $k \notin \mathcal{R}$, $|k| \leq N$, con un passo perturbativo si eliminano tutte le armoniche di f tranne quelle con $k \in \mathcal{R}$, ottenendo

$$\tilde{H}(\tilde{I}, \tilde{\varphi}) = h(\tilde{I}) + \varepsilon g(\tilde{I}, \tilde{\varphi}) + \varepsilon^2 \tilde{f}(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^3) ,$$

con \tilde{f} simile a sopra ma questa volta con

$$g(\tilde{I}, \tilde{\varphi}) = \sum_{k \in \mathcal{R}} \hat{f}_k(\tilde{I}) e^{ik \cdot \tilde{\varphi}} .$$

Trascuriamo il termine $\mathcal{O}(\varepsilon^2)$ e cerchiamo di capire come g muove le azioni. Per alleggerire la notazione aboliamo la tilde sulle variabili: tanto come sempre i risultati sulle nuove azioni si trasferiscono banalmente alle vecchie, a meno di una deformazione di ordine ε .

Dalle equazioni di Hamilton segue che \dot{I} è una combinazione di vettori $k \in \mathcal{R}$, precisamente

$$\dot{I} = \varepsilon \sum_{k \in \mathcal{R}} C_k(I, \varphi) k , \quad C_k(I, \varphi) = i \hat{f}_k(I) e^{ik \cdot \varphi} .$$

Pertanto la dinamica nello spazio delle azioni è schiacciata sul piano $\Pi_{\mathcal{R}}(I^o)$ parallelo a \mathcal{R} passante per il dato iniziale I^o , detto piano di scivolamento rapido⁸¹ (di *fast drift*); si veda la figura 53.

Perché il moto su tale piano risulti confinato, e con esso tutte le azioni, occorre un'ipotesi su h , e la più semplice che conduce a risultati interessanti è la quasi convessità introdotta sopra. La quasi convessità ha due conseguenze importanti.

- (i) In ogni punto $I^* \in \mathcal{M}_{\mathcal{R}}$, il piano $\Pi_{\mathcal{R}}(I^*)$ e la varietà $\mathcal{M}_{\mathcal{R}}$ sono trasversi (come implicitamente assunto in figura). Infatti, se non lo fossero, allora dovrebbe esistere un vettore v parallelo al piano, dunque

$$v = \sum_{k \in \mathcal{R}} v_k k ,$$

che è anche tangente a $\mathcal{M}_{\mathcal{R}}$ in I^* , e pertanto è ortogonale a qualunque vettore u ortogonale a $\mathcal{M}_{\mathcal{R}}$ in I^* . Ma poiché $\omega(I) \cdot v = 0$ identicamente su $\mathcal{M}_{\mathcal{R}}$, il gradiente $u(I) = \nabla(\omega(I) \cdot v) = h''(I)v$ è ortogonale a $\mathcal{M}_{\mathcal{R}}$ in I ; in particolare in I^* si avrebbe allora

$$h''(I^*)v \cdot v = 0 ,$$

ma poiché risulta anche $\omega(I^*) \cdot v = 0$, resta violata la quasi convessità. La trasversalità implica anche l'unicità dell'intersezione di $\Pi_{\mathcal{R}}$ con $\mathcal{M}_{\mathcal{R}}$.

- (ii) La funzione h , ristretta a $\Pi_{\mathcal{R}}$, ha un minimo o un massimo di secondo ordine in I^* . Infatti, scriviamo

$$h(I^* + v) = h(I^*) + \omega(I^*) \cdot v + \frac{1}{2} h''(I^*)v \cdot v + \mathcal{O}(\|v\|^3) ;$$

per v parallelo a $\Pi_{\mathcal{R}}(I^*)$ il termine di ordine uno si annulla, mentre per l'ipotesi di quasi convessità il segno della parte quadratica è definito.

⁸¹Rapido nel senso che la velocità \dot{I} è irrimediabilmente $\mathcal{O}(\varepsilon)$, senza speranza di dimostrare che sia più lenta.

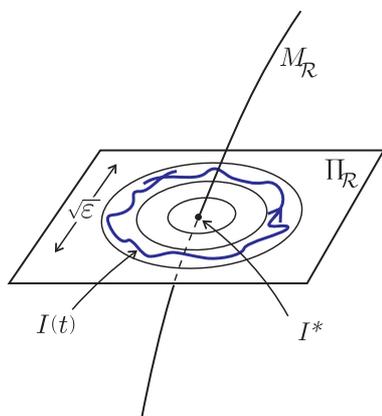


Figura 54: *Il confinamento delle azioni nel piano di scivolamento rapido per sistemi quasi convessi.*

La conservazione dell'energia garantisce allora il confinamento. Infatti, la “parte cinetica” h nell'hamiltoniana (J.3), ristretta a $\Pi_{\mathcal{R}}$, ha un massimo o un minimo (di secondo ordine) in I^* , e corrispondentemente le sue superfici di livello, per distanze piccole da I^* — non si dimentichi che siamo in un intorno di I^* di raggio $\mathcal{O}(\sqrt{\varepsilon})$ — hanno la struttura di ellissoidi concentrici attorno a I^* . Poiché la “parte potenziale” $\varepsilon g + \varepsilon^2 \tilde{f}$ ha oscillazioni limitate da $(\text{cost})\varepsilon$, il moto delle azioni è necessariamente confinato tra due superfici di livello vicine e si estende su una scala di ordine $\sqrt{\varepsilon}$; si veda la figura 54. La situazione è questa fintantoché il resto $\varepsilon^2 \tilde{f}$ non sposta in maniera apprezzabile le azioni in direzione trasversa al piano di scivolamento rapido; nel nostro caso possiamo garantirlo per tempi di ordine ε^{-1} (in tale intervallo lo spostamento trasverso è $\mathcal{O}(\varepsilon) \ll \sqrt{\varepsilon}$, per cui il moto è effettivamente appiattito su $\Pi_{\mathcal{R}}$). Un lavoro analitico a ordini più elevati garantirebbe un confinamento per tempi più lunghi, *ma non un confinamento migliore.*

Se in $\Pi_{\mathcal{R}}$ al posto della struttura ellittica attorno a I^* ci fosse una struttura iperbolica, non potremmo più garantire il confinamento delle azioni. E' il caso dell'hamiltoniana $h = \frac{1}{2}(I_1^2 - I_2^2)$ incontrata nel paragrafo 4.3 (esempio *iv*) che perde la quasi convessità sulle rette $I_2 = \pm I_1$ e corrispondentemente ammette moti che attraversano distanze $\mathcal{O}(1)$ su scala di tempo $1/\varepsilon$.

C. Variabili lente e veloci; riscalamanti. Consideriamo ora la risonanza particolare di ordine d associata al reticolo

$$\mathcal{R} = \{k \in \mathbb{Z}^n : k = (k_1, \dots, k_d, 0, \dots, 0)\} ; \tag{J.4}$$

$\mathcal{M}_{\mathcal{R}}$ corrispondentemente è definita da

$$\omega_j(I) = 0 , \quad j = 1, \dots, d .$$

La forma normale adattata a questa risonanza è della forma

$$\tilde{H}(\tilde{I}, \tilde{\varphi}) = h(\tilde{I}) + \varepsilon g(\tilde{I}, \tilde{\varphi}_1, \dots, \tilde{\varphi}_d) + \varepsilon^2 \tilde{f}(\tilde{I}, \tilde{\varphi}) ,$$

e il piano di scivolamento rapido $\Pi_{\mathcal{R}}(I)$ è il piano in cui restano costanti I_{d+1}, \dots, I_n . Distinguiamo allora le variabili in “lente” (S, σ) e “veloci” (F, ψ) ponendo

$$(S, \sigma) = (\tilde{I}_1, \dots, \tilde{I}_d, \tilde{\varphi}_1, \dots, \tilde{\varphi}_d) , \quad (F, \psi) = (\tilde{I}_{d+1}, \dots, \tilde{I}_n, \tilde{\varphi}_{d+1}, \dots, \tilde{\varphi}_n)$$

(gli aggettivi “lento” e “veloce” si riferiscono agli angoli, non alle azioni, o per meglio dire alla scala di tempo su cui il sottosistema evolve). L'hamiltoniana si scrive allora

$$\tilde{H}(S, F, \sigma, \psi) = h(S, F) + \varepsilon g(S, F, \sigma) + \mathcal{O}(\varepsilon^2) .$$

Fintantoché il termine $\mathcal{O}(\varepsilon^2)$ si può trascurare, le azioni F sono parametri costanti, e il moto delle variabili (S, σ) è determinato dall'hamiltoniana scritta sopra, privata del resto, interpretata come sistema a soli d gradi di libertà in cui le F sono parametri:

$$H_F(S, \sigma) = h_F(S) + \varepsilon g_F(S, \sigma) ,$$

con ovvio significato dei simboli h_F, g_F . Per ogni F , sia $I^* = (S^*(F), F)$; conviene prendere S^* come origine, tramite la traslazione $S = S^* + s$, e inoltre, poiché il dominio che ci interessa ha diametro $\mathcal{O}(\sqrt{\varepsilon})$, conviene introdurre, *esattamente come abbiamo fatto per i rotatori*, il riscaldamento

$$s = \sqrt{\varepsilon} s' , \quad t = \frac{t'}{\sqrt{\varepsilon}} , \quad H'_F = \varepsilon^{-1} H_F .$$

Si trova

$$H'_F(s', \sigma) = \frac{1}{\varepsilon} h(S^*, F) + \frac{1}{\sqrt{\varepsilon}} \omega(S^*, F) \cdot s' + \frac{1}{2} A_F s' \cdot s' + g_F(S^*, \sigma) + \mathcal{O}(\sqrt{\varepsilon}) ,$$

ove $A_F = h''(S^*, F)$. Rimosso il termine costante $\frac{1}{\varepsilon} h(S^*, F)$, e osservato che $\omega(S^*, F) \cdot s' = 0$, si resta con

$$H'_F(s', \sigma) = \frac{1}{2} A_F s' \cdot s' + g_F(\sigma) + \mathcal{O}(\sqrt{\varepsilon}) .$$

Questo è un sistema fortemente accoppiato, che in nessun modo è vicino a un sistema integrabile, e ci si devono aspettare, genericamente, moti complessi su scala di tempo e di azione di ordine uno, cioè indipendenti da ε . Moti limitati naturalmente, perchè A_F è definita positiva, o negativa, mentre g_F è limitata. A questi moti corrispondono, nelle variabili originarie non riscalate, analoghi moti non banali che si svolgono su scala di azione di ordine $\sqrt{\varepsilon}$ e su scala di tempo di ordine $1/\sqrt{\varepsilon}$ (sia per S che per σ), mentre nel contempo le azioni F sono stabili e gli angoli ψ ruotano regolarmente con velocità di ordine uno.

Tutte queste considerazioni sono state sviluppate in connessione a una risonanza molto particolare: ma in realtà quello che abbiamo trovato è generico perché, come abbiamo già avuto modo di ricordare, qualunque reticolo risonante di dimensione d tramite una trasformazione canonica lineare si porta alla forma (J.4).

In conclusione, dentro a una qualunque risonanza di ordine d :

- (i) Il movimento delle azioni è appiattito sul piano di scivolamento rapido, ed è ivi confinato in una regione di diametro $\mathcal{O}(\sqrt{\varepsilon})$;
- (ii) d gradi di libertà scelti opportunamente possono compiere moti complessi, benché localizzati, su una scala di tempo di ordine $1/\sqrt{\varepsilon}$, mentre gli altri gradi di libertà si muovono regolarmente: le azioni restano costanti, gli angoli avanzano regolarmente.
- (iii) Il quadro vale (almeno) per tempi di ordine $1/\varepsilon$.

D. Perturbazioni con infinite componenti di Fourier. Vediamo ora rapidamente come si generalizzano i risultati al caso di una funzione analitica f generica, con serie di Fourier piena. L'idea centrale di cui approfittiamo è quella introdotta da Arnol'd nel 1961, nella sua dimostrazione del teorema KAM: se introduciamo nella serie un taglio ultravioletto N e decomponiamo f nella sua "parte infrarossa" $f^{\leq N}$ e "parte ultravioletta" $f^{> N}$:

$$f = f^{\leq N} + f^{> N} \tag{J.5}$$

con

$$f^{\leq N}(I, \varphi) = \sum_{|k| \leq N} \hat{f}_k(I) e^{ik \cdot \varphi}, \quad f^{> N} = \sum_{|k| < N} \hat{f}_k(I) e^{ik \cdot \varphi}, \quad (\text{J.6})$$

allora la parte ultravioletta è esponenzialmente piccola in N . Precisamente:

LEMMA 53 *Se f è analitica e limitata nel dominio \mathcal{D}_ϱ , allora le singole componenti di Fourier $\hat{f}(I)$ sono stimate per ogni I da*

$$|\hat{f}_k(I)| \leq \|f\|_\varrho e^{-|k|\varrho_\varphi}; \quad (\text{J.7})$$

corrispondentemente la parte ultravioletta $f^{> N}$ è stimata da

$$\|f^{> N}\|_{\frac{1}{2}\varrho} \leq C e^{-\frac{1}{4}N\varrho_\varphi} \|f\|_\varrho, \quad (\text{J.8})$$

ove $\|\cdot\|$ denota la sup-norma, con a pedice l'estensione complessa del dominio su cui la norma è calcolata, mentre C è una opportuna costante dipendente da n .

DIMOSTRAZIONE. Dall'espressione dei coefficienti di Fourier

$$\hat{f}(I) = \frac{1}{(2\pi)^n} \int_0^{2\pi} \cdots \int_0^{2\pi} f(I, \varphi) e^{-ik \cdot \varphi} d\varphi_1 \cdots d\varphi_n,$$

alzando il cammino di integrazione del j -esimo angolo φ_j a $\text{Im } \varphi_j = \pm \varrho_\varphi$, con segno opposto al segno di k , si ottiene

$$\hat{f}_k(I) = \frac{1}{(2\pi)^n} \int_0^{2\pi} \cdots \int_0^{2\pi} f(I, \varphi' \pm i\varrho_\varphi) e^{-|k|\varrho_\varphi} e^{ik \cdot \varphi'} d\varphi'_1 \cdots d\varphi'_n,$$

con cammino di integrazione reale (come sempre, i tratti verticali in 0 e in 2π si compensano). La (J.7) è immediata.

Dalla definizione (J.6) di $f^{> N}$ si trova allora subito, per ogni ϱ' con $\varrho'_I \leq \varrho_I$ e $\varrho'_\varphi < \varrho_\varphi$,

$$\begin{aligned} \|f^{> N}\|_{\varrho'} &\leq \sum_{|k| > N} \|f\|_\varrho e^{-|k|\varrho_\varphi} e^{|k|\varrho'_\varphi} \\ &\leq \|f\|_\varrho \sum_{m > N} 2^n M(m, n) e^{-m(\varrho_\varphi - \varrho'_\varphi)}, \end{aligned}$$

ove $M(m, n)$ denota il numero di n -ple (k_1, \dots, k_n) con i $k_j \geq 0$ tali che $\sum_j k_j = m$ (il fattore 2^n tiene conto ad abbondanza dei $k_j < 0$). Per il coefficiente $M(m, n)$ si trova immediatamente la stima $M(m, n) \leq (m+1)^{n-1}$ (fissati ad arbitrio i primi $n-1$ dei k_j tra 0 e m l'ultimo ha al più un valore); il valore esatto, un po' migliore, risulta essere il binomiale $\binom{m+n-1}{m}$. Si adopera (per esempio) metà dell'esponenziale per compensare la potenza (cioè si usa il fatto che $M(m, n)e^{-\frac{1}{2}N(\varrho_\varphi - \varrho'_\varphi)}$ è limitato) giungendo a

$$\|f^{> N}\|_{\varrho'} \leq C e^{-\frac{1}{2}N(\varrho_\varphi - \varrho'_\varphi)} \|f\|_\varrho;$$

la (J.8) è questa espressione scritta, per semplicità, con $\varrho' = \frac{1}{2}\varrho$. □

L'utilizzo del lemma è elementare: se prendiamo N tale che $e^{-\varrho_\varphi N/4} = \varepsilon$, cioè $N(\varepsilon) = (4/\varrho_\varphi) \log \varepsilon^{-1}$ (non c'è bisogno che N sia intero), allora la parte ultravioletta della perturbazione, pur avendo un solo ε a fattore, è già piccola come ε^2 , e non c'è bisogno di lavorarci sopra. In questo

modo si riduce il problema, in buona sostanza, al caso visto sopra di uno sviluppo di Fourier finito: a meno della complicazione (sottile, insidiosa) che diminuendo ε si alza N e compaiono nuove armoniche, per cui dove sembrava esserci, poniamo, una risonanza di un certo ordine d , rimpicciolendo ε può comparire una risonanza di ordine più elevato. Non entreremo qui in questo problema.

E. Oltre la quasi convessità. La quasi convessità è solo la più semplice delle proprietà che conducono al confinamento delle azioni nel piano di scivolamento rapido. Come ha mostrato Nekhoroshev, la quasi convessità può essere sostituita da una condizione molto più generale (e generica, in un senso opportuno) chiamata da lui *steepness* (ripidità?), seppure a prezzo di un confinamento peggiore. Ad esempio, un indebolimento possibile della quasi convessità, che a sua volta è un caso particolare di *steepness*, è la “non degenerazione del 3-jet”:

$$\begin{cases} h'v = 0 \\ h''vv = 0 \\ h'''vvv = 0 \end{cases} \Rightarrow v = 0 ;$$

ma l'ulteriore indebolimento alla non degenerazione del 4-jet invece non funziona, e questo fa comprendere quanto sia delicata la materia. Non ci addentreremo ulteriormente in quest'argomento, rinunciando anche a dare anche la definizione, assai complessa, di *steepness*. E' un argomento ancora non del tutto capito e oggetto oggi di ricerca.