

### III. Introduzione alla teoria delle perturbazioni per sistemi Hamiltoniani

(2001/02)

La teoria delle perturbazioni per i sistemi hamiltoniani si è rivelata, dagli anni '50 in poi, particolarmente vitale e potente. Il cosiddetto *teorema KAM* (Kolmogorov 1954; Arnol'd 1962; Moser 1961) ha consentito di chiudere problemi antichi di uno o due secoli, come la stabilità del più semplice problema a tre corpi, e soprattutto ha introdotto problemi nuovi e un modo di lavorare nuovo. Poco dopo il *teorema di Nekhoroshev* (1974) a sua volta ha introdotto un punto di vista nuovo nell'analisi della stabilità di un sistema dinamico, ponendo il problema della stabilità dei moti per tempi lunghi (diciamo  $t \sim e^{1/\varepsilon}$ , se  $\varepsilon$  è un naturale parametro perturbativo) anche se non infiniti, e si è rivelato essere particolarmente fecondo per le applicazioni fisiche. Questo capitolo non ha la pretesa di offrire un panorama completo della moderna teoria hamiltoniana delle perturbazioni, ma (al pari dei capitoli precedenti) si prefigge di introdurre gli strumenti di base della teoria, illustrando i problemi, le difficoltà principali e le idee per superarle, così da condurre il lettore almeno alla soglia dei risultati importanti.

Il lavoro perturbativo è principalmente lavoro di analisi da compiersi in coordinate, per il quale basterebbe la nozione di sistema hamiltoniano locale (quella che si impara solitamente nei corsi di meccanica razionale). Pochi problemi<sup>1)</sup> richiedono in modo sostanziale un punto di vista geometrico, ovvero la nozione di sistema hamiltoniano su una varietà. Ma è opinione di chi scrive che, se non si ha almeno un'idea della formulazione "in grande" o geometrica, il procedere della teoria perturbativa, a cominciare dal teorema di Liouville-Arnol'd che ne costituisce il piedestallo e dalla nozione di trasformazione canonica che ne è lo strumento principale, può risultare oscuro. Per questo nel primo paragrafo, dovendo richiamare le nozioni di base sui sistemi hamiltoniani e la nozione di trasformazione canonica, prendiamo volentieri, benché provvisoriamente, il punto di vista globale.

#### 1. Il formalismo hamiltoniano

In questo paragrafo ricordiamo alcune nozioni di base sui sistemi hamiltoniani. Premettiamo pochi richiami di carattere più generale su forme e campi vettoriali su una varietà, limitandoci agli oggetti di immediato interesse per la meccanica hamiltoniana, ovvero a

---

<sup>1)</sup> Tra questi, segnatamente, il problema classico della stabilità delle rotazioni rigide veloci (in particolare dei moti giroscopici) nell'ambito della teoria di Nekhoroshev.

funzioni reali, campi vettoriali e relative derivate di Lie, 1 e 2-forme esterne. Come regola, in questo paragrafo i punti di  $\mathbb{R}^n$ , e in generale le  $n$ -ple (le matrici) rappresentative di vettori e 1-forme (2-forme) saranno (doppiamente) sottolineati, per distinguerli dai corrispondenti oggetti intrinseci; più avanti, quando non vi sarà possibilità di confusione e lavoreremo principalmente in coordinate, per alleggerire la notazione le sottolineature saranno abolite. Anche per le funzioni useremo qui, provvisoriamente, notazioni abbastanza rigide, precisamente lettere maiuscole in corsivo per le funzioni intrinseche e lettere minuscole per le rappresentative locali.

### 1.1. Funzioni, campi vettoriali, forme

#### Coordinate locali.

Sia  $M$  una varietà differenziabile  $n$ -dimensionale. Denoteremo le sue carte con  $(U, \varphi)$ , ove  $U \subset M$  e  $\varphi : U \rightarrow V$ ,  $V$  aperto di  $\mathbb{R}^n$ . Ogni carta si può interpretare come sistema di coordinate locali; se  $\varphi$  invia  $x \in U$  in  $\underline{x} = (x_1, \dots, x_n) \in V \subset \mathbb{R}^n$ , 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  $(U, \varphi)$  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

$$\psi := \varphi \circ \tilde{\varphi}^{-1}, \quad \psi : \tilde{V}_0 \rightarrow V_0$$

si può a sua volta interpretare come cambiamento locale di coordinate. La classe di regolarità delle funzioni di transizione è per definizione la classe di regolarità di  $M$ . Per tutti gli oggetti geometrici (funzioni, campi vettoriali, 1 e 2-forme) che introdurremo, la classe di regolarità è, per definizione, quella dei loro rappresentativi locali. Per semplicità, e senza farne esplicita menzione, supporremo sempre tutto di classe  $C^\infty$ .

#### 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  $(U, \varphi)$  resta definita una *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

$$\tilde{f} = f \circ \psi, \quad \text{ovvero} \quad \tilde{f}(\tilde{x}) = f(\psi(\tilde{x})); \quad (1.1)$$

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 (1.1), resta definita una funzione intrinseca  $\mathcal{F} : M \rightarrow \mathbb{R}$ . La funzione  $\tilde{f}$  è anche detta *pull-back* di  $f$  sotto  $\psi$  ( $\psi$  va da  $\tilde{V}$  a  $V$ ; il *pull-back* porta funzioni  $f$  su  $V$  "indietro" a funzioni  $\tilde{f}$  su  $\tilde{V}$ ).

#### Campi vettoriali.

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  $(U, \varphi)$ , e per ogni  $x \in U$ , resta definita una base in  $T_x M$  formata dai vettori tangenti alle linee coordinate, usualmente denotati  $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ ; 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  $T_x M$  a  $x \in U$  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  $\mathbb{R}^n$ . Per ogni campo vettoriale  $X$  su  $M$ , e per ogni carta  $(U, \varphi)$ , 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  $X_i \circ \varphi$ );  $\underline{X}$  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  $\psi$  dall'una all'altra come sopra, allora si ha la legge di trasformazione

$$\tilde{\underline{X}}(\tilde{\underline{x}}) = J^{-1}(\tilde{\underline{x}}) \underline{X}(\psi(\tilde{\underline{x}})) , \quad J := \left( \frac{\partial \psi_i}{\partial \tilde{x}_j} \right) . \quad (1.2)$$

Viceversa, se una famiglia di campi locali soddisfa questa relazione, allora resta definito un campo  $X$  sulla varietà. Il campo  $\tilde{\underline{X}}$  definito dalla (1.2) si dice *pull-back*<sup>2)</sup> 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).

### Derivata di Lie di funzioni.

Dato un campo vettoriale  $X$  su  $M$  restano definite un'equazione differenziale sulla varietà e un'equazione differenziale locale in ogni carta, 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 pertanto resta definita una soluzione in grande su  $M$ , detta *flusso* del

<sup>2)</sup> La definizione di *pull-back* di un campo è ben data grazie al fatto che  $\psi$  è un diffeomorfismo, per cui  $J$  è invertibile.

campo  $X$  e denotata  $\Phi_X^t$ . Se  $M$  è compatta, o comunque se  $\Phi_X^t$  è prolungabile per  $t \in \mathbb{R}$  (cosa che supporremo sempre), allora  $\{\Phi_X^t, t \in \mathbb{R}\}$  è un gruppo:

$$\Phi_X^{t+s} = \Phi_X^t \circ \Phi_X^s, \quad (\Phi_X^t)^{-1} = \Phi_X^{-t}, \quad \Phi_X^0 = \text{Id}.$$

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, 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

$$(L_X \mathcal{F})(x) = \left. \frac{d}{dt} \mathcal{F}(\Phi_X^t(x)) \right|_{t=0}.$$

Notazioni alternative a  $L_X \mathcal{F}$  e  $L_{\underline{X}}f$ , ove sia ben chiaro il campo vettoriale di riferimento, sono  $\dot{\mathcal{F}}$  e  $\dot{f}$ .

### 1-forme su $M$ .

Sia ora  $T_x^*M$  lo spazio duale a  $T_xM$  (lo spazio 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  $(x, \theta)$ , con  $x \in M$  e  $\theta$  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  $(x, \theta(x)) \in T^*M$ ; la 1-forma sarà denotata semplicemente con  $\theta$ . Data una 1-forma  $\theta$  e un campo vettoriale  $X$  su  $M$ , resta definita una funzione  $\mathcal{F} : M \rightarrow \mathbb{R}$ , nel modo ovvio  $\mathcal{F}(x) = (\theta(x))(X(x))$ . 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

$$\theta(x) = \sum_{i=1}^n \theta_i(x) dx_i, \quad \theta_i = \theta\left(\frac{\partial}{\partial x_i}\right).$$

Assegnata una 1-forma  $\theta$  su  $M$ , restano definite le sue rappresentative locali, o covettori, che (sempre con piccolo abuso) denoteremo

$$\underline{\theta}(\underline{x}) = (\theta_1(\underline{x}), \dots, \theta_n(\underline{x})).$$

La funzione  $\theta(X) : M \rightarrow \mathbb{R}$  ha come rappresentativa locale il prodotto di vettore e covettore:

$$\underline{\theta}(\underline{x}) \cdot \underline{X}(\underline{x}) := \sum_{i=1}^n \theta_i(\underline{x}) X_i(\underline{x}) .$$

Se  $\underline{\theta}$  e  $\tilde{\theta}$  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

$$\tilde{\theta}(\tilde{x}) = J^T(\tilde{x}) \underline{\theta}(\psi(\tilde{x})) \quad (1.3)$$

(attenzione al verso: il confronto con (1.2) mostra che se i vettori vanno con  $J$ , le forme vanno con  $J^{-T}$ , inversa della trasposta), e corrispondentemente risulta

$$\tilde{\theta} \cdot \tilde{X} = \underline{\theta} \cdot \underline{X} , \quad \text{precisamente} \quad \tilde{\theta}(\tilde{x}) \cdot \tilde{X}(\tilde{x}) = \underline{\theta}(\psi(\tilde{x})) \cdot \underline{X}(\psi(\tilde{x})) ;$$

viceversa se una famiglia di 1-forme locali, una per ogni carta dell'atlante, soddisfa la relazione di compatibilità (1.3), allora resta definita una 1-forma  $\theta$  su  $M$ . La 1-forma  $\tilde{\theta}$  è anche detta *pull-back* di  $\underline{\theta}$  sotto  $\psi$ .

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 $M$ .

Consideriamo infine il caso di una 2-forma  $\omega$  su  $M$ . Una 2-forma esterna, 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  $\underline{\underline{\omega}} = (\omega_{ij})$ , precisamente si ha

$$(\omega(X, Y))(x) = \sum_{i,j=1}^n X_i(\underline{x}) \omega_{ij}(\underline{x}) Y_j(\underline{x}) , \quad \omega_{ij} = \omega\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) .$$

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  $\underline{\underline{\omega}}$  e  $\tilde{\underline{\underline{\omega}}}$  sono rappresentative locali di una stessa 2-forma  $\omega$ , allora si ha

$$\tilde{\underline{\underline{\omega}}} = J^T \underline{\underline{\omega}} J , \quad \text{più precisamente} \quad \tilde{\underline{\underline{\omega}}}(\tilde{x}) = J^T(\tilde{x}) \underline{\underline{\omega}}(\psi(\tilde{x})) J(\tilde{x}) , \quad (1.4)$$

e viceversa se una famiglia di 2-forme locali, una per ogni carta dell'atlante, soddisfa la relazione di compatibilità (1.4), allora resta definita una 2-forma su  $M$ . La 2-forma  $\underline{\omega}$  è detta *pull-back* di  $\underline{\omega}$  sotto  $\psi$ .

Si ottiene una 2-forma esterna moltiplicando esternamente due 1-forme. Ricordiamo che il *prodotto esterno*  $\theta \wedge \theta'$  di due 1-forme  $\theta$  e  $\theta'$  su  $M$  è la 2-forma esterna definita da

$$(\theta \wedge \theta')(X, X') = \theta(X)\theta'(X') - \theta(X')\theta'(X) ; \quad (1.5)$$

si ha evidentemente  $\theta' \wedge \theta = -\theta \wedge \theta'$  e dunque  $\theta \wedge \theta = 0$ . 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:

$$\omega(x) = \sum_{1 \leq i < j \leq n} \omega_{ij}(x) dx_i \wedge dx_j .$$

Un altro modo di generare 2-forme esterne da 1-forme è la *differenziazione esterna*: se, con riferimento a una carta qualsiasi,  $\theta = \sum \theta_i(x) dx_i$  è una 1-forma su  $M$ , il suo differenziale  $d\theta$  è per definizione la 2-forma esterna definita da

$$d\theta = \sum_{j=1}^n d\theta_j \wedge dx_j = \sum_{1 \leq i < j \leq n} \left( \frac{\partial \theta_j}{\partial x_i} - \frac{\partial \theta_i}{\partial x_j} \right) dx_i \wedge dx_j ;$$

il risultato non dipende dalla carta. Se  $d\theta = 0$ , la forma  $\theta$  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  $\omega$  si dice *non degenerare* se

$$\omega(X, Y) = 0 \quad \forall Y \quad \Rightarrow \quad X = 0 ;$$

equivalentemente, in ogni carta si ha  $\det \underline{\omega}(\underline{x}) \neq 0$  per ogni  $\underline{x}$ . Data una 2-forma non degenerare  $\omega$  resta stabilita una corrispondenza biunivoca tra 1-forme e campi vettoriali: infatti per ogni campo vettoriale  $Y$  è ben definita la 1-forma  $\theta = \omega(\cdot, Y)$ , ovvero la 1-forma  $\theta$  tale che  $\theta(X) = \omega(X, Y)$ , e viceversa, grazie alla non degenerazione, assegnata  $\theta$  esiste unico  $Y$  tale che  $\theta(X) = \omega(X, Y)$  (è immediato in coordinate, usando  $\det \underline{\omega} \neq 0$ ). Se  $\omega$  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 (delle quali però non faremo uso). Una  $k$ -forma esterna  $\alpha$  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 (1.5): precisamente, date una  $k$ -forma  $\alpha$  e una  $k'$ -forma  $\alpha'$ , il loro prodotto esterno  $\alpha \wedge \alpha'$  è la  $(k + k')$ -forma  $\beta$  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  $\theta^{(1)}, \dots, \theta^{(k)}$ ,

$$\alpha = \theta^{(1)} \wedge \dots \wedge \theta^{(k)} ,$$

e precisamente è la  $k$ -forma

$$\alpha(X^{(1)}, \dots, X^{(k)}) = \sum_{j_1, \dots, j_k} (-1)^p \theta^{(1)}(X^{(j_1)}) \dots \theta^{(k)}(X^{(j_k)}) .$$

Una base per le  $k$ -forme è data dai prodotti  $dx_{j_1} \wedge \dots \wedge dx_{j_k}$ ,  $j_1 < \dots < j_k$ , e dunque si potrà scrivere

$$\alpha = \sum_{0 \leq j_1 < \dots < j_k \leq n} \alpha_{j_1, \dots, j_k}(x) dx_{j_1} \wedge \dots \wedge dx_{j_k} .$$

Infine il differenziale  $d\alpha$  della  $k$ -forma esterna  $\alpha$  è per definizione la  $(k + 1)$ -forma esterna

$$\beta = \sum_{j_1 < \dots < j_k} d\alpha_{j_1, \dots, j_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_k} .$$

Una  $k$ -forma  $\alpha$  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  $\alpha$  è chiusa, in generale essa non è il differenziale di una  $(k - 1)$ -forma (lo è solo localmente). Se una  $k$  forma  $\alpha$  è il differenziale di una  $(k - 1)$  forma, allora essa si dice *esatta*.

## 1.2. 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:

$$\underline{q} = \beta(\tilde{\underline{q}}) , \quad \underline{v} = B(\tilde{\underline{q}})\tilde{\underline{v}} , \quad B := \left( \frac{\partial \beta_i}{\partial \tilde{q}_j} \right) , \quad (1.6)$$

ove  $\beta$  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 ha<sup>3)</sup>  $l = k(\underline{q}, \underline{v}) + u(\underline{q})$ , con

$$k(\underline{q}, \underline{v}) = \sum_{i,j=1}^n g_{ij}(\underline{q}) v_i v_j ,$$

$\underline{\underline{g}} = (g_{ij})$  simmetrica positiva. Le equazioni di Lagrange hanno allora la forma (implicita, ma esplicitabile grazie a  $\det \underline{\underline{g}} \neq 0$ )

$$\dot{q}_i = v_i , \quad \frac{d}{dt} \frac{\partial l}{\partial v_i} - \frac{\partial l}{\partial q_i} = 0 , \quad i = 1, \dots, n . \quad (1.7)$$

Resta così definito per ogni carta un campo vettoriale locale  $\underline{X} = (X_{q_1}, \dots, X_{v_n})$ . 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 (1.7), 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,

$$\tilde{\underline{\underline{g}}} = B^T \underline{\underline{g}} B \quad (1.8)$$

(tutte le quantità sono calcolate in  $\tilde{q}$ ), e ciò garantisce  $\tilde{k}(\tilde{q}, \tilde{v}) = k(\underline{q}, \underline{v})$ . In linea di principio nulla vieta di arricchire l'atlante di  $TQ$  di ulteriori carte "non adattate", con funzioni di transizione diverse dalle (1.6): a prezzo però di perdere la forma (1.7) 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

$$\underline{p} = \frac{\partial k}{\partial \underline{v}} = g \underline{v} , \quad (1.9)$$

detto momento, che in conseguenza alla (1.8) trasforma secondo la regola dei covettori

$$\tilde{\underline{p}} = B^T \underline{p} . \quad (1.10)$$

Alla lagrangiana  $l$  si associa  $h$ , detta hamiltoniana locale, definita da<sup>4)</sup>

$$h(\underline{p}, \underline{q}) = \underline{p} \cdot \underline{\underline{g}}^{-1} \underline{p} - l(\underline{q}, \underline{\underline{g}}^{-1} \underline{p}) ,$$

---

<sup>3)</sup> Non prendiamo qui in considerazione il caso di vincoli mobili, di coordinate dipendenti dal tempo, o di potenziali dipendenti dalla velocità.

<sup>4)</sup> E' abitudine di chi scrive, che ha studiato sui libri russi, anteporre le  $p$  alle  $q$  nel formalismo hamiltoniano.



e si osserva che le equazioni locali per  $(\underline{p}, \underline{q})$ , equivalenti alle equazioni di Lagrange (1.7), sono le equazioni di Hamilton, o equazioni canoniche:

$$\dot{p}_i = -\frac{\partial h}{\partial q_i}, \quad \dot{q}_i = \frac{\partial h}{\partial p_i}, \quad i = 1, \dots, n. \quad (1.11)$$

A  $h$  è dunque associato il campo vettoriale locale  $\underline{X}^h$ , detto campo vettoriale hamiltoniano (locale) di hamiltoniana  $h$ ,

$$\underline{X}^h = (X_{\underline{p}}^h, X_{\underline{q}}^h) = \left(-\frac{\partial h}{\partial \underline{q}}, \frac{\partial h}{\partial \underline{p}}\right).$$

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 (1.6),

$$\underline{p} = B^{-T}(\underline{q})\tilde{\underline{p}}, \quad \underline{q} = \beta(\tilde{\underline{q}}). \quad (1.12)$$

La corrispondenza tra vettori e covettori stabilita dalla (1.9) è intrinseca; ne segue che le diverse funzioni di Hamilton locali sono rappresentative di una funzione  $\mathcal{H} : T^*Q \rightarrow \mathbb{R}$ .<sup>5)</sup> Infine, i diversi campi vettoriali hamiltoniani locali si incollano bene: la matrice jacobiana del cambio di carta è infatti

$$J(\tilde{\underline{p}}, \tilde{\underline{q}}) = \left( \begin{array}{c|c} B^{-T} & \\ \hline & B \end{array} \right), \quad (1.13)$$

e si verifica subito che  $\tilde{\underline{X}} = J^{-1}\underline{X}$ . Resta così definito 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 (1.7) sono rappresentative locali nelle carte dell'atlante adattato.

Uno dei maggiori motivi di interesse del formalismo hamiltoniano, che lo distingue da quello lagrangiano, è il fatto che l'atlante adattato di  $T^*Q$  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 (1.11), queste abbiano carattere geometrico, cioè siano rappresentative locali di un'equazione differenziale sulla varietà.

---

<sup>5)</sup> La 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.

Sia allora  $h$  la rappresentativa locale di  $\mathcal{H}$  in una carta, e  $\underline{X}^h$  il suo campo vettoriale hamiltoniano locale,  $\underline{X}^h = (-\frac{\partial h}{\partial q}, \frac{\partial h}{\partial p})$ . Introduciamo la matrice  $2n \times 2n$

$$E = \left( \begin{array}{c|c} & I \\ \hline -I & \end{array} \right), \quad E^T = -E = E^{-1},$$

ove  $I$  denota la matrice identità  $n \times n$ . La relazione tra  $h$  e  $\underline{X}^h$  si scrive allora nella forma sintetica

$$E\underline{X}^h = \left( \frac{\partial h}{\partial x_1}, \dots, \frac{\partial h}{\partial x_{2n}} \right), \quad (1.14)$$

adatta a discuterne il carattere geometrico. Infatti a destra dell'uguale appare l' $n$ -pla delle derivate di  $h$ , che è rappresentativa locale di una 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  $\mathcal{H}$ , se<sup>6)</sup> la matrice  $E$  antisimmetrica (e non singolare) è rappresentativa locale di una 2-forma  $\omega$  (non degenera) su  $M$ . In tal caso infatti la (1.14) diventa globale, e precisamente si innalza a*

$$\omega(\cdot, X^{\mathcal{H}}) = d\mathcal{H}. \quad (1.15)$$

Sia  $\psi : \tilde{x} \mapsto \underline{x}$  la funzione di transizione tra due carte di  $\mathcal{A}$ , e  $J$  la corrispondente matrice jacobiana. Dalla regola (1.4) sul *pull-back* di forme si vede che perché una 2-forma  $\omega$  su  $M$  sia rappresentata da  $E$  in entrambe le carte, deve risultare

$$J^T E J = E. \quad (1.16)$$

Matrici  $J$  che soddisfino la (1.16) 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  $J^T$ ;  $E$  stessa è simplettica, e anzi è detta *identità simplettica*.

**Definizione.** 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*.

La compatibilità tra le equazioni di Hamilton locali dunque è 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 (1.13) 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).

---

<sup>6)</sup> La condizione è ovviamente sufficiente, ma è anche necessaria se vogliamo che la proprietà sia vera per ogni  $\mathcal{H}$ .

- **Esercizio.** Posto  $M = T^*\mathbb{R} = \mathbb{R}^2$ , si verifichi che la trasformazione  $(p, q) = \psi(I, \varphi)$  definita da  $p = \sqrt{2I\omega} \cos \varphi$ ,  $q = \sqrt{2I/\omega} \sin \varphi$ ,  $\omega > 0$ , è simplettica. Posto  $h = \frac{1}{2}(p^2 + \omega^2 q^2)$ , si scrivano l'hamiltoniana e le equazioni del moto nella nuova carta.
- **Osservazione.** La nozione di simpletticità ricorda la nozione di ortogonalità ( $E$  sostituisce l'identità), e come le matrici ortogonali preservano il prodotto scalare euclideo  $\underline{X} \cdot \underline{Y}$ , così le matrici simplettiche preservano il cosiddetto *prodotto scalare simplettico*

$$(\underline{X}, \underline{Y})_\wedge = \underline{X} \cdot E \underline{Y} = \sum_{i=1}^n X_{p_i} Y_{q_i} - X_{q_i} Y_{p_i}$$

(che però non è un prodotto scalare:  $(\underline{X}, \underline{X})_\wedge = 0$ ).

### 1.3. Varietà simplettiche

Riprendiamo il caso generale di una varietà  $M$  di dimensione pari munita di un atlante simplettico. La 2-forma  $\omega$  su  $M$ , di cui la matrice  $E$  è rappresentativa locale in ogni carta, è con evidenza

$$\omega = dp \wedge dq := \sum_{i=1}^n dp_i \wedge dq_i ; \quad (1.17)$$

la forma, oltre che ben definita su  $M$ , è non degenera è anche chiusa (come segue dal fatto che ha coefficienti costanti). Essa è detta *2-forma canonica*. Un modo assai frequente di scrivere la relazione di compatibilità canonica (1.16) è

$$dp \wedge dq = d\tilde{p} \wedge d\tilde{q} ; \quad (1.18)$$

la relazione si legge comunemente dicendo che *le trasformazioni canoniche preservano la 2-forma canonica* (1.17). La (1.18) è adoperata spesso come definizione di trasformazione canonica.<sup>7)</sup>

E' spontanea ora la seguente definizione:

**Definizione.** Una varietà differenziabile di dimensione pari munita di un atlante simplettico si dice *varietà simplettica*.

E poi:

**Definizione.** 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 (1.15), ove  $\omega$  è la 2-forma canonica su  $M$ . L'equazione differenziale sulla varietà

$$\dot{x} = X^{\mathcal{H}} , \quad (1.19)$$

di cui la (1.11) è rappresentativa locale, è detta *sistema hamiltoniano* su  $M$ .

---

<sup>7)</sup> Si vede facilmente che la 2-forma canonica è chiusa ma in generale non è esatta. Localmente essa è il differenziale della 1-forma  $\theta = p dq = \sum_i p_i dq_i$ , detta a sua volta *1-forma canonica*:  $\omega = d\theta$ . Ma in generale la 1-forma  $\theta$  non esiste globalmente su  $M$ , e la differenza  $p dq - \tilde{p} d\tilde{q}$  è chiusa ma non nulla. Si dimostra facilmente che la 1-forma canonica è preservata dalle trasformazioni puntuali estese.

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 data<sup>8)</sup>

Le trasformazioni canoniche, per distinguerle da più generali cambiamenti di coordinate, nel seguito saranno denotate con il simbolo  $\mathcal{C}$ ; scriveremo pertanto

$$(\underline{p}, \underline{q}) = \mathcal{C}(\tilde{\underline{p}}, \tilde{\underline{q}}), \quad \mathcal{C} = (\underline{C}_p, \underline{C}_q) = (C_{p_1}, \dots, C_{p_n}, C_{q_1}, \dots, C_{q_n}).$$

Useremo anche le notazioni leggere

$$L_{\mathcal{H}}, \quad \Phi_{\mathcal{H}}^t \quad \text{al posto di} \quad L_{X^{\mathcal{H}}}, \quad \Phi_{X^{\mathcal{H}}}^t;$$

in coordinate  $L_h, \Phi_h^t$ . Infine, tre elementari ma utili relazioni:

$$X^{c\mathcal{H}} = cX^{\mathcal{H}}, \quad L_{c\mathcal{H}} = cL_{\mathcal{H}}, \quad \Phi_{c\mathcal{H}}^t = \Phi_{\mathcal{H}}^{ct} \quad (1.20)$$

per ogni  $c \neq 0$ .

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 può essere estesa globalmente; in questo modo ci siamo imbattuti nella 2-forma canonica  $\omega = dp \wedge dq$ , e abbiamo preteso che  $\omega$  fosse ben definita su  $M$  grazie alla (1.16) o equivalentemente alla (1.18). 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  $(M, \omega)$ , con  $M$  varietà differenziabile di dimensione pari e  $\omega$  una qualsiasi 2-forma chiusa antisimmetrica non degenera su  $M$  (senza chiedere a priori che  $M$  sia munito di un atlante particolare, nel quale  $\omega$  abbia la forma particolare  $dp \wedge dq$ ). La definizione (1.15) 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:<sup>9)</sup>

**Proposizione 1.1** (Teorema di Darboux). *Sia  $M$  una varietà differenziabile di dimensione pari, e  $\omega$  una 2-forma chiusa antisimmetrica non degenera su  $M$ . Allora esiste un atlante di  $M$  (atlante simplettico) tale che per tutte le sue carte la 2-forma  $\omega$  assume la forma canonica (1.17).*

<sup>8)</sup> Una 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 (1.11), ma il sistema — che non è cambiato — resta hamiltoniano.

<sup>9)</sup> Il teorema è conseguenza di un lemma di algebra lineare, secondo il quale *se la matrice  $\Omega$  di ordine  $2n$  è antisimmetrica e non degenera, allora esiste  $J$  tale che*

$$J^T \Omega J = E.$$

Il risultato ricorda il più noto risultato secondo cui, se  $\Omega$  è simmetrica e definita positiva, allora esiste  $J$  tale che  $J^T \Omega J = I$ .

## 1.4. Parentesi di Poisson di funzioni, parentesi di Lie di campi vettoriali

Parentesi di Poisson.

A livello elementare locale ci si imbatte naturalmente, nell'ambito della meccanica hamiltoniana, in una operazione binaria tra funzioni (chiusa per funzioni di classe  $C^\infty$ ) detta *parentesi di Poisson*,

$$\{f, g\} = \sum_{i=1}^n \left( \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right); \quad (1.21)$$

in notazione compatta

$$\{f, g\} = E \partial_{\underline{x}} f \cdot \partial_{\underline{x}} g, \quad (1.22)$$

ove si è anche introdotta la notazione agile  $\partial_{\underline{x}} = \frac{\partial}{\partial \underline{x}}$ . L'operazione è bilineare, antisimmetrica, e soddisfa l'*identità di Jacobi*:

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0 \quad (1.23)$$

(per la verifica si veda più avanti, proposizione 1.3).

Sia ora  $M$  una varietà di dimensione pari, e siano  $\mathcal{F}, \mathcal{G}$  due funzioni  $M \rightarrow \mathbb{R}$ . E' spontaneo chiedersi a quali condizioni l'operazione (1.21) tra rappresentative locali definisca una operazione intrinseca su  $M$ , ovvero quando le diverse funzioni locali  $\{f, g\}$ , una per ogni carta dell'atlante, siano rappresentative di una funzione su  $M$ . Ricordando la (1.1) si vede che ciò corrisponde a richiedere che per una qualunque coppia di carte di  $M$  (con intersezione non vuota dei domini), detto  $\mathcal{C}$  il cambiamento di variabili, risulti

$$\{f \circ \mathcal{C}, g \circ \mathcal{C}\}_{\tilde{p}, \tilde{q}} = \{f, g\}_{p, q} \circ \mathcal{C} \quad (1.24)$$

(per maggior chiarezza si sono messe in evidenza a piede le variabili rispetto alle quali si deriva, ma è ridondante).

**Definizione.** Si dice che il cambiamento di variabili  $\mathcal{C}$  *preserva la parentesi di Poisson* se la (1.24) è soddisfatta per ogni coppia di funzioni regolari  $f, g$ .

Si dimostra facilmente la seguente (non sorprendente) proposizione:

**Proposizione 1.2.** *Condizione necessaria e sufficiente perchè  $\mathcal{C}$  preservi le parentesi di Poisson è che essa sia canonica.*

**Dimostrazione.** Dalla (1.22) segue subito (usando  $\partial_{\tilde{x}} \tilde{f} = J^T(\partial_{\underline{x}} f) \circ \mathcal{C}$ )

$$\{\tilde{f}, \tilde{g}\} = E \partial_{\tilde{x}} \tilde{f} \cdot \partial_{\tilde{x}} \tilde{g} = (J E J^T)(\partial_{\underline{x}} f \cdot \partial_{\underline{x}} g) \circ \mathcal{C};$$

pertanto le parentesi di Poisson sono preservate se e solo se  $J^T$ , e dunque  $J$ , è simplettica. ■

Dunque su di una varietà simplettica la parentesi di Poisson è un'operazione intrinseca; una notazione appropriata è  $\{\mathcal{F}, \mathcal{G}\}$ . La relazione intrinseca corrispondente alla (1.22) è

$$\{\mathcal{F}, \mathcal{G}\} = -\omega(X^{\mathcal{F}}, X^{\mathcal{G}})$$

(e questa si sarebbe potuta usare direttamente come definizione di parentesi di Poisson, verificando poi che in coordinate si ha la consueta definizione (1.21)).

Si vede facilmente (la verifica un po' noiosa è lasciata per esercizio) che tutte le parentesi di Poisson sono preservate se sono preservate le *parentesi di Poisson fondamentali*:

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

(ove si intende  $p_i = p_i(\tilde{p}, \tilde{q}) = \mathcal{C}_{p_i}(\tilde{p}, \tilde{q})$ , e similmente per  $q_i$ , e le derivate sono eseguite rispetto alle  $\tilde{p}_i, \tilde{q}_i$ ).

Si osserva poi che si ha

$$L_{\mathcal{F}} = \{ \cdot, \mathcal{F} \}, \quad L_{\mathcal{G}} = -L_{\mathcal{G}} \mathcal{F},$$

e che l'identità di Jacobi si riscrive

$$L_{\mathcal{F}} L_{\mathcal{G}} - L_{\mathcal{G}} L_{\mathcal{F}} + L_{\{\mathcal{F}, \mathcal{G}\}} = 0. \quad (1.25)$$

Infine, le stesse equazioni canoniche si possono scrivere in termini di parentesi di Poisson:

$$\dot{p}_i = \{p_i, h\}, \quad \dot{q}_i = \{q_i, h\}, \quad \dot{\underline{x}} = \{\underline{x}, h\}.$$

### Parentesi di Lie di campi vettoriali.

Sia  $M$  una varietà qualsiasi, e siano  $X, Y$  campi vettoriali su  $M$ . Il prodotto  $L_X L_Y$  delle corrispondenti derivate di Lie *non* è un operatore di derivazione ( $L_X L_Y \mathcal{F}$  contiene le derivate seconde di  $\mathcal{F}$ ). Lo è invece il commutatore

$$[L_X, L_Y] := L_X L_Y - L_Y L_X,$$

precisamente si ha  $[L_X, L_Y] = L_Z$ , con  $Z$  definito, con riferimento a una carta qualsiasi, da

$$Z_i = L_X Y_i - L_Y X_i; \quad (1.26)$$

la verifica è immediata (i termini di derivata seconda si semplificano), e la definizione è indipendente dalla carta.

**Definizione.** Il campo vettoriale definito dalla (1.26) si dice *parentesi di Lie* di  $X$  e  $Y$ , e si denota

$$Z = [X, Y].$$

Si ha dunque, per definizione di parentesi di Lie,

$$[L_X, L_Y] = L_{[X, Y]}. \quad (1.27)$$

Per campi vettoriali hamiltoniani avviene che la loro parentesi di Lie è ancora un campo hamiltoniano: precisamente

**Proposizione 1.3.** *Risulta*

$$[X^{\mathcal{F}}, X^{\mathcal{G}}] = X^{\{\mathcal{G}, \mathcal{F}\}}. \quad (1.28)$$

**Dimostrazione.** Per la (1.27), la (1.28) è equivalente a

$$[L_{\mathcal{F}}, L_{\mathcal{G}}] + L_{\{\mathcal{F}, \mathcal{G}\}} = 0 ,$$

e questa non è altro che l'identità di Jacobi (1.25) rappresentata in coordinate dalla (1.23). Dimostriamo allora qui l'identità di Jacobi. Per questo osserviamo che, se si sviluppano le derivate nella (1.23), si trova una somma di termini contenenti ciascuno una derivata seconda di una delle tre funzioni. D'altra parte è evidente che il risultato non può contenere derivate seconde: non di  $h$ , perchè la somma dei primi due termini della (1.23) risulta essere  $[L_f, L_g]h$ , e il commutatore si è visto essere un operatore di derivazione del primo ordine, mentre l'ultimo certamente non contiene derivate seconde di  $h$ ; per simmetria neanche di  $f$  e di  $g$ . Si conclude che tutti i termini si semplificano e la somma è nulla. ■

### Commutazione di campi e flussi.

Si dice che due campi  $X$  e  $Y$  commutano se  $[X, Y] = 0$ ; si dice anche che due funzioni  $\mathcal{F}$  e  $\mathcal{G}$  su una varietà симплетica commutano, o che *sono in involuzione*, se esse hanno parentesi di Poisson nulla. Per quanto appena visto, due campi vettoriali hamiltoniani certamente commutano se commutano le rispettive funzioni di Hamilton (precisamente: i campi commutano se e solo se la parentesi di Poisson delle rispettive Hamiltoniane è costante).

Infine, dati due campi vettoriali si dice che i loro flussi  $\Phi_X^t, \Phi_Y^t$  commutano se per ogni  $s, t \in \mathbb{R}$  si ha

$$\Phi_Y^t \circ \Phi_X^s - \Phi_X^s \circ \Phi_Y^t = 0 .$$

Per ogni  $\mathcal{F} : M \rightarrow \mathbb{R}$  la “mancata commutazione” è ben rappresentata da

$$\Delta_{\mathcal{F}}(s, t) := \mathcal{F} \circ \Phi_Y^t \circ \Phi_X^s - \mathcal{F} \circ \Phi_X^s \circ \Phi_Y^t .$$

**Proposizione 1.4.** *Con riferimento alla notazione sopra introdotta,*

- i) *per ogni  $\mathcal{F}$  si ha  $\Delta_{\mathcal{F}}(s, t) = st L_{[X, Y]} \mathcal{F} + \mathcal{O}_3(s, t)$ ;*
- ii) *se  $[X, Y] = 0$  allora i flussi commutano (e viceversa, per il punto precedente).*

La dimostrazione è demandata all'appendice A.

### 1.5. Generazione di trasformazioni canoniche

La teoria delle perturbazioni richiede di saper generare, all'occorrenza, cambiamenti di variabili  $\mathcal{C} : (\tilde{p}, \tilde{q}) \mapsto (p, q)$  la cui canonicità sia assicurata a priori. Introduciamo qui due metodi che si riveleranno molto utili nel seguito.

#### Funzioni generatrici.

Sia  $A$  un aperto di  $\mathbb{R}^{2n}$ , e  $S$  una funzione:  $A \rightarrow \mathbb{R}$ . La scrittura

$$p_i = \frac{\partial S}{\partial q_i}(\tilde{p}, \tilde{q}) , \quad \tilde{q}_i = \frac{\partial S}{\partial \tilde{p}_i}(\tilde{p}, \tilde{q}) , \quad i = 1, \dots, n , \quad (1.29)$$

definisce implicitamente un cambiamento di variabili  $(\underline{p}, \underline{q}) = \mathcal{C}(\underline{\tilde{p}}, \underline{\tilde{q}})$ , che invia un opportuno aperto  $\tilde{V} \subset \mathbb{R}^{2n}$  nella sua immagine  $V = \mathcal{C}(\tilde{V}) \subset \mathbb{R}^{2n}$ , non appena sia soddisfatta in  $A$  la condizione di non degenerazione

$$\det \left( \frac{\partial^2 S}{\partial \tilde{p}_i \partial \tilde{q}_j} \right) \neq 0 . \quad (1.30)$$

Infatti, in questa ipotesi si può invertire la seconda delle (1.29) ottenendo, in un qualche  $\tilde{V}$ ,  $q_i = \mathcal{C}_{q_i}(\underline{\tilde{p}}, \underline{\tilde{q}})$ , e poi per sostituzione nella prima resta definita  $p_i = \mathcal{C}_{p_i}(\underline{\tilde{p}}, \underline{\tilde{q}})$ .

Si dimostra senza sforzo che

**Proposizione 1.5.** *La trasformazione  $\mathcal{C}$ , ove definita, è canonica.*

**Dimostrazione.** Dalle (1.29) si deduce immediatamente

$$\underline{p} \cdot d\underline{q} + \underline{\tilde{q}} \cdot d\underline{\tilde{p}} = \frac{\partial S}{\partial \underline{q}} \cdot d\underline{q} + \frac{\partial S}{\partial \underline{\tilde{p}}} \cdot d\underline{\tilde{p}} = dS ,$$

ovvero la 1-forma  $\theta = \underline{p} \cdot d\underline{q} + \underline{\tilde{q}} \cdot d\underline{\tilde{p}}$  è chiusa (è localmente esatta). Differenziando segue  $d\underline{p} \wedge d\underline{q} - d\underline{\tilde{p}} \wedge d\underline{\tilde{q}} = 0$ , pertanto la 2-forma canonica è preservata. ■

Si osservi che  $S(\underline{\tilde{p}}, \underline{q}) := \underline{\tilde{p}} \cdot \underline{q}$  genera l'identità, e che funzioni del tipo

$$S(\underline{\tilde{p}}, \underline{q}) := \underline{\tilde{p}} \cdot \underline{q} + W(\underline{\tilde{p}}, \underline{q}) , \quad (1.31)$$

con  $W$  piccola in norma  $C^2$  (tipicamente  $W = \varepsilon \mathcal{W}$ , con  $\varepsilon$  piccolo) soddisfano sicuramente la (1.30), e generano trasformazioni prossime all'identità (intendendosi con questo che la differenza  $\mathcal{C}(\underline{\tilde{p}}, \underline{\tilde{q}}) - (\underline{\tilde{p}}, \underline{\tilde{q}})$  è piccola in norma  $C^1$ ). La funzione  $S$  (a volte anche  $W$ ) è detta *funzione generatrice* della trasformazione canonica  $\mathcal{C}$ .

**Proposizione 1.6.** *Tutte le trasformazioni canoniche prossime all'identità in norma  $C^1$  si possono scrivere attraverso una opportuna funzione generatrice  $S$  della forma (1.29).*

La proposizione è utile in teoremi di non esistenza.

**Dimostrazione.** Se

$$(\underline{p}, \underline{q}) = (\underline{\tilde{p}}, \underline{\tilde{q}}) + \gamma(\underline{\tilde{p}}, \underline{\tilde{q}}) , \quad \gamma \text{ piccola in norma } C^1 ,$$

allora sicuramente per inversione si ottiene

$$\underline{\tilde{q}} = \underline{q} + \alpha(\underline{\tilde{p}}, \underline{q}) , \quad \underline{p} = \underline{\tilde{p}} + \beta(\underline{\tilde{p}}, \underline{q}) , \quad \alpha, \beta \text{ piccole in norma } C^1 ;$$

si tratta di dimostrare che  $\alpha$  e  $\beta$  sono derivate parziali di una opportuna  $W(\underline{\tilde{p}}, \underline{q})$  rispetto a  $\underline{\tilde{p}}$  e rispettivamente  $\underline{q}$ . Percorrendo a ritroso la dimostrazione della proposizione precedente, e sfruttando questa volta come ipotesi la chiusura della 1-forma  $\theta$  (conseguenza della canonicità), si trova facilmente  $W$  locale, piccola in norma  $C^2$ , tale che

$$\beta(\underline{\tilde{p}}, \underline{q}) \cdot d\underline{q} + \alpha(\underline{\tilde{p}}, \underline{q}) \cdot d\underline{\tilde{p}} = dW ,$$



e questo basta. ■

Un altro modo di generare trasformazioni canoniche attraverso una funzione generatrice “in variabili miste” è quello di scrivere

$$\underline{p} = \frac{\partial F}{\partial \underline{q}}(\underline{q}, \underline{\tilde{q}}), \quad \underline{\tilde{p}} = -\frac{\partial F}{\partial \underline{\tilde{q}}}(\underline{q}, \underline{\tilde{q}}), \quad (1.32)$$

ancora con la condizione di non degenerazione

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

Anche in questo caso si dimostra immediatamente che, per inversione nella seconda e sostituzione nella prima, si ottiene una trasformazione  $(\underline{p}, \underline{q}) = \mathcal{C}(\underline{\tilde{p}}, \underline{\tilde{q}})$  canonica. Si osservi che  $F(\underline{q}, \underline{\tilde{q}}) = \underline{q} \cdot \underline{\tilde{q}}$  genera la trasformazione che scambia coordinate e momenti:  $p_i = \tilde{q}_i$ ,  $q_i = -\tilde{p}_i$ . Si possono anche mescolare le (1.29) e (1.32), operando diversamente, nell’uno o nell’altro modo, a seconda dei gradi di libertà (e si dimostra che, operando in questo modo, si scrivono tutte le trasformazioni canoniche possibili).

Il flusso hamiltoniano come trasformazione canonica.

Mettiamoci in coordinate, e dunque in un aperto  $V \subset \mathbb{R}^{2n}$ ; sia  $h$  una funzione di Hamilton locale, e  $\Phi_h^t$ ,  $t \in \mathbb{R}$ , il suo flusso. Per  $s$  fissato sufficientemente piccolo (in relazione a  $h$ : si ricordino le (1.20)) l’insieme

$$\tilde{V} = \{(\underline{\tilde{p}}, \underline{\tilde{q}}) \in V : \Phi_h^t(\underline{\tilde{p}}, \underline{\tilde{q}}) \in V \text{ per } |t| \leq s\}$$

è aperto e non vuoto.

**Proposizione 1.7.** *La trasformazione di variabili*

$$(\underline{p}, \underline{q}) = \Phi_h^t(\underline{\tilde{p}}, \underline{\tilde{q}}), \quad \Phi_h^t : \tilde{V} \rightarrow \Phi_h^t(\tilde{V}) \in V, \quad (1.33)$$

è canonica.

**Dimostrazione.** Poniamo  $J_t(\underline{\tilde{x}}) = \left( \frac{\partial \Phi_h^t}{\partial \underline{\tilde{x}}} \right)$ ,  $\underline{\tilde{x}} = (\underline{\tilde{p}}, \underline{\tilde{q}})$ , e mostriamo che questa matrice è симплетica; per questo basta verificare che per ogni  $t$  risulta

$$\frac{\partial}{\partial t} [J_t^T(\underline{\tilde{x}}) E J_t(\underline{\tilde{x}})] = 0. \quad (1.34)$$

Ci si riporta a  $t = 0$  usando  $J_{t+\tau}(\underline{\tilde{x}}) = J_\tau(\Phi^t(\underline{\tilde{x}}))J_t(\underline{\tilde{x}})$ : infatti, ponendo per brevità  $\underline{y}_t = \Phi^t(\underline{\tilde{x}})$ , la (1.34) si riscrive

$$J_t^T(\underline{\tilde{x}}) \frac{\partial}{\partial \tau} [J_\tau^T(\underline{y}_t) E J_\tau(\underline{y}_t)]_{\tau=0} J_t(\underline{\tilde{x}}) = 0.$$

Ora si vede facilmente che per ogni  $\underline{y}$  si ha  $\frac{\partial}{\partial \tau} [J_\tau^T(\underline{y}) E J_\tau(\underline{y})]_{\tau=0} = 0$ . Infatti si ottiene subito

$$\frac{\partial}{\partial \tau} J_\tau(\underline{y}) \Big|_{\tau=0} = \frac{\partial \underline{X}}{\partial \underline{x}}(\underline{y}) = E^{-1} H(\underline{y}), \quad H := \left( \frac{\partial^2 h}{\partial x_i \partial x_j} \right),$$

e similmente  $\frac{\partial}{\partial \tau} J_\tau^T(\underline{y})|_{\tau=0} = H(\underline{y})E$ ; ricordando  $J_0 = I$  e  $E^2 = -I$  segue allora

$$\frac{\partial}{\partial \tau} [J_\tau^T E J_\tau]_{\tau=0} = \frac{\partial}{\partial \tau} J_\tau^T|_{\tau=0} E + E \frac{\partial}{\partial \tau} J_\tau|_{\tau=0} = 0 .$$

■

Per  $s$  piccolo, o equivalentemente per  $\|h\|$  piccolo, si ottengono in questo modo trasformazioni prossime all'identità, con  $\tilde{V}$  di poco ridotto rispetto a  $V$ . Useremo massicciamente questo metodo nel seguito (benché in questo modo *non* si ottengano tutte le trasformazioni canoniche prossime all'identità).

Trasformazioni canoniche con valenza diversa da uno.

Consideriamo il banale riscaldamento

$$p_i = \alpha \tilde{p}_i , \quad q_i = \beta \tilde{q}_i , \quad \alpha, \beta \in \mathbb{R} \setminus \{0\} ,$$

che con evidenza muta la 2-forma canonica  $\omega$  in

$$\tilde{\omega} = \gamma \omega , \quad \gamma = (\alpha\beta)^{-1} ,$$

e dunque è una trasformazione canonica se e solo se  $\alpha\beta = 1$ . Presa una hamiltoniana locale  $h$  qualsiasi, e posto

$$\tilde{h}(\tilde{p}, \tilde{q}) = h(\alpha\tilde{p}, \beta\tilde{q})$$

(*pull-back* di  $h$ ), le variabili  $(\tilde{p}, \tilde{q})$  *non* seguono le equazioni di Hamilton relative a  $\tilde{h}$ , bensì le equazioni

$$\dot{\tilde{p}}_i = -\gamma \frac{\partial \tilde{h}}{\partial \tilde{q}_i} , \quad \dot{\tilde{q}}_i = \gamma \frac{\partial \tilde{h}}{\partial \tilde{p}_i} .$$

Ma si vede subito che ci si riporta alle consuete equazioni di Hamilton pur di sostituire a  $\tilde{h}$  la nuova hamiltoniana locale  $\hat{h} = \gamma \tilde{h}$ ,

$$\hat{h}(\tilde{p}, \tilde{q}) = \gamma h(\alpha\tilde{p}, \beta\tilde{q}) .$$

Strettamente parlando non possiamo dire che  $\hat{h}$  e  $h$  siano rappresentative locali di una funzione su  $M$  (la funzione è cambiata per una costante moltiplicativa). Tuttavia è molto comodo a volte poter disporre di questi riscaldamenti “quasi canonici” in aggiunta alle trasformazioni canoniche sopra definite. In diversi testi, orientati più al lavoro locale che all'interpretazione geometrica, tali trasformazioni sono considerate canoniche a pieno titolo, e la costante  $\gamma$  è detta *valenza* della trasformazione; le trasformazioni canoniche in senso stretto, introdotte precedentemente, sono chiamate *univalenti*. Il termine “simplettiche” è comunque sempre riferito soltanto ad esse.

Osserviamo che  $\gamma$  scompare se si riscalda anche il tempo: precisamente, del tutto in generale si ha che con riferimento al nuovo tempo  $\tau = ct$  le equazioni di Hamilton di  $h$  sono mutate in

$$\frac{dp_i}{d\tau} = -c^{-1} \frac{\partial h}{\partial p_i} , \quad \frac{dq_i}{d\tau} = -c^{-1} \frac{\partial h}{\partial q_i} ,$$

e dunque invece di rimpiazzare  $\tilde{h}$  con  $\hat{h} = \gamma \tilde{h}$  come sopra si può equivalentemente riscaldare il tempo, ponendo  $\tau = \gamma t$ .

## 2. Il teorema di Liouville–Arnol'd

### 2.1. Le variabili di azione–angolo nel pendolo.

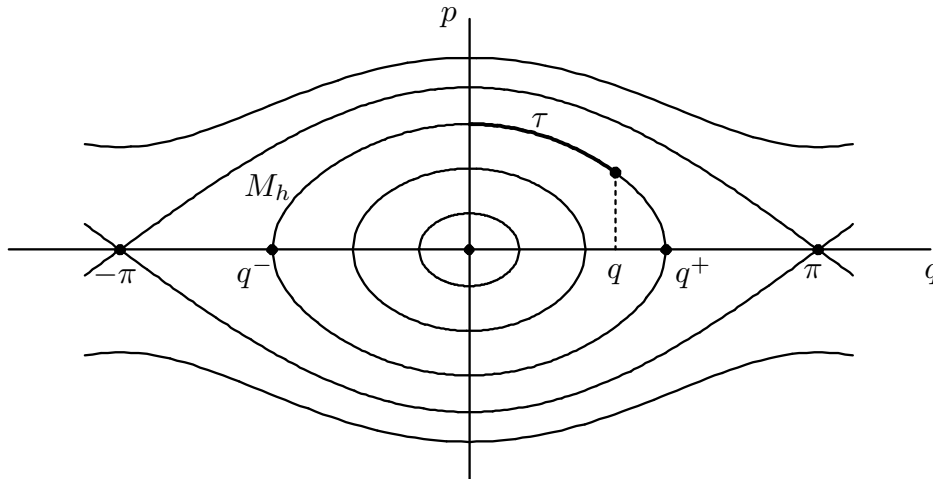
Consideriamo un sistema dinamico autonomo a un grado di libertà; per essere definiti, facciamo riferimento al pendolo:

$$H(p, q) = \frac{p^2}{2} + V(q), \quad V(q) = -\omega^2 \cos q, \quad (2.1)$$

ma quanto diremo in realtà dipende assai poco dalla scelta di  $V$ . Il ritratto in fase del sistema è riportato in figura 2.1; le traiettorie coincidono con le superfici di livello (curve di livello) dell'energia, che denotiamo

$$M_h = \{(p, q) \in \mathbb{R}^2 : H(p, q) = h\}. \quad (2.2)$$

Più precisamente,  $M_h$  risulta essere una curva chiusa regolare se  $\nabla H$  non si annulla mai su  $M_h$  stessa. Per questo vanno esclusi i valori di energia “eccezionali” in corrispondenza ai quali vi sono punti singolari: nell'esempio, il valore  $h = -\omega^2$  (la curva di livello degenera nel punto di equilibrio stabile), e il valore  $h = +\omega^2$  (la curva di livello è spigolosa nei punti di equilibrio instabili, e inoltre è decomposta in moti diversi).



**Figura 2.1.** Il ritratto in fase del pendolo; la variabile  $\tau$ .

Consideriamo, per fissare le idee, il caso dei moti di librazione, ovvero i moti interni alle separatrici; il caso delle rotazioni presenta variazioni minime (e anzi, è anche più semplice, grazie al fatto che la velocità non si annulla mai). Ricordando che si ha  $\dot{q} = p$ , e inoltre, a  $h$  fissato,

$$p = \pm P(h, q) = \pm \sqrt{2(h - V(q))},$$

il periodo  $T = T(h)$  dei moti si ottiene con una quadratura, precisamente si ha

$$T(h) = 2 \int_{q^-(h)}^{q^+(h)} \frac{dx}{\sqrt{2(h - V(x))}}, \quad (2.3)$$

ove  $q^\pm$  denotano i punti di arresto (le intercette di  $M_h$  con l'asse  $q$ ). Ancora con una quadratura si determina la legge oraria: per ogni dato iniziale  $(p_0, q_0)$ , posto  $h_0 = H(p_0, q_0)$ , si ha infatti (in un tratto interno all'intervallo tra due punti di arresto)

$$t = \int_{q_0}^q \frac{dx}{\pm \sqrt{2(h_0 - V(x))}} \quad (2.4)$$

(il segno è quello di  $p$ ), e un'inversione fornisce poi  $q$  come funzione di  $t$ .

Si vede così che nel pendolo, ma in generale in ogni sistema hamiltoniano a un grado di libertà, la presenza di un integrale del moto (l'hamiltoniana stessa) consente di "risolvere per quadrature" le equazioni del moto, ovvero di riportare un problema di integrazione di equazioni differenziali, "difficile", alle operazioni "elementari" di integrazione ordinaria e inversione. Questo fatto è già molto significativo,<sup>10)</sup> ma si può fare di più, precisamente si possono introdurre nel sistema coordinate canoniche "adattate", nelle quali il moto appare particolarmente semplice.

Una prima coppia di coordinate che è spontaneo considerare è la coppia energia-tempo  $(h, \tau)$ , precisamente

$$\begin{aligned} h(p, q) &= H(p, q) \\ \tau(p, q) &= \text{tempo di arrivo a } (p, q) \in M_h, \\ &\text{partendo da } q_0 = 0, p_0 = +\sqrt{2(h(p, q) - V(0))} ; \end{aligned}$$

$\tau$  è in realtà funzione multivoca di  $(p, q)$ , definita a meno di multipli del periodo  $T(h(p, q))$ , e viceversa il punto  $(p, q)$  dipende periodicamente da  $\tau$  per ogni  $h$ . L'espressione analitica di  $\tau$  come funzione di  $p$  e  $q$  è data da un integrale significativamente analogo a (2.4): precisamente (a meno di multipli del periodo  $T(h(p, q))$ )

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

La canonicità della trasformazione da  $(p, q)$  a  $(h, \tau)$  si verifica facilmente, ad esempio osservando che si ha<sup>11)</sup>  $\{\tau, h\} = 1$ , o equivalentemente facendo uso della funzione generatrice in variabili miste

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

Nelle nuove coordinate canoniche la nuova hamiltoniana, che denotiamo  $H'$ , è semplicemente

$$H'(h, \tau) = h ;$$

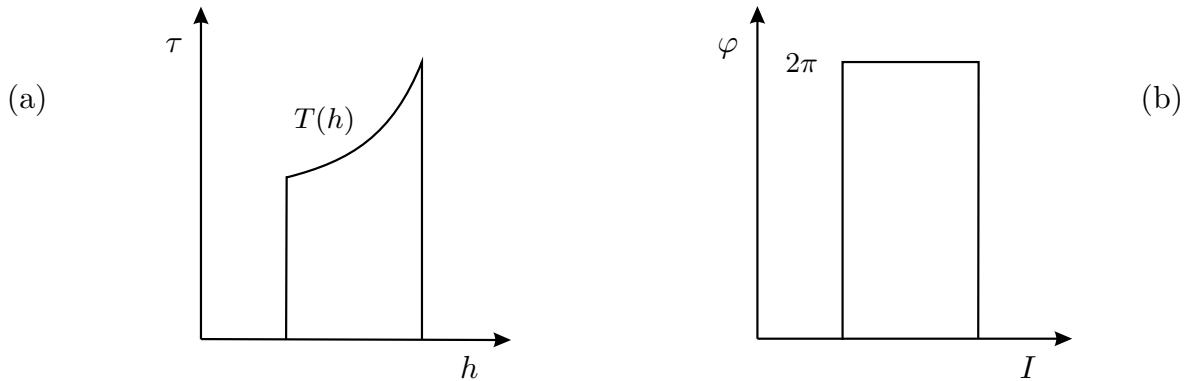
<sup>10)</sup> La speranza ultima della meccanica ottocentesca era proprio la riduzione di ogni sistema alle quadrature; la stessa equazione di Hamilton-Jacobi va pensata, almeno negli intenti di Jacobi, come un grosso sforzo in questo senso. Fu Poincaré a capire che bisognava andare oltre questa concezione troppo angusta.

<sup>11)</sup> Conviene osservare preliminarmente che la dipendenza di  $\tau$  da  $p$  e  $q$  attraverso  $h(p, q)$  (a denominatore della funzione integranda) non contribuisce alla parentesi di Poisson di  $\tau$  con  $h$ .

corrispondentemente (come si sarebbe potuto dire anche *a priori*) le nuove equazioni del moto sono

$$\dot{h} = 0, \quad \dot{\tau} = 1.$$

L'hamiltoniana  $H'$  ha in realtà una complicazione nascosta, dovuta al fatto che il dominio della variabile  $\tau$  dipende da  $h$  (il dominio complessivo di  $H'$  ha la forma illustrata in figura 2.1a). In effetti,  $\tau$  è sostanzialmente un angolo, perchè  $p$  e  $q$  dipendono periodicamente da  $\tau$ , ma la periodicità, come mostra la (2.3), dipende da  $h$ .



**Figura 2.2.** (a) Il dominio delle variabili  $h, \tau$ ; (b) il dominio delle variabili  $I, \varphi$ .

Per avere un “vero” angolo, con periodicità fissata  $2\pi$ , dobbiamo introdurre la variabile  $\varphi$  definita da

$$\varphi(h, \tau) = 2\pi \frac{\tau}{T(h)}, \quad (2.5)$$

e completare canonicamente la trasformazione. Si verifica facilmente che la buona variabile coniugata a  $\varphi$  è funzione della sola  $h$ , ed è data dall’area sottesa dalla curva  $M_h$ , divisa per  $2\pi$ :

$$I(h) = \frac{1}{2\pi} \oint_{M_h} p dq = \frac{1}{\pi} \int_{q^-(h)}^{q^+(h)} \sqrt{2(h - V(x))} dx. \quad (2.6)$$

Per verificare la canonicità del cambiamento di variabili, è sufficiente osservare che si ha

$$\frac{\partial I}{\partial h} = \frac{1}{\pi} \int_{q^-(h)}^{q^+(h)} \frac{dx}{\sqrt{2(h - V(x))}} = \frac{T(h)}{2\pi} \quad (2.7)$$

(gli estremi, pur dipendenti da  $h$ , non contribuiscono alla derivata perchè in essi l’integrando si annulla), e quindi

$$\{\varphi, I\} = \frac{\partial \varphi}{\partial \tau} \frac{\partial I}{\partial h} = \frac{2\pi}{T(h)} \frac{T(h)}{2\pi} = 1.$$

La nuova hamiltoniana ha ora la forma

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

ove  $h(I)$  è la funzione che inverte la (2.6), ed è definita in un dominio rettangolare (cioè con condizioni indipendenti sulle due variabili; in questo caso un cilindro), del tipo

$$(I, \varphi) \in \mathcal{B} \times \mathbb{T}^1, \quad \mathcal{B} \subset \mathbb{R}, \quad (2.8)$$

come illustrato in figura 2.2b. Come ci si aspetta, le equazioni del moto hanno ora la forma

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

con  $\omega = \frac{\partial h}{\partial I} = \left(\frac{\partial I}{\partial h}\right)^{-1} = 2\pi/T$ . Le variabili  $(I, \varphi)$  sono dette *variabili di azione–angolo* del sistema.

- **Esercizio.** Si verifichi che l'espressione (2.6) di  $I$  è obbligata (a meno di una banale costante additiva), se si impone che la forma differenziale  $I d\varphi - p dq$  sia chiusa (cioè che la trasformazione complessiva da  $(p, q)$  a  $(I, \varphi)$  sia canonica), che  $I$  sia funzione della sola  $h$ , e che  $\varphi$  abbia l'espressione (2.5). [*Suggerimento:  $\oint (I d\varphi - p dq) = \text{cost}$ , eccetera.*]
- **Osservazione.** E' bene dire con chiarezza che l'interesse vero per le variabili di azione–angolo non sta nel fatto, pur innegabile, che in tali variabili risulta poi facile risolvere le equazioni del moto: è chiaro infatti che tali variabili si sono potute costruire solo perché, in realtà, già si sapevano risolvere le equazioni del moto per quadrature. Va detto piuttosto che *queste variabili si rivelano preziose, praticamente indispensabili, per continuare lo studio*, e comprendere il comportamento dei sistemi hamiltoniani *prossimi* al sistema assegnato, o *perturbati*. Ad esempio, all'hamiltoniana (2.1) possiamo aggiungere una piccola forzante, oppure possiamo introdurre un debole accoppiamento con un altro semplice sistema hamiltoniano, come un rotatore;<sup>12)</sup> in quest'ultimo caso otteniamo una hamiltoniana perturbata della forma

$$H(p, q, J, \psi) = \frac{p^2}{2} - \omega^2 \cos q + \frac{J^2}{2C} + \varepsilon V(q, \psi).$$

Non sarebbe facile condurre lo studio in queste variabili. Ben diversa è la situazione dopo l'introduzione delle variabili d'azione–angolo del pendolo: l'hamiltoniana prende infatti la forma

$$K(I, J, \varphi, \psi) = h(I) + \frac{J^2}{2C} + \varepsilon f(I, J, \varphi, \psi)$$

( $f$  in realtà non dipende da  $J$ ), e questa è una tipica hamiltoniana che si studia con successo nella *teoria delle perturbazioni*.

Un minimo di riflessione mostra che la breve trattazione qui svolta si applica, in pratica, a tutti i sistemi hamiltoniani a un grado di libertà che abbiano traiettorie chiuse: in effetti, le uniche proprietà importanti che abbiamo usato sono la presenza di un integrale del moto (l'hamiltoniana stessa), e il fatto che  $M_h$  fosse una curva regolare chiusa. Grazie a queste semplici proprietà è stato possibile sia risolvere per quadrature le equazioni del moto, sia introdurre, in un intorno di  $M_h$ , le variabili d'azione–angolo  $(I, \varphi)$ .

---

<sup>12)</sup> Il rotatore (punto materiale vincolato a un circolo, ovvero oggetto rigido che ruota attorno a un asse fisso, in assenza di forze attive) è descritto da una hamiltoniana della forma  $\tilde{H}(J, \psi) = J^2/(2C)$ ,  $C =$  momento di inerzia. Si osservi che tale sistema è già scritto in variabili di azione–angolo.

## 2.2. Il teorema di Liouville–Arnol’d: enunciato.

Il teorema di Liouville–Arnol’d si può considerare, in un certo senso, una naturale<sup>13)</sup> generalizzazione del caso sopra studiato a sistemi hamiltoniani con  $n$  gradi di libertà. Una generalizzazione banale si ha, ovviamente, nel caso di  $n$  sistemi hamiltoniani a un grado di libertà tra loro non interagenti, ciascuno del tipo sopra considerato:

$$H(p, q) = \sum_i F_i(p_i, q_i) , \quad (p, q) = (p_1, \dots, p_n, q_1, \dots, q_n) \in M \subset \mathbb{R}^n .$$

In effetti, la presenza di  $n$  integrali del moto (le  $n$  hamiltoniane  $F_i$ ) fa sì che il moto sia confinato a una varietà  $n$ -dimensionale<sup>14)</sup>  $M_f \subset M$ , definita da

$$M_f = \{(p, q) \in M : F_i(p_i, q_i) = f_i , \quad i = 1, \dots, n\};$$

tale varietà è il prodotto delle  $n$  curve invarianti delle singole hamiltoniane, di equazione  $F_i(p_i, q_i) = f_i$ , dunque è un toro  $n$  dimensionale; procedendo come sopra, in tutto un intorno  $M_{\mathcal{F}}$  di  $M_f$ , ottenuto facendo variare  $f = (f_1, \dots, f_n)$  in un aperto (non troppo grande)  $\mathcal{F} \subset \mathbb{R}^n$ , le equazioni del moto si risolvono per quadrature, e inoltre (procedendo separatamente per ciascuna delle  $F_i$ ) si possono introdurre nel sistema coordinate di azione–angolo

$$(I, \varphi) = (I_1, \dots, I_n, \varphi_1, \dots, \varphi_n) , \quad I \in \mathcal{B} \subset \mathbb{R}^n , \quad \varphi \in \mathbb{T}^n ,$$

che danno all’hamiltoniana la forma “integrabile”

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

con (in questo caso disaccoppiato)  $h(I) = h_1(I_1) + \dots + h_n(I_n)$ . Le corrispondenti equazioni canoniche sono

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

con  $\omega = (\omega_1, \dots, \omega_n)$ ,  $\omega_i = \omega_i(I_i) = \frac{\partial h_i}{\partial I_i}$ . Lo spazio delle fasi (più precisamente, il dominio  $\mathcal{D}$  sul quale siamo riusciti a mettere l’hamiltoniana in forma integrabile) è così decomposto in *tori invarianti*, sui quali il moto è quasi periodico.

La grande intuizione di Liouville (metà ’800), completata (inizio anni ’60) dal lavoro di Arnol’d, si può così riassumere: anche nel caso interagente questo quadro continua a essere vero, se il sistema ammette  $n$  integrali del moto (indipendenti)  $F_1, \dots, F_n$ , purché *gli  $n$  flussi  $\Phi_1, \dots, \Phi_n$ , che si ottengono prendendo  $F_1, \dots, F_n$  come funzioni di Hamilton, commutino tra loro*: ovvero sia

$$\Phi_i^{\tau_i} \circ \Phi_j^{\tau_j} = \Phi_j^{\tau_j} \circ \Phi_i^{\tau_i}$$

per ogni scelta di  $i, j$  e  $\tau_i, \tau_j$ . Tecnicamente, la condizione che assicura la commutazione dei flussi è che le corrispondenti hamiltoniane abbiano mutua parentesi di Poisson nulla:

$$\{F_i, F_j\} = 0 \quad \forall i, j ,$$

<sup>13)</sup> Ma tutt’altro che banale!

<sup>14)</sup> Più precisamente, si trova una varietà  $n$ -dimensionale se per ogni  $i$  la curva  $F_i(p_i, q_i) = f_i$  non contiene punti singolari di  $F_i$ ; in tal caso la matrice  $n \times 2n$   $\partial F_i / \partial (p_j, q_j)$  ha rango  $n$ .

ovvero (paragrafo 1.4) siano a due a due in involuzione.<sup>15)</sup> Comunemente si prende una delle  $F_i$  coincidente con l'hamiltoniana  $H$ .

Per formulare il teorema, ci mettiamo nella situazione semplice in cui una sola carta copre una regione interessante dello spazio delle fasi, e ragioniamo in coordinate; il nostro spazio delle fasi sarà dunque un aperto  $M \subset \mathbb{R}^{2n}$ , munito di coordinate canoniche  $(p, q) = (p_1, \dots, p_n, q_1, \dots, q_n)$ .<sup>16)</sup> Una possibile formulazione del teorema di Liouville–Arnol'd è la seguente:

**Proposizione 2.1** (Teorema di Liouville–Arnol'd). *Sia  $M$  un aperto di  $\mathbb{R}^{2n}$  munito di coordinate canoniche  $(p, q) = (p_1, \dots, p_n, q_1, \dots, q_n)$ , e siano  $F_1, \dots, F_n, F_i : M \rightarrow \mathbb{R}$ , funzioni regolari a due a due in involuzione:*

$$\{F_i, F_j\} = 0 \quad \forall i, j. \quad (2.9)$$

Posto  $f = (f_1, \dots, f_n) \in \mathbb{R}^n$ , denotiamo

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

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

$$\text{rango} \frac{\partial(F_1, \dots, F_n)}{\partial(p_1, \dots, p_n, q_1, \dots, q_n)} = n \quad (2.10)$$

in ogni punto di  $M_{f^*}$ . Allora:

- i.  $M_{f^*}$  è una varietà  $n$ -dimensionale, invariante per ciascuno degli  $n$  flussi hamiltoniani  $\Phi_1^t, \dots, \Phi_n^t$  associati alle funzioni di Hamilton  $F_1, \dots, F_n$ ;
- ii.  $M_{f^*}$ , se connessa<sup>17)</sup> e compatta, è diffeomorfa al “toro standard”  $\mathbb{T}^n = \mathbb{R}^n / (2\pi\mathbb{Z})^n$ ; se non compatta, e ugualmente i flussi  $\Phi_1^t, \dots, \Phi_n^t$  sono prolungabili a tutto  $\mathbb{R}$ , è diffeomorfa al cilindro  $\mathcal{C}^{k, n-k} = \mathbb{T}^k \times \mathbb{R}^{n-k}$ , per un valore di  $k$  tra 0 e  $n$ ;
- iii. nel caso compatto connesso, esiste un intorno  $\mathcal{F} \subset \mathbb{R}^n$  di  $f^*$ , tale che l'insieme

$$M_{\mathcal{F}} = \bigcup_{f \in \mathcal{F}} M_f \quad (2.11)$$

(“intorno tubolare di  $M_{f^*}$ ”) è diffeomorfo a  $\mathcal{F} \times M_{f^*}$ ; in  $M_{\mathcal{F}}$  esistono coordinate canoniche di azione–angolo

$$(I, \varphi) = (I_1, \dots, I_n, \varphi_1, \dots, \varphi_n), \quad I \in \mathcal{B} \subset \mathbb{R}^n, \quad \varphi \in \mathbb{T}^n,$$

<sup>15)</sup> Si osservi l'analogia con l'*insieme completo di osservabili che commutano*, di ben nota importanza in meccanica quantistica.

<sup>16)</sup> L'enunciato che segue, pur formalmente vero per ogni scelta di  $M$ , è interessante solo se  $M$  è abbastanza grande: precisamente, se la varietà invariante  $M_f$  che vi compare “ci sta per intero in  $M$ ”, senza incontrarne il bordo, in modo che i moti su  $M_f$  siano prolungabili per ogni  $t \in \mathbb{R}$  (nell'esempio del pendolo: il caso interessante è quello in cui  $M$  contiene almeno una curva di livello di  $H$ ).

<sup>17)</sup> Ci si può comunque restringere a una sua componente connessa.



con le variabili  $I$  funzioni invertibili delle sole  $f$ ,  $I = I(f)$ ,  $\mathcal{B} = I(\mathcal{F})$ , tali che, posto  $(p, q) = \mathcal{C}(I, \varphi)$ , il pull-back  $F'_i = F_i \circ \mathcal{C}$ ,  $i = 1, \dots, n$ , è funzione delle sole azioni, precisamente si ha

$$F'_i(I, \varphi) = f_i(I) .$$

- **Osservazione.** Nell'enunciato che abbiamo dato tutte le funzioni  $F_1, \dots, F_n$  compaiono “alla pari”, e viceversa non si fa riferimento ad alcuna hamiltoniana alla quale si è interessati in modo particolare. Come già si è ricordato, il teorema è invece frequentemente enunciato avendo in mente una particolare hamiltoniana  $H$ : in tal caso si prende una delle  $F_i$ , diciamo  $F_1$ , uguale a  $H$ ; l'ipotesi che le  $n$  funzioni  $F_1, \dots, F_n$  siano in involuzione si enuncia allora dicendo che l'hamiltoniana  $H$  ammette  $n$  integrali primi in involuzione, e corrispondentemente nella tesi si dice che esiste una trasformazione canonica  $(p, q) = \mathcal{C}(I, \varphi)$ , definita (nel caso compatto) in  $\mathcal{B} \times \mathbb{T}^n$ , che coniuga  $H$  a una hamiltoniana  $H'$  dipendente dalle sole azioni (senza curarsi delle  $F'_i$  che restano coniugate alle  $F_i$ ).
- **Osservazione.** Non appaiono nell'enunciato per non appesantirlo troppo, ma giocano un ruolo fondamentale nella dimostrazione e sono comunque interessanti, le variabili “intermedie”, corrispondenti alla coppia energia–tempo nell'esempio del pendolo. Come si vedrà, tali variabili sono date dalle  $n$  “energie”  $f_1, \dots, f_n$  associate alle  $n$  funzioni di Hamilton  $F_1, \dots, F_n$ , e dagli  $n$  “tempi”, che saranno denotati  $\tau_1, \dots, \tau_n$ , dei corrispondenti flussi. In effetti, la trattazione classica di Liouville si fermava a queste variabili, senza effettuare l'ultimo passaggio alle variabili di azione–angolo. Va anche detto che la trattazione classica sottolineava soprattutto gli aspetti locali: il punto di vista globale, e specificamente il passaggio alle variabili di azione–angolo  $(I, \varphi)$ , sono dovuti a Arnol'd.
- **Osservazione.** Come si vedrà, la costruzione delle variabili intermedie  $(f, \tau)$ , e con esse l'integrazione delle equazioni del moto, si riduce (come già si è visto nel caso del pendolo) alle operazioni “elementari” di quadratura e di inversione di funzioni; si dice comunemente che, nelle ipotesi del teorema di Liouville, l'integrazione delle equazioni differenziali è “riportata alle quadrature”. Di qui nasce la terminologia “sistema integrabile” (o “integrabile alla Liouville”).<sup>18)</sup>

### 2.3. Dimostrazione dei punti $i$ – $i$ .

La condizione (2.10) garantisce innanzitutto che  $M_{f^*}$  stessa è una sottovarietà  $n$ –dimensionale di  $M$ . Per l'ipotesi (2.9), ciascuna delle  $F_1, \dots, F_n$  è conservata dal sistema di hamiltoniana  $F_i$ ,  $i = 1, \dots, n$ , e questo basta per il primo punto.

---

<sup>18)</sup> Si rifletta su questo punto: dire che si sanno integrare le equazioni del moto è una espressione vuota di senso, se non si specifica *a priori* la classe di funzioni cui si restringe la ricerca delle soluzioni. Paradossalmente, essendovi un teorema di esistenza e unicità, le equazioni stesse definiscono in linea di principio una funzione che è soluzione, e non si può dire molto di più se non si indica in anticipo una famiglia di funzioni considerate in qualche modo “già note”. Nelle ipotesi del teorema di Liouville–Arnol'd si sa di poter esprimere la soluzione tramite funzioni costruite, a partire dalle funzioni che appaiono nelle equazioni del moto, con sole operazioni algebriche, quadrature e inversioni.

Denotiamo ora con

$$X_i = -\frac{\partial F_i}{\partial q} \cdot \frac{\partial}{\partial p} + \frac{\partial F_i}{\partial p} \cdot \frac{\partial}{\partial q}$$

il campo vettoriale associato all'hamiltoniana  $F_i$ ; per l'ipotesi (2.10) gli  $n$  campi vettoriali  $X_1, \dots, X_n$  sono linearmente indipendenti in  $M_{f^*}$ , per la (2.9) commutano tra di loro, inoltre essi sono tangenti a  $M_{f^*}$ . Quest'ultima proprietà si può vedere sia semplicemente osservando che ciascuno dei flussi  $\Phi_i$  si svolge su  $M_{f^*}$ , cosicchè  $X_i$  è necessariamente tangente a  $M_{f^*}$ , sia anche usando direttamente la (2.9): infatti,  $M_{f^*}$  è ortogonale a ciascuno degli  $n$  gradienti  $\nabla F_1, \dots, \nabla F_n$ , e d'altra parte, per la (2.9), si ha

$$X_i \cdot \nabla F_j = -\{F_i, F_j\} = 0 \quad \forall i, j .$$

Il punto *ii* allora si riporta interamente al seguente lemma (ove in luogo di  $M_{f^*}$  si discute di una generica varietà  $n$ -dimensionale  $N$ ):

**Lemma 2.2.** *Se su una varietà  $n$ -dimensionale connessa  $N$  esistono  $n$  campi vettoriali tangenti  $X_1, \dots, X_n$  linearmente indipendenti in ogni suo punto, che commutano a due a due, ed i cui flussi  $\Phi_1^t, \dots, \Phi_n^t$  sono prolungabili per  $t \in \mathbb{R}$ , allora esiste  $k$  intero,  $0 \leq k \leq n$ , tale che  $N$  stessa è diffeomorfa al cilindro*

$$C^{k, n-k} = \mathbb{T}^k \times \mathbb{R}^{n-k} .$$

- **Osservazione.** Per la successiva dimostrazione del punto *iii* serviranno, oltre all'enunciato di questo lemma, anche alcune parti essenziali della sua dimostrazione, che dunque non si può omettere neanche in prima lettura (le variabili  $\tau$  e  $\varphi$  sono introdotte qui). Questo lemma è, si può dire, il cuore del teorema di Liouville–Arnol'd.

**Dimostrazione.** Preso  $\tau = (\tau_1, \dots, \tau_n) \in \mathbb{R}^n$  qualsiasi, sia  $\Phi^\tau : N \rightarrow N$  la mappa definita da

$$\Phi^\tau = \Phi_n^{\tau_n} \circ \dots \circ \Phi_1^{\tau_1} ;$$

per la commutatività dei campi vettoriali, e dunque dei flussi ad essi associati, l'ordine in cui i diversi flussi  $\Phi_i^{\tau_i}$  sono applicati non è rilevante. Se ne deduce immediatamente che si ha

$$\Phi^0 = \text{Id} , \quad (\Phi^\tau)^{-1} = \Phi^{-\tau} , \quad \Phi^{\tau+\tau'} = \Phi^\tau \circ \Phi^{\tau'} ,$$

ovvero  $\{\Phi^\tau, \tau \in \mathbb{R}^n\}$  è un gruppo commutativo a  $n$  parametri di diffeomorfismi:  $N \rightarrow N$ ; si osservi che questa è una naturale generalizzazione della nozione di flusso. La mappa che a ogni  $x \in N$  e a ogni  $\tau \in \mathbb{R}^n$  associa  $\Phi^\tau(x) \in N$  è detta *azione del gruppo commutativo  $\mathbb{R}^n$  sulla varietà  $N$* .

Fissiamo ora un qualsiasi punto  $y$  (dato iniziale, origine) in  $N$ , e definiamo  $\Psi_y : \mathbb{R}^n \rightarrow N$ , ponendo

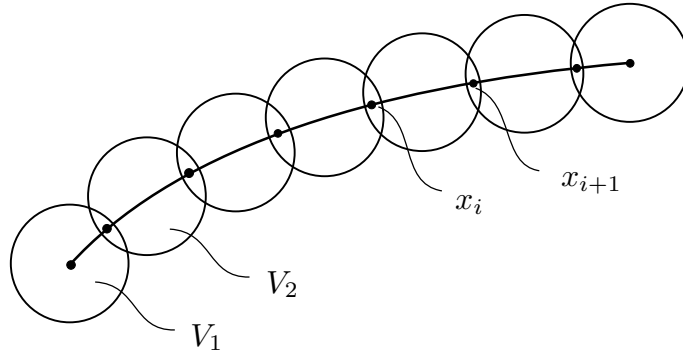
$$\Psi_y(\tau) = \Phi^\tau(y)$$

(in altre parole: ora pensiamo il punto  $y$  fisso, e  $\tau$  variabile; ma è solo un cambiamento di notazione). La mappa  $\Psi_y$  ha alcune proprietà rilevanti, di facile dimostrazione:

- a)  $\Psi_y$  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  $\tau$  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  $N$ . 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  $x'' = \Phi^{\tau'' - \tau'}(x')$ ).



**Figura 2.3.** Gli aperti  $V_1, \dots, V_l$  ricoprono  $\gamma$ .

- b) L'applicazione  $\Psi_y : \mathbb{R}^n \rightarrow N$  è *suriettiva*. Infatti, preso  $x \in N$ , sia  $\gamma$  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  $\gamma$ , ne basta un numero finito  $V_1, \dots, V_l$  a ricoprire  $\gamma$  stessa. Si costruisce allora facilmente (si veda la figura 2.3) una catena finita

$$y = x_0, x_1, \dots, x_l = x,$$

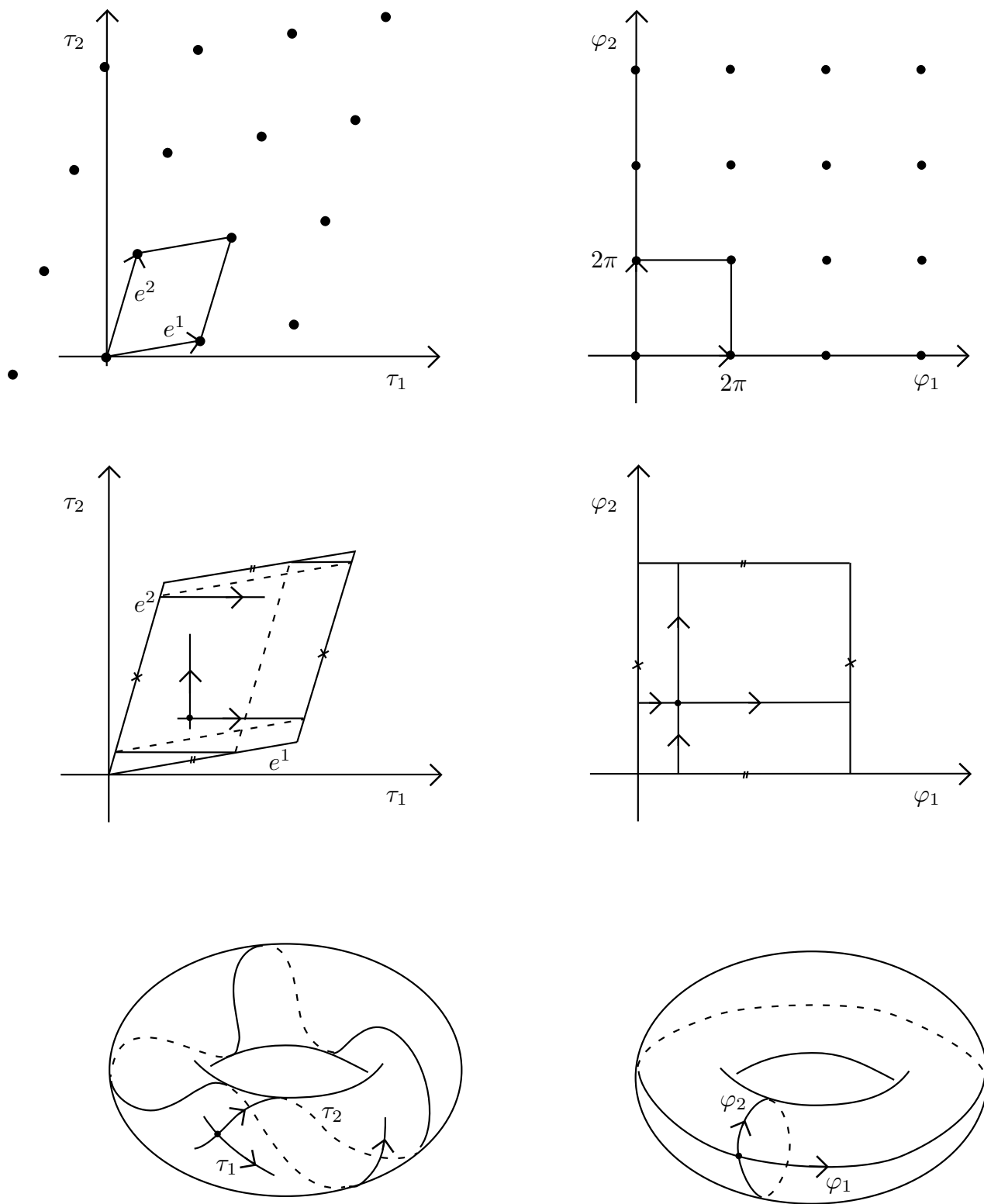
con  $x_{j-1}$  e  $x_j$  nello stesso aperto  $V_j$ ; per il punto precedente si trovano allora  $\tau^1, \dots, \tau^l$ , tali che  $x_j = \Phi^{\tau^j}(x_{j-1})$ , e dunque, posto  $\tau = \tau^1 + \dots + \tau^l$ , si ha  $\Phi^\tau(y) = x$ , ovvero  $x = \Psi_y(\tau)$ .

- c) Se  $\Psi_y$  non è *biiettiva* (altrimenti il lemma è già dimostrato, con  $N$  diffeomorfo a  $\mathbb{R}^n$  ovvero con  $k = 0$ ), allora *l'insieme*  $\mathcal{T} \subset \mathbb{R}^n$  *dei periodi di*  $y$ , *ovvero l'insieme*

$$\mathcal{T} = \{ \tau \in \mathbb{R}^n : \Phi^\tau(y) = y \},$$

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  $\Psi_y$  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  $\tau', \tau''$  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  $N$ , se  $\tau$  è periodo di  $y$  allora è anche periodo di  $x$  (e di conseguenza anche viceversa). Infatti, se  $x = \Phi^\sigma(y)$ , allora

$$\Phi^\tau(x) = \Phi^\tau(\Phi^\sigma(y)) = \Phi^\sigma(\Phi^\tau(y)) = \Phi^\sigma(y) = x.$$



**Figura 2.4.** Le variabili  $\tau$  e le variabili  $\varphi$ .

Identifichiamo ora valori diversi di  $\tau$  che differiscano per un periodo, ovvero prendiamo il quoziente  $\tilde{N} = \mathbb{R}^n / \mathcal{T}$ . Come mostra il diagramma, resta allora definita l'applicazione quoziente  $\tilde{\Psi}_y = \Psi_y / \mathcal{T} : \tilde{N} \rightarrow N$ , che per quanto sopra detto risulta essere un diffeomorfismo.

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\Psi_y} & N \\ \downarrow / \mathcal{T} & \nearrow \tilde{\Psi}_y \text{ (difeo)} & \\ \tilde{N} & & \end{array}$$

Studiamo allora  $\tilde{N}$ . Per questo ci viene in aiuto un lemma di natura algebrica, la cui dimostrazione, poco interessante per noi, è riportata nell'appendice B.1.

**Lemma 2.3.** *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:*

$$\mathcal{T} = \left\{ \sum_{i=1}^k m_i e^i, m_1, \dots, m_k \in \mathbb{Z} \right\}.$$

Ammessso questo lemma, la conclusione del lemma 2.2 è immediata: 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  $2\pi$ , mentre  $\varphi_{k+1}, \dots, \varphi_n$  sono nulli: l'elemento di matrice  $T_{ij}$  ci dice precisamente di quanto avanza  $\tau_i$ , quando  $\varphi_j$ ,  $j \leq k$ , avanza di  $2\pi$ . La matrice  $T$  è detta *matrice dei periodi*. Il gruppo, 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  $\mathcal{C}^{k, n-k}$ , corrisponde all'operazione di quoziente  $\mathbb{R}^n / \mathcal{T}$ , che dà  $\tilde{N}$ . Resta allora ben definito il diffeomorfismo  $\tilde{T}$  che appare nel diagramma

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{\frac{1}{2\pi}T} & \mathbb{R}^n & \xrightarrow{\Psi_y} & N \\ \downarrow / (2\pi\mathbb{Z})^k & & \downarrow / \mathcal{T} & \nearrow \tilde{\Psi}_y & \\ \mathcal{C}^{k, n-k} & \xrightarrow{\tilde{T}} & \tilde{N} & & \end{array}$$

e con esso il diffeomorfismo  $\tilde{\Psi}_y \circ \tilde{T}$ , che manda il cilindro standard in  $N$ , come asserito nel lemma. Se  $N$  è compatta, si ha evidentemente  $k = n$ , e il cilindro altro non è che il toro  $\mathbb{T}^n$ . La figura 2.4 illustra la situazione per  $k = n = 2$ . ■

## 2.4. Conclusione della dimostrazione

Proseguiamo con la dimostrazione del teorema, venendo al punto *iii*. Quello che dobbiamo fare per andare avanti è passare dallo studio del singolo toro  $M_{f^*}$  allo studio di un intorno tubolare  $M_{\mathcal{F}}$  di  $M_{f^*}$ , del tipo (2.11). A questo scopo prendiamo, in un primo momento ad arbitrio, una superficie  $n$ -dimensionale  $\Sigma$  trasversa a  $M_{f^*}$  in un suo punto fissato  $y^*$ . Se l'intorno  $\mathcal{F}$  di  $f^*$  è piccolo, allora<sup>19)</sup>  $\Sigma$  interseca ogni superficie  $M_f$  in uno e un solo punto  $y = y(f)$ , ovvero  $\Sigma$  stessa ammette la forma parametrica  $y = y(f)$ , e inoltre (sempre se  $\mathcal{F}$  è abbastanza piccolo) la (2.10) è soddisfatta in tutto  $M_{\mathcal{F}}$ . Possiamo allora ripetere la costruzione precedente su ciascuna delle  $M_f$ , ottenendo così buone coordinate  $(f, \tau)$  in  $M_{\mathcal{F}}$ . Precisamente, se denotiamo con  $\mathcal{S}$  la striscia (il cilindro pieno)  $\mathcal{F} \times \mathbb{R}^n$ , quello che otteniamo è una mappa suriettiva

$$\mathcal{W} : \mathcal{S} \rightarrow M_{\mathcal{F}} , \quad (f, \tau) \mapsto (p, q) ,$$

che è periodica in  $\tau$  e ha come periodi gli elementi del gruppo  $\mathcal{T}(f)$ . Si ha evidentemente  $\mathcal{W}(f, \tau) = \Phi^\tau(y(f))$ , e grazie a questa espressione non è difficile verificare che  $\mathcal{W}$  è un diffeomorfismo locale.

- **Esercizio.** Si dimostri che  $\mathcal{W}$  è un diffeomorfismo locale. [*Suggerimento:* per  $\tau = 0$  la matrice jacobiana è

$$\frac{\partial \mathcal{W}}{\partial (f, \tau)}(f, 0) = \left( \frac{\partial y}{\partial f_j}(f) \mid X_j(y(f)) \right) ,$$

e il determinante è diverso da zero per la trasversalità di  $M_{f^*}$  con  $\Sigma$ ; per  $\tau$  qualsiasi si trova

$$\frac{\partial \mathcal{W}}{\partial (f, \tau)}(f, \tau) = D\Phi_{y(f)}^\tau \frac{\partial \mathcal{W}}{\partial (f, \tau)}(f, 0) ,$$

ove  $D\Phi_y^\tau$  denota l'applicazione tangente a  $\Phi^\tau$  in  $y$ .]

Quozientando per ogni  $f$  con  $\mathcal{T}(f)$ , che come è facile vedere dipende in modo regolare da  $f$ , si ottiene allora un diffeomorfismo  $\tilde{\mathcal{W}}$  da  $\tilde{\mathcal{S}} := \mathcal{S}/\mathcal{T}$  a  $M_{\mathcal{F}}$ , come illustrato in figura 2.5.

- **Esercizio.** Si verifichi che la matrice dei periodi dipende in modo regolare da  $f$ . [*Suggerimento:* ogni periodo  $\tau$  è soluzione dell'equazione

$$\delta(f, \tau) = 0 , \quad \delta(f, \tau) = \Phi^\tau(y(f)) - y(f) .$$

Si applica allora il teorema della funzione implicita, ottenendo  $\tau = \tau(f)$ , dopo aver osservato che risulta  $\det(\partial\delta/\partial\tau) \neq 0$ .]

Per ciascun punto  $f \in \mathcal{F}$  si possono poi introdurre le coordinate  $\varphi$  al posto delle  $\tau$ , portando così a  $2\pi$  i periodi e rendendo rettangolari i domini; si ottiene una mappa di  $\mathcal{F} \times \mathbb{R}^n$  sopra  $M_{\mathcal{F}}$ , che con l'operazione di quoziente genera un diffeomorfismo  $\mathcal{F} \times \mathbb{T}^n \rightarrow M_{\mathcal{F}}$ , come illustrato in figura 2.6.

---

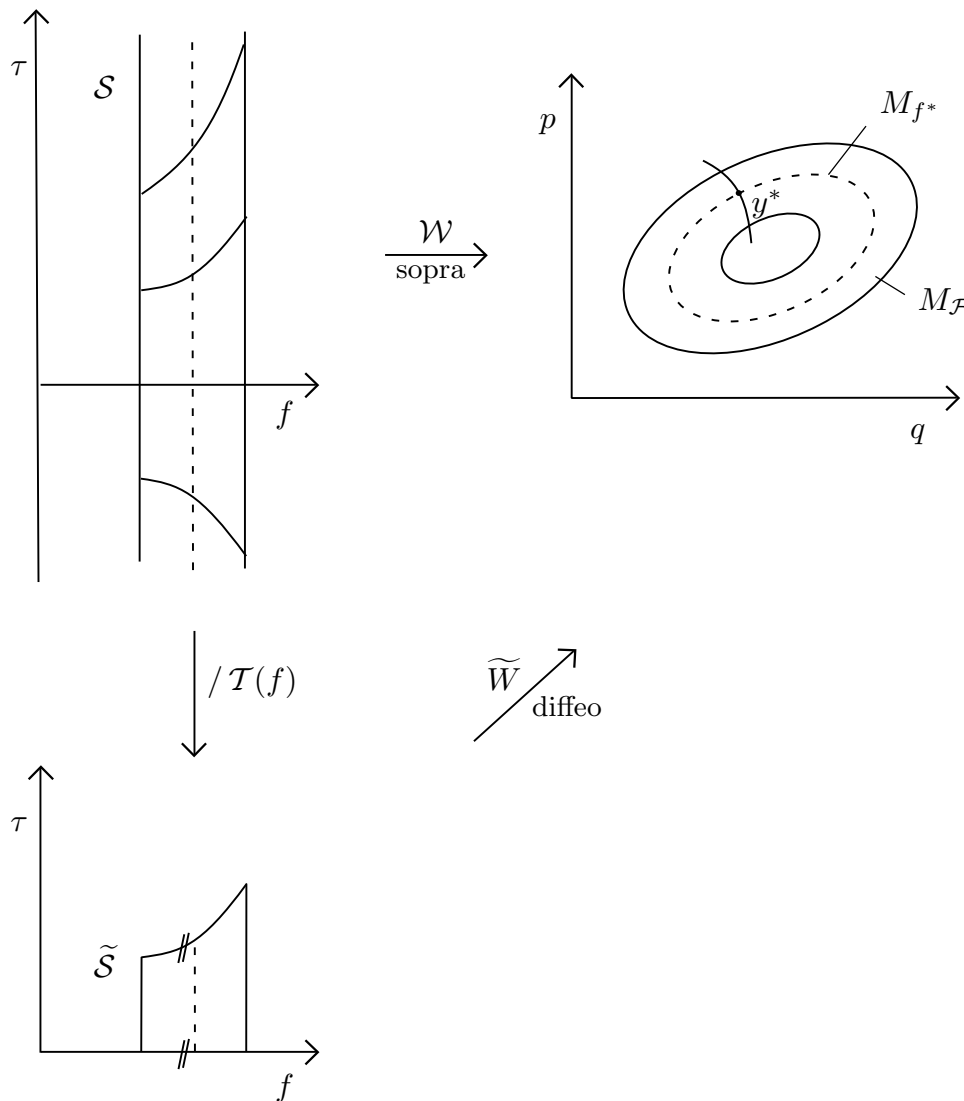
<sup>19)</sup> Siano  $\sigma_i(x) = 0$ ,  $i = 1, \dots, n$ , le equazioni di  $\Sigma$ ; ricordando che le equazioni di  $M_f$  hanno la forma  $F_i(x) = f_i$ , la condizione di trasversalità delle due superfici in  $y^*$  si scrive

$$\det \frac{\partial (F, \sigma)}{\partial x}(y^*) \neq 0 .$$

Per inversione si ottiene localmente  $x = \xi(f, s)$ , tale che  $F_i(\xi(f, s)) = f_i$ ,  $\sigma_i(\xi(f, s)) = s_i$ . Si pone allora  $y(f) = \xi(f, 0)$ . La trasversalità si mantiene per continuità.

A questo punto disponiamo di due coppie di coordinate,  $(f, \tau)$  e  $(f, \varphi)$ , su  $M_{\mathcal{F}}$  (multivoche come lo sono sempre gli angoli). Per concludere la dimostrazione proveremo che:

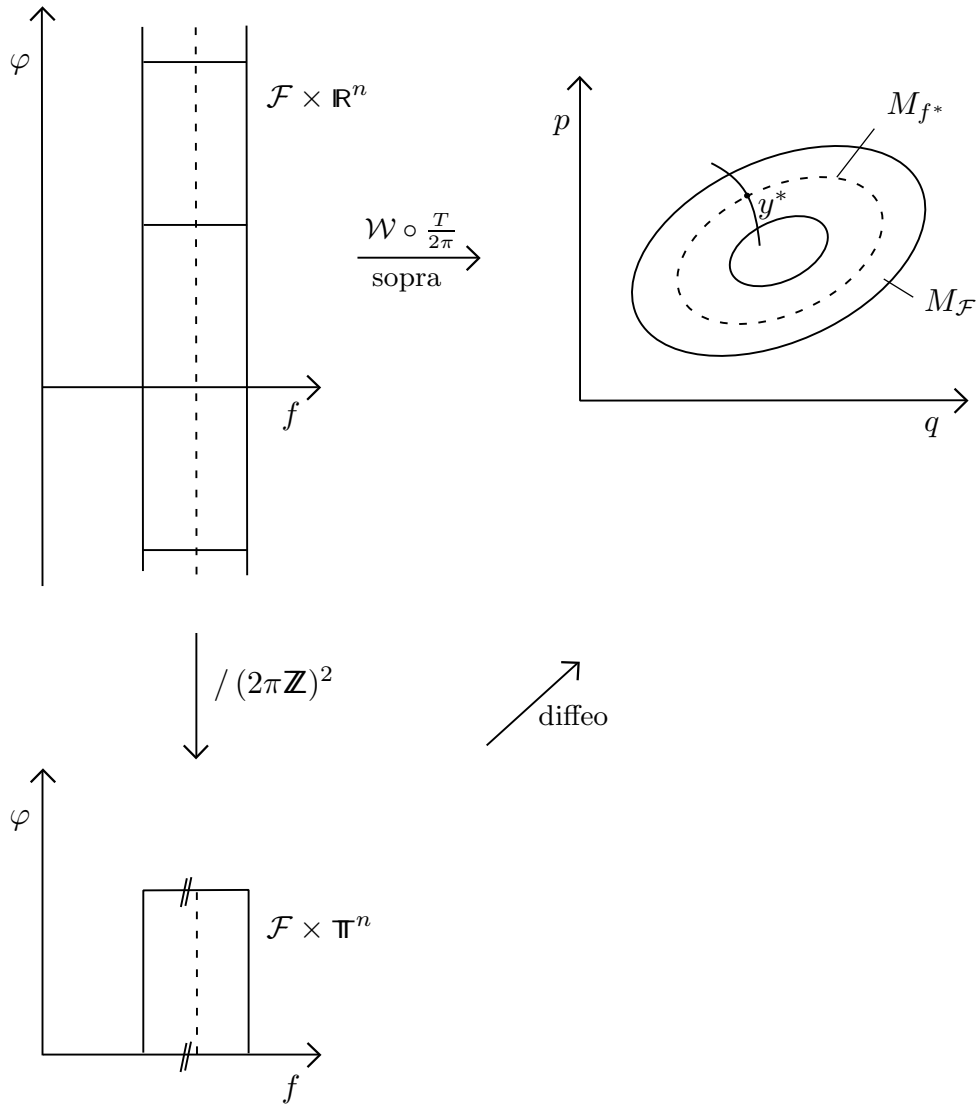
- è possibile scegliere la sezione  $\Sigma$  in modo che le coordinate  $(f, \tau)$  siano canoniche;
- è poi possibile introdurre azioni  $I$ , funzioni delle sole  $f$ , in modo che la trasformazione  $(t, \tau) \mapsto (I, \varphi)$  sia canonica.



**Figura 2.5.** Le variabili  $f, \tau$  in un intorno  $M_{\mathcal{F}}$  di  $f^*$ .

- **Osservazione.** E' chiaro che il moto hamiltoniano associato all'hamiltoniana  $F_i$  ha per immagine, nelle nuove variabili  $(f, \tau)$ , il moto banale in cui tutte le  $f_j$  sono costanti, tutte le  $\tau_j$  con  $j \neq i$  sono anch'esse costanti, mentre la sola  $\tau_i$  avanza uniformemente con velocità uno: ovvero le nuove equazioni del moto, immagine delle equazioni di Hamilton relative all'hamiltoniana  $F_i$ , sono

$$\dot{f}_j = 0 \quad \forall j, \quad \dot{\tau}_j = 0 \quad \forall j \neq i, \quad \dot{\tau}_i = 1.$$



**Figura 2.6.** Le variabili  $f, \varphi$  in un intorno  $M_{\mathcal{F}}$  di  $f^*$ .

Queste equazioni sono con evidenza le equazioni di Hamilton relative all'hamiltoniana  $\tilde{F}_i(f, \tau) = f_i$ , che a sua volta è proprio  $F_i(\mathcal{W}(f, \tau))$ , e questo vale per ogni  $i$ . Ma *il fatto che  $n$  particolari sistemi hamiltoniani siano mutati dalla trasformazione in altrettanti sistemi hamiltoniani non basta a dire che la trasformazione è canonica.*<sup>20)</sup> Come si vedrà, la canonicità di  $\mathcal{W}$  non solo va dimostrata, ma non è neanche assicurata *a priori*, e richiede una scelta particolare della sezione  $\Sigma$  (ovvero dell'origine delle coordinate  $\tau$

<sup>20)</sup> Un facile controesempio in due gradi di libertà è dato dalla trasformazione

$$\mathcal{C} : (p_x, p_y, x, y) \mapsto (p_x, p_y, x, y + p_x)$$

che non è canonica (non preserva la parentesi di Poisson), tuttavia manda ogni sistema hamiltoniano con hamiltoniana  $H$  dipendente dai soli momenti nel sistema hamiltoniano di hamiltoniana  $H' = H \circ \mathcal{C}$ , di fatto coincidente con  $H$ . Si osservi che  $\mathcal{C}$  non è canonica perché la nuova origine della coordinata  $y$  è scelta male (in modo dipendente da  $p_x$ ). *Questo è esattamente il problema che abbiamo con la nostra trasformazione  $\mathcal{W}$* : perché  $\mathcal{W}$  sia canonica bisogna prendere in modo non qualsiasi ma opportuno l'origine delle  $\tau$ , ovvero il punto  $y(f)$ , ovvero la sezione  $\Sigma$ .



per ciascuna  $f$ ).

Il prossimo passo consiste allora nel mostrare che con una scelta opportuna della sezione  $\Sigma$ , ovvero di  $y(f)$ , le coordinate  $(f, \tau)$  sono canoniche. Ciò si ottiene lavorando dapprima localmente attorno a  $y^* \in M_{f^*}$ , poi estendendo la dimostrazione. Il lavoro si articola in alcuni lemmi.

**Lemma 2.4.** *La 1-forma  $p \cdot dq$  è chiusa su ciascuna superficie  $M_f$ .*

**Dimostrazione.** Per questo è sufficiente mostrare che il differenziale esterno della 1-forma  $p \cdot dq$ , ovvero la 2-forma  $dp \wedge dq = \sum_i dp_i \wedge dq_i$ , ristretta a  $M_f$  è nulla (dà zero quando lavora su due qualsiasi vettori tangenti a  $M_f$ ). Poiché come sappiamo  $X_1(x) \dots, X_n(x)$  costituiscono una base nello spazio tangente a  $M_f$  in  $x$ , basta osservare che per ogni coppia  $i, j$  si ha

$$\begin{aligned} (dp \wedge dq)(X_i, X_j) &= \sum_l (dp_l \wedge dq_l)(X_i, X_j) \\ &= \sum_l (X_i^{p_l} X_j^{q_l} - X_i^{q_l} X_j^{p_l}) = -\{F_i, F_j\} = 0 . \end{aligned}$$

■

Una dimostrazione che non fa uso esplicito delle 2-forme è riportata nell'appendice B.2.

Proseguiamo con il

**Lemma 2.5.** *Se  $F_1, \dots, F_n$  sono in involuzione, e soddisfano la (2.10) in un punto  $x \in M$ , allora attorno a  $x$  si possono scegliere coordinate canoniche locali  $(p, q)$ , tali che risulti*

$$\det \left( \frac{\partial F_i}{\partial p_j}(x) \right) \neq 0 . \quad (2.12)$$

La dimostrazione, poco interessante, è riportata nell'appendice B.3. Dalla (2.12) segue subito che la particolare superficie  $\Sigma$  di equazione  $q_i = q_i^*$ ,  $i = 1, \dots, n$ , è trasversa a  $M_{f^*}$  in  $y^* = (p^*, q^*)$ . Inoltre, se denotiamo con

$$p_i = P_i(f, q) , \quad i = 1, \dots, n ,$$

le funzioni che che invertono localmente  $F_1, \dots, F_n$  rispetto a  $p_1, \dots, p_n$ , ovvero

$$F_i(P(f, q), q) = f_i , \quad i = 1, \dots, n , \quad (2.13)$$

l'intersezione tra  $\Sigma$  e  $M_f$  è

$$y(f) = (P(f, q^*), q^*)$$

(e questa è una rappresentazione parametrica di  $\Sigma$ ).

Verifichiamo allora innanzitutto la canonicità del cambiamento locale di coordinate.

**Lemma 2.6.** *La funzione generatrice in variabili miste*

$$S(f, q) = \int_{q^*}^q P(f, q') \cdot dq' ,$$

con integrale lungo un cammino qualsiasi su  $M_f$  da  $q^*$  a  $q$ , è ben definita in un intorno di  $(f^*, q^*)$ , e genera implicitamente una trasformazione canonica locale  $\mathcal{Z} : (p, q) \mapsto (f, g)$ , tale che la variabile  $g$  coniugata a  $f$  coincide, ove definita, con  $\tau$  (ovvero  $\mathcal{Z}^{-1}$  e  $W$  coincidono).

**Dimostrazione.** Innanzitutto,  $S$  è ben definita perché, per il lemma 2.4, l'integrale non dipende dal cammino di integrazione. E' una buona funzione generatrice locale perché risulta

$$\frac{\partial^2 S}{\partial f \partial q} = \frac{\partial P}{\partial f} ,$$

e poiché  $P$  viene dall'inversione di  $F$ ,  $\frac{\partial P}{\partial f}$  ha sicuramente determinante diverso da zero. La trasformazione canonica da essa generata è implicitamente definita da

$$p = P(f, q) , \quad g = \int_{q^*}^q \frac{\partial P}{\partial f}(f, q') \cdot dq' ,$$

e si tratta di dimostrare che le coordinate  $g$  così definite coincidono con le  $\tau$ .

A questo scopo osserviamo che per la canonicità di  $\mathcal{Z}$ , denotando con  $\tilde{\Phi}^\tau$  l'evoluzione hamiltoniana nelle nuove variabili, il diagramma qui sotto commuta:

$$\begin{array}{ccc} (P(f, q^*), q^*) & \xrightarrow{\Phi^\tau} & (p, q) \\ \mathcal{Z} \downarrow & & \downarrow \mathcal{Z} \\ (f, 0) & \xrightarrow{\tilde{\Phi}^\tau} & (f, g) \end{array}$$

(la prima colonna si spiega osservando che per  $\tau = 0$  si ha  $q = q^*$ , e dunque  $g = 0$ ). D'altra parte, la trasformazione canonica  $\mathcal{Z}$  associa a ciascuna hamiltoniana  $F_i$  la nuova hamiltoniana  $\tilde{F}_i$  definita da

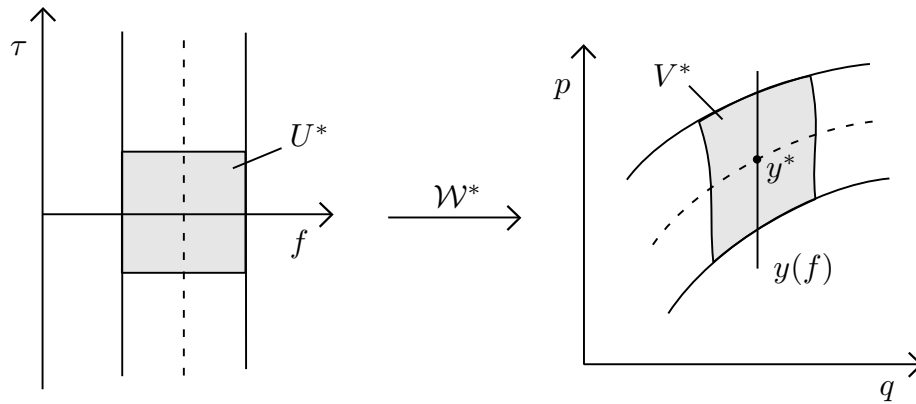
$$\tilde{F}_i(f, g) = F_i(P(f, q), q) = f_i ,$$

cosicché sotto il suo flusso le variabili  $g$  seguono le equazioni del moto  $\dot{g}_i = 1$  e  $\dot{g}_j = 0$  per  $j \neq i$ . Si ha dunque in particolare  $\tilde{\Phi}^t(f, 0) = (f, \tau)$ , dunque  $(f, t) = (f, g)$  ovvero  $g$  e  $\tau$  coincidono. ■

Grazie a questo lemma sappiamo che il cambio di coordinate  $\mathcal{Z}^{-1} : (f, \tau) \mapsto (p, q)$  manda canonicamente un intorno  $U^*$  di  $(f^*, 0)$  in un intorno  $V^*$  di  $(p^*, q^*)$ , e coincide con la restrizione  $\mathcal{W}^*$  di  $\mathcal{W}$  al dominio  $U^*$ . Il dominio  $U^*$  si può prendere ad esempio (ma non è importante) di tipo rettangolare, come indicato simbolicamente in figura 2.7; può essere necessaria una restrizione a  $\mathcal{F}$ .

Per mostrare che la mappa  $\mathcal{W} : (f, \tau) \mapsto (p, q)$  è ovunque canonica dobbiamo ora prendere un qualsiasi punto  $(\bar{f}, \bar{\tau})$ , e mostrare che un intorno  $U$  di tale punto è inviato canonicamente nella sua immagine  $V = \mathcal{W}(U) \subset M_{\mathcal{F}}$ , che a sua volta è un intorno di  $(\bar{p}, \bar{q}) = \mathcal{W}(\bar{f}, \bar{\tau})$ . Per questo è sufficiente prendere  $U = \mathcal{A}^{\bar{\tau}}(U^*)$ , ove  $\mathcal{A}^{\bar{\tau}}$  è la traslazione  $(f, \tau) \mapsto (f, \tau + \bar{\tau})$ , e corrispondentemente  $V = \Phi^{\bar{\tau}}(V^*)$ , e considerare il diagramma

$$\begin{array}{ccc} U^* & \xrightarrow{\mathcal{W}^*} & V^* \\ \mathcal{A}^{\bar{\tau}} \downarrow & & \downarrow \Phi^{\bar{\tau}} \\ U & \xrightarrow{\mathcal{W}} & V \end{array}$$

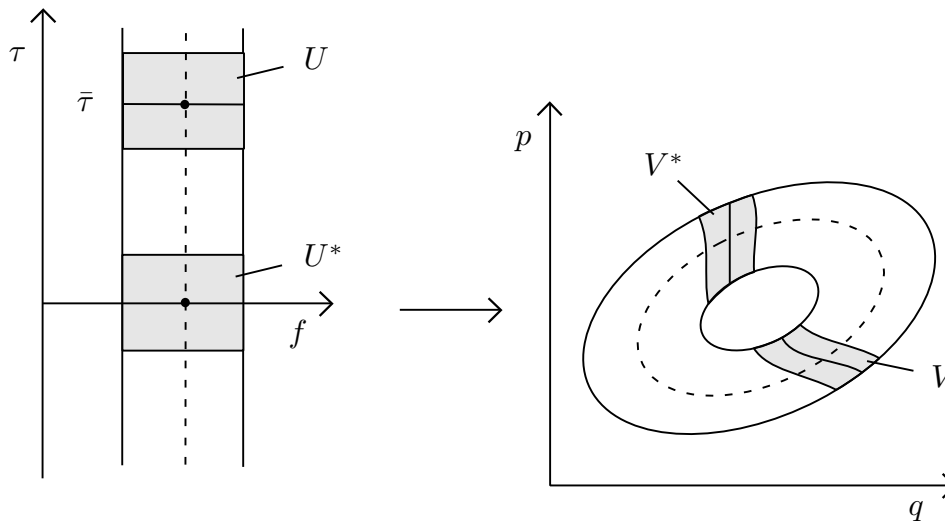


**Figura 2.7.** Gli intorni  $U^*$  e  $V^*$ .

che per la definizione stessa delle variabili  $(f, \tau)$  risulta essere commutativo. Un lato verticale del diagramma è l'evoluzione hamiltoniana  $\Phi^{\bar{\tau}}$ , che è canonica; l'altro lato verticale è la traslazione  $\mathcal{A}^{\bar{\tau}}$ , anch'essa ovviamente canonica; il lato alto  $\mathcal{W}^*$  è stato appena dimostrato essere canonico; di conseguenza è canonico anche il quarto lato, ovvero la trasformazione

$$\mathcal{W} = \Phi^{\bar{\tau}} \circ \mathcal{W}^* \circ \mathcal{A}^{-\bar{\tau}}$$

che manda  $U$  nella sua immagine  $V$ . La composizione delle trasformazioni è illustrata in figura 2.8.



**Figura 2.8.** Gli intorni  $U^*$ ,  $V^*$ ,  $U$ ,  $V$ .

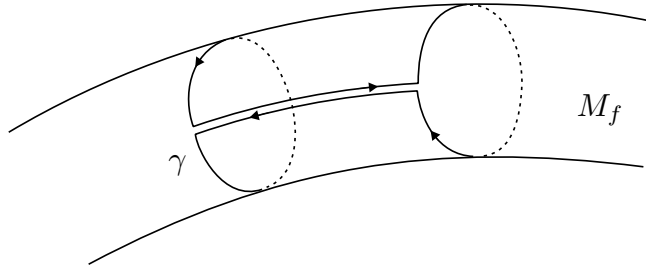
Veniamo infine alle variabili di azione–angolo. Buona parte del lavoro in realtà è già stato fatto nel momento in cui abbiamo stabilito la relazione tra  $\tau$  e  $\varphi$ ; ricordiamo che si ha  $\tau = \frac{1}{2\pi}T(f)\varphi$ , ove  $T(f)$  è la matrice dei periodi sopra introdotta. Mostriamo ora che è possibile completare canonicamente la trasformazione: precisamente, mostriamo che si possono introdurre nuove variabili  $I$  dipendenti solo da  $f$ ,  $I = I(f)$ , tali che la trasformazione

$$I = I(f), \quad \varphi = \varphi(f, \tau) = 2\pi T^{-1}(f) \tau \tag{2.14}$$

risulti canonica. La definizione buona delle  $I$  risulta essere

$$I_i(f) = \frac{1}{2\pi} \oint_{\gamma_i(f)} p \cdot dq , \quad (2.15)$$

ove  $\gamma_i$  è il circuito che si ottiene partendo da un punto qualsiasi  $x$  di  $M_f$ , e facendo variare di  $2\pi$  il solo angolo  $\varphi_i$ . Si osservi che la definizione è corretta, perchè l'integrale non dipende dal particolare  $x$  prescelto: cicli omotopi su  $M_f$  danno lo stesso risultato perchè, come si è già osservato, la 1-forma  $p \cdot dq$  è chiusa su  $M_f$ , e la differenza tra due cicli omotopi è un ciclo banale (si veda la figura 2.9).



**Figura 2.9.** Il cammino  $\gamma$  su  $M_f$

Si osservi che la (2.14), di cui dobbiamo verificare la canonicità, è una trasformazione puntuale nelle azioni, estesa agli angoli. Vale allora il

**Lemma 2.7.** *Se la trasformazione puntuale estesa (2.14), che riscriviamo in variabili miste nella forma*

$$I = I(f) , \quad \tau = \frac{1}{2\pi} T(f) \varphi ,$$

soddisfa la relazione

$$\frac{\partial I}{\partial f} = \frac{1}{2\pi} T^T(f) , \quad (2.16)$$

allora essa è canonica.

Con  $T^T$  si denota qui la trasposta di  $T$ .

**Dimostrazione.** La trasformazione è generata da

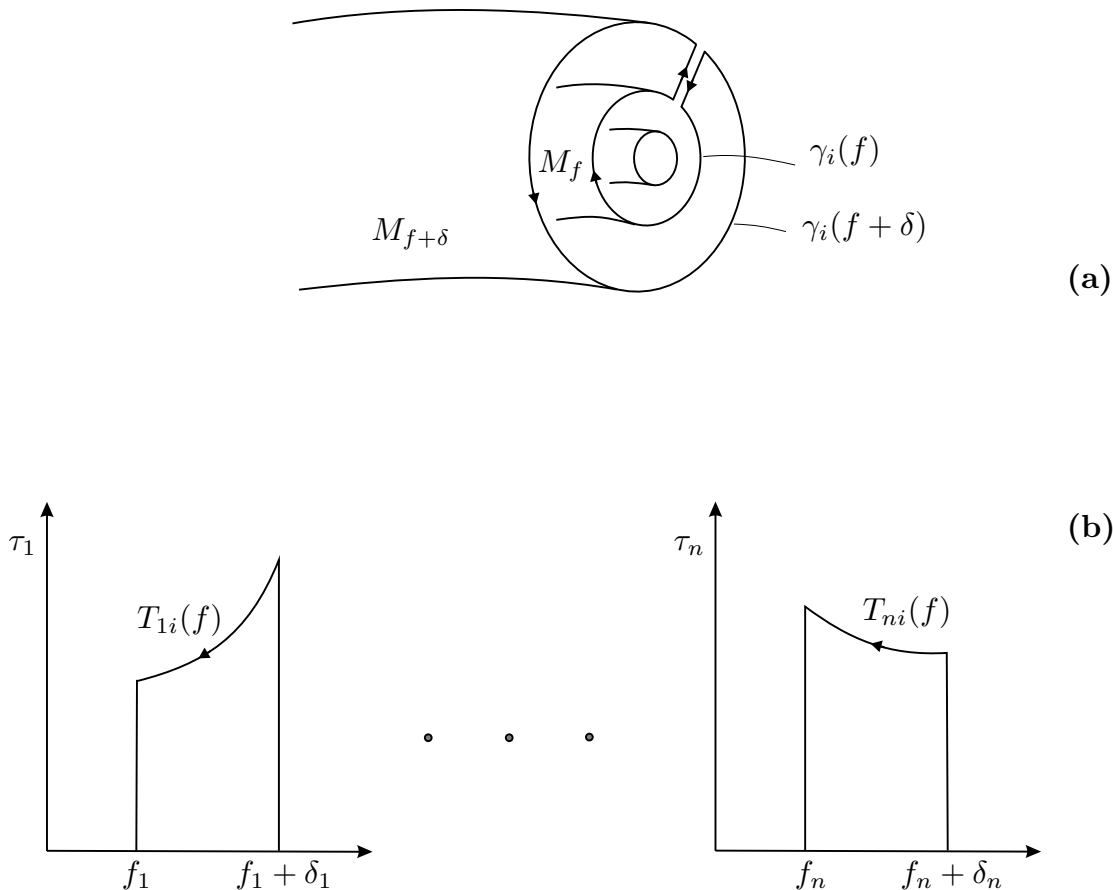
$$S(f, \varphi) = \varphi \cdot I(f) ,$$

come è immediato verificare. ■

- **Osservazione.** Si confrontino la definizione (2.15) con la (2.6), e la relazione (2.16) con la (2.7).

Per applicare questo lemma, e così praticamente concludere la dimostrazione del teorema, dobbiamo mostrare che per le  $I_1, \dots, I_n$  definite dalla (2.15) l'ipotesi (2.16) è soddisfatta. La verifica si può fare così: per un incremento  $\delta$  a  $f$  si ha

$$2\pi [I_i(f + \delta) - I_i(f)] = \oint_{\gamma_i(f+\delta)} p \cdot dq - \oint_{\gamma_i(f)} p \cdot dq , \quad (2.17)$$



**Figura 2.10.** (a) Il ciclo  $\Gamma_i^{f, f+\delta} = \gamma_i(f + \delta) - \gamma_i(f)$ ; (b) i cicli corrispondenti su ciascun piano  $f_j, \tau_j$ .

e questa differenza è equivalente a un unico ciclo  $\Gamma_i^{f, f+\delta}$  omotopo all'identità, con un qualunque percorso di andata e ritorno tra i cicli  $\gamma_i(f + \delta)$  e  $\gamma_i(f)$ ; si veda la figura 2.10a. Se uno dei due lati del percorso di andata e ritorno è a  $\tau = 0$ , l'altro è esattamente dopo che l'angolo  $\varphi_i$  è avanzato di  $2\pi$ , e corrispondentemente ciascuno dei  $\tau_j$  è avanzato del periodo  $T_{ij}$ . Il ciclo  $\Gamma_i^{f, f+\delta}$  ha dunque per immagine, nelle diverse coppie di variabili  $(f_j, \tau_j)$ , cicli  $\tilde{\Gamma}_{ij}^{f, f+\delta}$  come quelli rappresentati in figura 2.10b, ove il tratto curvo in alto è il grafico di  $T_{ij}(f)$  nel percorso da  $f$  a  $f + \delta$ . Ma per la canonicità della trasformazione  $\mathcal{W}$ , la forma differenziale

$$p \cdot dq - f \cdot d\tau$$

è chiusa, cosicché il suo integrale lungo un ciclo omotopo all'identità è nullo: si ha pertanto<sup>21)</sup>

$$2\pi [I_i(f + \delta) - I_i(f)] = \sum_{j=1}^n \oint_{\tilde{\Gamma}_{ij}^{f, f+\delta}} f_j d\tau_j = \sum_{j=1}^n \text{Area}(\tilde{\Gamma}_{ij}^{f, f+\delta}).$$

<sup>21)</sup> A questa stessa conclusione si giunge un po' più in fretta osservando che alla differenza di integrali (2.17) corrisponde l'integrale della 2-forma  $dp \wedge dq$  sull'area spazzata dal ciclo  $\gamma_i$  nella sua deformazione, e che la 2-forma è invariante per trasformazioni canoniche.

Se ora, per calcolare  $\frac{\partial I_i}{\partial f_i}$ , consideriamo variazioni  $\delta$  con una sola componente diversa da zero, diciamo la componente  $l$ , allora il solo termine  $l$ -esimo della sommatoria è non nullo, e si conclude immediatamente che si ha  $2\pi \frac{\partial I_i}{\partial f_i} = T_{li}$ , come richiesto.

Siamo giunti alla fine della dimostrazione: il dominio delle variabili  $(I, \varphi)$ , dopo l'operazione di quoziente sugli angoli, è chiaramente della forma voluta, precisamente

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

con  $\mathcal{B}$  immagine di  $\mathcal{F}$  sotto la prima delle (2.14),  $\mathcal{B} = I(\mathcal{F})$ . A sua volta ciascuna delle  $F_i$ , che sappiamo essere coniugata a  $\tilde{F}(f, \tau) = f_i$  dalla prima trasformazione, è mutata in  $F'(I, \varphi) = f_i(I)$  dalla composizione delle due trasformazioni, ove  $f(I)$  è la funzione che inverte la prima delle (2.14), e questo basta. Il teorema di Liouville–Arnol'd è così completamente dimostrato.

## 2.5. Esempi di sistemi integrabili

Non sono molti i sistemi (non banali) cui si può applicare il teorema di Liouville–Arnol'd (come sarà chiaro più avanti, la proprietà di integrabilità è da considerarsi eccezionale). Tra essi tuttavia sono compresi sistemi molto rilevanti fisicamente.

**Esempio 2.1.** Tutti i sistemi hamiltoniani a un grado di libertà. Per costruire le variabili di azione–angolo si procede esattamente come per il pendolo. Vanno escluse le curve di energia costante contenenti punti singolari.

**Esempio 2.2.** Tutti i sistemi lineari (punti materiali legati da forze elastiche). Si passa alle coordinate normali, e i sistemi si decompongono in oscillatori armonici disaccoppiati. Perché la (2.10) sia soddisfatta vanno escluse le sottovarietà dello spazio delle fasi in cui uno o più oscillatori sono a riposo.

**Esempio 2.3.** I sistemi a due gradi di libertà con una coordinata ignorabile (moto centrale, pendolo sferico). Per questi sistemi, che è spontaneo scrivere in coordinate polari  $(r, \theta)$  con relativi momenti  $(p_r, p_\theta)$ , si conservano l'hamiltoniana  $H$  e  $p_\theta$ . Non è difficile vedere direttamente che la superficie di livello di  $H$  e  $p_\theta$  è un toro  $\mathbb{T}^2$  nel caso di stati legati, e un cilindro  $\mathbb{T} \times \mathbb{R}$  nel caso di orbite aperte (a meno di valori eccezionali delle costanti, per i quali la (2.10) non è soddisfatta). Infatti si ha

$$H(p_r, p_\theta, r, \varphi) = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V(r),$$

pertanto la superficie di equazioni

$$p_\theta = P, \quad \frac{p_r^2}{2m} + W(r; P) = E, \quad \text{con } W(r; P) := V(r) + \frac{P^2}{2mr^2},$$

è il prodotto del circolo su cui corre  $\theta$  per una curva aperta o chiusa, a seconda della forma di  $V$  e del valore di  $E$ , nel piano  $r, p_r$ . La costruzione delle variabili di azione–angolo per questo problema è riportata nell'appendice C. Il caso del potenziale kepleriano

$V(r) = -\frac{\alpha}{r}$  è eccezionale perchè tutte le orbite sono chiuse: ovvero tutti i tori  $\mathbb{T}^2$  si decompongono ulteriormente in circoli, e il sistema ammette un ulteriore integrale del moto (corrispondente al fatto che il perielio del pianeta non avanza ma mantiene direzione costante; si veda, in testi classici di meccanica, il *vettore di Laplace–Runge–Lenz*). Nel caso kepleriano l'insieme dei moti con fissata energia e fissato momento angolare è l'insieme delle ellissi kepleriane con fissato semiasse maggiore e fissata eccentricità. Il toro bidimensionale corrispondente altro non è che il prodotto dell'ellisse per il circolo su cui corre  $\theta$ . Ma tali tori  $\mathbb{T}^2$  sono fittizi, e ulteriormente decomposti in circoli (le ellissi).

**Esempio 2.4.** Il moto centrale in tre dimensioni. Si conservano *quattro* quantità, precisamente l'energia e le tre componenti del momento angolare; da esse si estraggono tre integrali primi in involuzione, precisamente  $E$ ,  $M^2$  e  $M_z$  (energia, quadrato del momento angolare, componente  $z$  qualsiasi). Per valori generici delle costanti tali integrali sono indipendenti, e dunque lo spazio delle fasi si decompone in tori invarianti  $\mathbb{T}^3$ . La presenza di un ulteriore integrale del moto fa sì che i tori  $\mathbb{T}^3$  siano ulteriormente decomposti in sottovarietà, che risultano essere tori  $\mathbb{T}^2$ . Nel caso del potenziale kepleriano si ha un quinto integrale del moto (si veda ancora il vettore di Laplace–Runge–Lenz), e le orbite sono chiuse. I tori tridimensionali si ottengono dai tori bidimensionali descritti nell'esempio precedente aggiungendo il prodotto per il circolo su cui corre l'angolo  $\varphi$  coniugato a  $M_z$ . L'indipendenza degli integrali in involuzione è garantita (il calcolo è un po' fastidioso) al di fuori delle orbite circolari e di quelle con  $M$  parallelo all'asse  $z$  prefissato.

**Esempio 2.5.** Il *corpo rigido di Eulero*, ovvero il corpo rigido con un punto fisso in assenza di forze attive. Come sopra si conservano energia e momento angolare, dunque vi sono quattro costanti del moto di cui tre in involuzione,  $E$ ,  $M^2$  e  $M_z$ . La condizione (2.10) sul rango è garantita (c'è un po' da lavorare) al di fuori delle rotazioni proprie (rotazioni attorno agli assi principali di inerzia) e dei moti con  $M$  parallelo all'asse  $z$ . La costante del moto addizionale fa sì che i tori  $\mathbb{T}^3$  siano ulteriormente decomposti in tori  $\mathbb{T}^2$ . Nel caso simmetrico i due moti circolari che combinati danno un toro  $\mathbb{T}^2$  sono le precessioni sui coni di Poincot.

**Esempio 2.6.** La *trottola di Lagrange*, ovvero il corpo rigido con un punto fisso, a struttura giroscopica (ellissoide di inerzia simmetrico, baricentro sull'asse di simmetria), pesante. Qui gli integrali del moto in involuzione sono l'energia, la componente verticale del momento angolare, e la sua componente parallela all'asse del corpo. Non vi sono ulteriori integrali del moto. I tre angoli corrispondono ai tre ben noti movimenti della trottola (rotazione propria, precessione attorno alla verticale e nutazione). Vanno esclusi anche qui i moti giroscopici e i moti con  $M$  parallelo all'asse  $z$ .

**Esempio 2.7.** Sistemi molto speciali, assai difficili da scovare. Uno tra i più noti è il caso di  $n$  punti materiali su una retta, con potenziale a primi vicini  $V(r) = e^r - 1 - r$  (*sistema di Toda*). L'espressione degli integrali del moto è complicata (Hénon, 1974).

**Definizione.** Sistemi con un numero di integrali del moto superiore al numero di gradi di libertà si dicono *superintegrabili*.

Come mostrano gli esempi 2.3–2.5, i sistemi superintegrabili, pur eccezionali, giocano

un ruolo assai importante nelle applicazioni fisiche.

## 2.6. Un commento

La trasformazione canonica  $(p, q) \mapsto (I, \varphi)$  costruita nel corso della dimostrazione del teorema di Liouville–Arnol’d non è univocamente determinata, e di conseguenza non sono univocamente definite le variabili di azione–angolo. Ci sono tre ovvie libertà:

- a) si può variare l’origine degli angoli sul toro, ovvero passare da  $(I, \varphi)$  a  $(I', \varphi')$  con

$$I' = I, \quad \varphi' = \varphi + \psi(I).$$

Perchè la trasformazione sia canonica  $\psi$  non può essere qualsiasi, ma dovendo risultare

$$0 = \{\varphi'_i, \varphi'_j\} = \{\varphi_i, \psi_j\} = \frac{\partial \psi_j}{\partial I_i} - \frac{\partial \psi_i}{\partial I_j}$$

bisogna e basta che sia

$$\psi_j(I) = \frac{\partial \Psi}{\partial I_j}$$

con opportuna  $\Psi : \mathbb{T}^n \rightarrow \mathbb{R}$ . L’indeterminazione dell’origine corrisponde a una parziale libertà nella scelta della sezione  $\Sigma$ .

- b) Si può eseguire la sostituzione lineare

$$\varphi' = A\varphi, \quad I' = A^{-T}I$$

con  $A$  matrice intera,  $\det A = 1$ . L’indeterminazione corrisponde all’arbitrarietà della scelta della base nello spazio dei periodi (dei vettori di base  $e^j$ , che con le loro componenti  $(e^j)_i$  danno la matrice dei periodi  $T_{ij}$ ), ovvero degli  $n$  cicli indipendenti su ciascun toro lungo i quali, mediante la (2.15), sono calcolate le azioni.

- c) Infine si può eseguire una banale traslazione nelle azioni:

$$I' = I + c, \quad \varphi' = \varphi.$$

Mettendo insieme:

$$\varphi' = A\varphi + (\nabla \Psi)(I), \quad I' = A^{-T}I + c.$$

Si verifica facilmente che questa è la più generale trasformazione canonica che gode delle seguenti due proprietà:

- i) manda tori in tori, ovvero si ha  $I' = f(I)$ ;
  - ii) manda angoli in angoli, ovvero i nuovi angoli  $\varphi'$  sono periodici di periodo  $2\pi$  nei vecchi e viceversa.
- **Esercizio.** Si dimostri quanto sopra affermato (ad esempio imponendo la preservazione delle parentesi di Poisson).

Consideriamo ora il caso esattamente isocrono  $h(I) = \omega \cdot I$  con  $\omega \in \mathbb{R}^n$  (oscillatori armonici in variabili di azione–angolo). Dalla teoria ergodica sappiamo che se l’ $n$ -pla delle



frequenze  $\omega$  è non risonante, allora il moto  $\varphi(t) = \varphi(0) + \omega t$  ristretto al singolo toro è ergodico. Pertanto non vi sono ulteriori integrali del moto (nemmeno misurabili) e la decomposizione dello spazio delle fasi in tori  $\mathbb{T}^n$  è la più fine decomposizione in sottovarietà invarianti ammissibile. Se invece esistono  $r$  risonanze,  $1 \leq r < n$ , in corrispondenza a  $r$  vettori interi indipendenti  $k^{(1)}, \dots, k^{(r)}$ , allora ogni toro è ulteriormente decomponibile in tori di dimensione  $n - r$  (come sappiamo, con un cambiamento di coordinate ci si porta a  $\omega'$  con  $\omega'_1, \dots, \omega'_r = 0$ ). Oltre alle  $n$  azioni ci sono ora  $r$  ulteriori integrali del moto indipendenti, ad esempio  $G_s = e^{ik^{(s)} \cdot \varphi}$  per  $s = 1, \dots, r$ , e dunque il sistema è superintegrabile. Se non vi sono ulteriori risonanze la decomposizione non si può ulteriormente raffinare.

Consideriamo poi, all'estremo opposto, il caso strettamente anisocrono, precisamente il caso in cui è soddisfatta la *condizione di non degenerazione*

$$\det \left( \frac{\partial^2 h}{\partial I_i \partial I_j} \right) \neq 0. \quad (2.18)$$

L'esempio più semplice è

$$H(I, \varphi) = \sum_{j=1}^n \frac{I_j^2}{2C} + \varepsilon f(\varphi)$$

( $n$  rotatori con momento di inerzia  $C$ , tra loro debolmente accoppiati). In tal caso la mappa  $I \mapsto \omega(I) = \frac{\partial h}{\partial I}$  è un diffeomorfismo locale, e si vede facilmente che

**Proposizione 2.8.** *Se la condizione (2.18) è soddisfatta in un aperto  $\mathcal{B} \in \mathbb{R}^n$ , allora gli insiemi  $B_1, \dots, B_{n-1} \in U$ , caratterizzati dalla presenza di  $1, \dots, n-1$  risonanze indipendenti, sono tutti densi in  $\mathcal{B}$ . La loro unione ha tuttavia misura nulla.*

- **Esercizio.** Si esegua la dimostrazione.

Per i sistemi superintegrabili la condizione di non degenerazione non può mai essere soddisfatta (tutti i tori sono ulteriormente decomposti in sottovarietà di dimensione più bassa).

## 2.7. Sistemi prossimi a sistemi integrabili: l'“approssimazione zero”

Consideriamo un sistema hamiltoniano a  $n$  gradi di libertà dipendente da un parametro  $\varepsilon$ , della forma

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

sia  $\Phi_\varepsilon^t : (p, q) \mapsto \Phi_\varepsilon^t(p, q)$ ,  $t \in \mathbb{R}$ , il suo flusso. E' naturale chiedersi che relazione ci sia tra  $\Phi_\varepsilon^t$  e il flusso del sistema imperturbato  $\Phi_0^t$ . In assenza di altre informazioni su  $h$  non si può dire molto: se non che, sulla base del teorema di esistenza e unicità della soluzione per le equazioni differenziali ordinarie, nella parte di continuità al variare dei parametri, per ogni  $(p, q)$  vale una stima *a priori* del tipo

$$\|\Phi_\varepsilon^t(p, q) - \Phi_0^t(p, q)\| < C e^{\lambda|t|} \varepsilon, \quad (2.19)$$

ove  $\lambda$  è una costante di Lipshitz e  $C$  una opportuna costante positiva (la scelta della norma non è rilevante). L'esponenziale in  $t$  in generale non è migliorabile.

- **Esercizio.** Si mostri che per  $n = 1$  e

$$H_\varepsilon(p, q) = \frac{1}{2}p^2 + (1 + \varepsilon)V(p, q), \quad V(p, q) = -\frac{1}{2}\omega^2 q^2$$

(più in generale, per potenziali che hanno un massimo) la stima esponenziale (2.19) non è migliorabile per  $(p, q)$  prossimo al punto di equilibrio instabile.

La (2.19) va letta come un sostanziale limite alla capacità di previsione: in sostanza, si perde ogni informazione<sup>22)</sup> dopo tempi di ordine  $\log \varepsilon^{-1}$ .

Le cose sono sostanzialmente diverse se invece si sa che l'hamiltoniana imperturbata è integrabile. Infatti, supponendo di aver già fatto il passaggio alle variabili di azione–angolo, ovvero di avere

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

si deduce immediatamente

$$\|I_\varepsilon(t) - I_0(t)\| = \|I_\varepsilon(t) - I^o\| \leq |t| \left\| \frac{\partial f}{\partial \varphi} \right\|,$$

mentre per gli angoli si trova senza difficoltà una stima poco peggiore del tipo

$$\|\varphi_\varepsilon(t) - \varphi_0(t)\| \leq a|t| + bt^2, \quad a, b > 0.$$

Entrambe le stime, e soprattutto quella per le azioni che è l'unica veramente importante (raramente le fasi sono oggetto di interesse), sono drasticamente migliori della (2.19).

Questa semplice analisi “di ordine zero” mostra la sostanziale differenza che si ha perturbando un sistema integrabile anziché un sistema qualsiasi, e mette anche in evidenza l'interesse delle variabili di azione–angolo in questo genere di studi. Tale interesse sarà ancor più evidente negli sviluppi della teoria delle perturbazioni che affronteremo nei prossimi paragrafi.

### 3. Il principio della media

Col termine *principio della media*, o metodo della media, non ci si riferisce a un risultato ma piuttosto a un metodo di lavoro, ad un insieme di idee, che costituiscono l'anima della teoria delle perturbazioni.<sup>23)</sup> In questo paragrafo cercheremo di introdurre l'idea principale in modo critico, basandoci soprattutto su esempi.

<sup>22)</sup> Si prenda  $\omega = 10 \text{ sec}^{-1}$ , come è per un pendolo di lunghezza 10 cm linearizzato attorno alla configurazione di equilibrio instabile, e si ponga ad esempio  $\|(p, q)\| = |p/\omega| + |q|$ . Se  $\omega$  è nota con la ragguardevole precisione di 9 cifre decimali, e inizialmente  $\|(p, q)\| = d$ , allora il tempo entro il quale l'errore sulla traiettoria è contenuto a sua volta entro  $d$  è circa 2 sec; portando la precisione al valore surreale di 18 cifre decimali il tempo raddoppia soltanto.

<sup>23)</sup> Non soltanto: c'è una classe molto ampia di problemi fisici in cui si passa da una descrizione dettagliata a una su scala più macroscopica, tramite una operazione di media su variabili rapidamente fluttuanti. Si pensi a equazioni di trasporto come la legge di Ohm, in cui si trascurano le (enormi) fluttuazioni locali di campo elettrico, alla nozione stessa di pressione in un gas, e alle leggi macroscopiche dell'idrodinamica (come le equazioni di Eulero dei fluidi ideali), o infine al 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).

## 3.1. Il sistema mediato.

Consideriamo un sistema di equazioni differenziali in  $m + n$  variabili, della forma<sup>24)</sup>

$$\begin{aligned} \dot{I} &= \varepsilon F(I, \varphi) \\ \dot{\varphi} &= \omega(I) + \varepsilon G(I, \varphi) , \end{aligned} \quad (3.1)$$

con

$$I = (I_1, \dots, I_m) \in \mathbb{R}^m , \quad \varphi = (\varphi_1, \dots, \varphi_n) \in \mathbb{T}^n .$$

Il caso che più ci interessa è quello di un sistema hamiltoniano quasi integrabile a  $n$  gradi di libertà con hamiltoniana del tipo

$$H(I, \varphi) = h(I) + \varepsilon f(I, \varphi) ,$$

ovvero il caso  $m = n$  con

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

Il discorso che andiamo a fare però non solo è più generale del caso hamiltoniano, ma anzi, acquista forse chiarezza se non si tiene conto fin dall'inizio della peculiarità delle equazioni di Hamilton, e per questo ci collochiamo, provvisoriamente, in un contesto più generale.

Ciò che caratterizza il sistema (3.1) è la distinzione “a vista” delle variabili in *lente* (le  $I$ ) e *veloci* (le  $\varphi$ ). Perché la distinzione abbia senso (posto che ne abbia: vedremo presto che occorre cautela) bisogna evidentemente supporre che  $\varepsilon$  sia piccolo. In queste condizioni è naturale scomporre  $F$  in una parte media<sup>25)</sup> e una parte oscillante:

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

e pensare che in qualche modo, se  $\varepsilon$  è piccolo, la parte oscillante  $\tilde{F}$  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 (3.1), il *sistema mediato*

$$\dot{J} = \varepsilon \mathcal{F}(J) , \quad (3.2)$$

e chiedersi che relazione vi sia tra le soluzioni  $J_\varepsilon(t)$  e  $I_\varepsilon(t)$  a parità di dato iniziale  $J^0 = I^0$ , ovvero che tipo di informazioni fornisca il sistema mediato sul sistema vero. Si osservi che il confronto diviene interessante solo su tempi  $t \sim \varepsilon^{-1}$  (per  $t \ll \varepsilon^{-1}$  né  $I$  né  $J$  si staccano dal valore iniziale comune, e dunque banalmente sono anche vicini tra loro), e che su tali tempi lo studio del sistema mediato (3.2) è *molto* più semplice di quello del sistema completo (3.1): per il sistema mediato infatti  $\varepsilon^{-1}$  è una unità naturale di tempo,

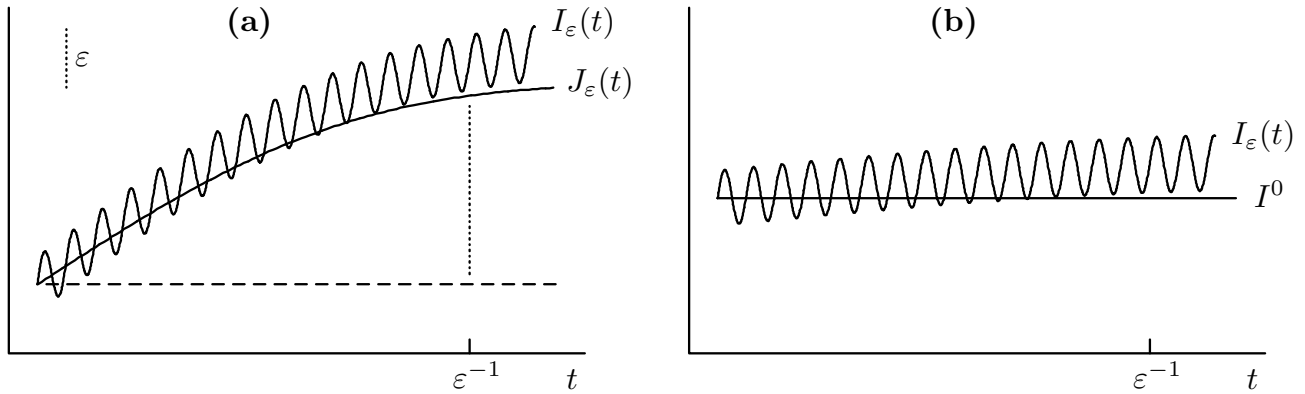
<sup>24)</sup> Non vi sarebbe alcuna difficoltà a considerare il caso, poco più generale, in cui  $F$  e  $G$  dipendono a loro volta da  $\varepsilon$ ; lo escludiamo qui solo per semplicità.

<sup>25)</sup> La notazione coerente con quella del primo capitolo (ma qui pesante) sarebbe, al posto di  $\mathcal{F}(I)$ ,  $\langle F \rangle(I)$ .

mentre per il sistema vero, a causa del moto veloce degli angoli,  $\varepsilon^{-1}$  è un tempo lungo. Ci chiediamo dunque se, ed eventualmente a quali condizioni, si ha

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

nell'intervallo  $0 \leq t \leq \varepsilon^{-1}$ . Infine, osserviamo che il sistema mediato è banale nel caso hamiltoniano:<sup>26)</sup> per i sistemi hamiltoniani si ha infatti  $F = -\frac{\partial H}{\partial \varphi}$  con  $H$  periodica in  $\varphi$ , e dunque  $\mathcal{F} = 0$ , perciò la soluzione dell'equazione mediata non è altro che  $J_\varepsilon(t) = I^0$ . Il modo appropriato di porre la questione in ambito hamiltoniano è allora di chiedersi se le azioni, che sono costanti del moto per  $\varepsilon = 0$ , per  $\varepsilon$  piccolo e  $t \sim \varepsilon^{-1}$  evolvano significativamente, o invece, a somiglianza dell'equazione mediata, restino vicine al valore iniziale  $I^0$ . La figura 3.1 illustra simbolicamente l'ipotesi di lavoro che sta dietro al principio della media, nel caso generale (a) e hamiltoniano (b).



**Figura 3.1.** L'ipotesi di lavoro su  $I_\varepsilon(t)$  e  $J_\varepsilon(t)$  secondo il principio della media, nel caso generale (a) e hamiltoniano (b).

Gli esempi che seguono mostreranno che la domanda posta è effettivamente significativa, ma ancora grossolana, perché basata su una distinzione tra variabili lente e veloci un po' troppo ingenua.

### 3.2. Esempi.

Cominciamo con un esempio totalmente elementare con  $m = n = 1$ , precisamente<sup>27)</sup>

$$\dot{I} = \varepsilon(a + b \cos \varphi), \quad \dot{\varphi} = \omega \neq 0.$$

La soluzione si trova all'istante: si ha  $\varphi(t) = \omega t + \varphi^0$ , e dunque  $\dot{I} = \varepsilon[a + b \cos(\omega t + \varphi^0)]$ ,

$$\begin{aligned} I_\varepsilon(t) &= I^0 + \varepsilon a t + \varepsilon b \int_0^t \cos(\omega t' + \varphi^0) dt' \\ &= I^0 + \varepsilon a t + \frac{\varepsilon b}{\omega} [\sin(\omega t + \varphi^0) - \sin \varphi^0]. \end{aligned}$$

<sup>26)</sup> Ci riferiamo qui al vero caso hamiltoniano, con  $H$  funzione periodica degli angoli (si confronti con il primo esempio sotto).

<sup>27)</sup> Il sistema è solo localmente hamiltoniano, con  $H = \omega I - \varepsilon(a\varphi + b \sin \varphi)$  non periodico in  $\varphi$ .

Corrispondentemente si ha l'equazione mediata  $\dot{J} = \varepsilon a$ , risolta da  $J_\varepsilon(t) = I^0 + \varepsilon at$ ; si conclude che

$$|I_\varepsilon(t) - J_\varepsilon(t)| \leq \frac{2b}{\omega} \varepsilon ,$$

addirittura uniformemente in  $t$ . Vale la pena di osservare che si è implicitamente supposto  $\omega \neq 0$ : per  $\omega = 0$  invece si trova  $I_\varepsilon(t) - J_\varepsilon(t) = \varepsilon b t \cos \varphi^0$ , e la differenza è  $\mathcal{O}(1)$  a  $t = \varepsilon^{-1}$ . La cosa non fa meraviglia: per  $\omega = 0$  l'angolo  $\varphi$  cessa di essere una variabile veloce, e dunque viene meno il presupposto di tutto il discorso.

Un esempio più interessante con  $n \geq 2$  frequenze, questa volta hamiltoniano, è dato da

$$H = \omega \cdot I + \varepsilon f(\varphi) , \quad I \in \mathbb{R}^n , \quad \varphi \in \mathbb{T}^n , \quad \omega \in \mathbb{R}^n ,$$

con

$$f(\varphi) = \sum_{k \in \mathcal{K}} \hat{f}_k e^{ik \cdot \varphi} ,$$

ove  $\mathcal{K}$  denota un sottoinsieme finito di  $\mathbb{Z}^n$ . Le equazioni di Hamilton si scrivono

$$\dot{I} = -\varepsilon \sum_{k \in \mathcal{K}} (ik \hat{f}_k) e^{ik \cdot \varphi} , \quad \dot{\varphi} = \omega$$

(si noti l'assenza del termine con  $k = 0$ , cioè della media, dalla prima equazione), e la soluzione è immediata: si ha  $\varphi(t) = \varphi^0 + \omega t$ , e poi

$$\begin{aligned} I_\varepsilon(t) &= I^0 - \varepsilon \sum_{k \in \mathcal{K}} (ik \hat{f}_k) \int_0^t e^{ik \cdot \varphi(t')} dt' \\ &= I^0 - \varepsilon t \sum_{k \in \mathcal{K}_0} (ik \hat{f}_k) e^{ik \cdot \varphi^0} + \varepsilon \sum_{k \in \mathcal{K} \setminus \mathcal{K}_0} \frac{k \hat{f}_k}{k \cdot \omega} e^{ik \cdot \varphi^0} (e^{ik \cdot \omega t} - 1) , \end{aligned}$$

ove  $\mathcal{K}_0$  (eventualmente vuoto) denota il sottoinsieme risonante:

$$\mathcal{K}_0 = \{k \in \mathcal{K} : k \cdot \omega = 0\} .$$

Si vede così che alle risonanze si accompagnano *termini secolari*, ovvero termini in cui il tempo non è più ad argomento di una funzione trigonometrica, ma appare a fattore (più in generale: appare in forma polinomiale), e dunque produce una deriva sistematica che nel corso del tempo si accumula.<sup>28)</sup> Proprio a causa di tali termini, se presenti,  $I_\varepsilon(t)$  si allontana di una quantità  $\mathcal{O}(1)$  dal dato iniziale  $I^0$  (ovvero, ricordando che nel caso hamiltoniano si ha  $J_\varepsilon(t) = I^0$ , dalla soluzione dell'equazione mediata) in un tempo  $t \sim \varepsilon^{-1}$ . In assenza di risonanze invece la deviazione dal dato iniziale è piccola per  $\varepsilon$  piccolo (ma si faccia attenzione: è cruciale qui il fatto che  $\mathcal{K}$  sia finito, altrimenti i denominatori  $k \cdot \omega$  possono essere arbitrariamente piccoli, e tutto diventa più complicato).

<sup>28)</sup> Il problema dei termini secolari associati alle risonanze ha tormentato la Meccanica Celeste fin dai tempi di Laplace.

- **Esercizio.** Si studi l'esempio precedente, con  $k \in \mathbb{Z}^n$  anziché  $k \in \mathcal{K}$ . Si cerchino condizioni su  $f_k$  e su  $\omega$  per le quali il principio della media si applica (non è facile; si ritorni qui più avanti). E' sufficiente ragionare su due soli gradi di libertà.
- **Esercizio.** Si complichino ulteriormente il problema, prendendo il sistema a due gradi di libertà

$$H(I_1, I_2, \varphi_1, \varphi_2) = \omega_1 I_1 + \omega_2 I_2 + \varepsilon [I_1 + f(\varphi_1, \varphi_2)]$$

(l'elemento nuovo significativo è lo spostamento della frequenza  $\omega_1$  con  $\varepsilon$ )

- **Esercizio.** Si studi il sistema hamiltoniano non isocrono

$$H = \frac{1}{2}(I_1^2 - I_2^2) + \varepsilon \cos(\varphi_1 + \varphi_2) , \tag{3.3}$$

e si determini una soluzione con  $I_\varepsilon(t) - I^0 = \mathcal{O}(1)$  per  $t = \varepsilon^{-1}$ .

### 3.3. Discussione.

Per comprendere in modo critico il principio della media e il ruolo delle risonanze, è utile considerare il sistema isocrono, non necessariamente hamiltoniano,

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

con  $I \in \mathbb{R}^m$ ,  $\varphi \in \mathbb{T}^n$ ,  $\omega \in \mathbb{R}^n$ . Procedendo in modo euristico, tentiamo di “dedurre” l'equazione mediata, e per questo consideriamo una scala di tempo  $T_\varepsilon$  intermedia tra 1 e  $\varepsilon^{-1}$  (ad esempio, per fissare le idee,  $T_\varepsilon = \varepsilon^{-1/2}$ ), cosicchè per  $0 \leq t \leq T_\varepsilon$  valgono le stime *a priori*

$$|I_\varepsilon(t) - I^0| \leq \varepsilon \|F\| t , \quad |\varphi_\varepsilon(t) - \varphi^0 - \omega t| \leq \varepsilon \|G\| t ,$$

avendo denotato con  $\|\cdot\|$  la sup-norma (per  $\varphi \in \mathbb{T}^n$  e  $I$  in un qualche fissato intorno del dato iniziale, nel quale il moto resta consistentemente per  $\varepsilon$  piccolo). Segue, scrivendo l'equazione per  $I$  in forma integrale,

$$I_\varepsilon(T_\varepsilon) - I^0 = \varepsilon \int_0^{T_\varepsilon} F(I_\varepsilon(t), \varphi_\varepsilon(t)) dt = \varepsilon \int_0^{T_\varepsilon} F(I^0, \varphi^0 + \omega t) dt + \mathcal{O}(\varepsilon^2 T_\varepsilon^2) , \tag{3.4}$$

ovvero anche

$$\frac{I_\varepsilon(T_\varepsilon) - I_\varepsilon(0)}{T_\varepsilon} = \varepsilon \frac{1}{T_\varepsilon} \int_0^{T_\varepsilon} F(I^0, \varphi^0 + \omega t) dt + \mathcal{O}(\varepsilon^2 T_\varepsilon) .$$

Ora, per  $\varepsilon$  piccolo la quantità a sinistra dell'uguale approssima la derivata  $\frac{dI}{dt}$  (apparentemente non va bene, perché  $T_\varepsilon$  tende all'infinito anziché a zero. Ma la scala di tempo naturale, per  $I$ , è  $t \sim \varepsilon^{-1} \gg T_\varepsilon$ , e dunque il tempo  $T_\varepsilon$ , per  $\varepsilon \rightarrow 0$ , in realtà è sempre più breve). A destra invece l'integrale diviso per  $T_\varepsilon$  approssima la media temporale

$$\bar{F}(I, \varphi) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T F(I, \varphi + \omega t) dt ,$$

mentre il termine  $\mathcal{O}(\varepsilon^2 T_\varepsilon)$  diventa trascurabile. Si “deduce” così per  $I$  un'equazione del tipo

$$\dot{I} = \varepsilon \bar{F}(I, \varphi) ,$$

che differisce dall'equazione mediata perché in essa appare la media temporale  $\overline{F}$  al posto della media in fase  $\mathcal{F}$ . Si vede così (pur nei limiti di questa trattazione del tutto euristica) che *condizione perché l'equazione mediata descriva bene il moto delle  $I$ , qualunque sia  $F$ , è che la traslazione sul toro:  $\varphi \mapsto \varphi + \omega t$  sia ergodica.*

Come sappiamo (capitolo primo, paragrafo 4) la condizione per l'ergodicità è che  $\omega$  sia non risonante: in tal caso, ma solo in tal caso, è ragionevole attendersi che l'equazione mediata sia significativa, mentre in presenza di risonanze le due medie (per funzioni generiche) non hanno alcuna relazione. Si osservi che nel caso di un solo angolo la traslazione sul toro è sempre ergodica, e corrispondentemente l'equazione mediata è sempre buona, a meno che non sia  $\omega = 0$ .

- **Esercizio.** Come deve essere fatta  $F$  perchè, pur nel caso non ergodico, le due medie coincidano?

Dietro la mancata corrispondenza tra il comportamento vero delle  $I$  e quello descritto dal sistema mediato sta dunque il fatto che in presenza di risonanze l' $n$ -pla degli angoli, nel corso del tempo, non invade il toro  $\mathbb{T}^n$ , ma soltanto un sotto-toro di dimensione inferiore. La questione è particolarmente chiara se si introducono nuovi angoli adattati al reticolo risonante, secondo quanto si è visto nel capitolo sulla teoria ergodica (paragrafo 2 e appendice A). Precisamente, se vi sono  $r$  risonanze indipendenti, e  $A$  è la matrice intera definita nella (2.8) del primo capitolo (lì denotata  $J$ ), introduciamo nuovi angoli  $\psi = (\psi_1, \dots, \psi_n)$  ponendo<sup>29)</sup>  $\psi = A\varphi$ , e denotiamo

$$\psi = (\psi_1, \dots, \psi_n) = (\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{n-r}) .$$

Ricordiamo che le prime  $r$  righe di  $A$  contengono vettori risonanti, cosicché ciascuno degli  $\alpha_i$  è definito da  $\alpha_i = k^{(i)} \cdot \varphi$  con  $k^{(i)}$  risonante, mentre le nuove frequenze sono

$$A\omega = (0, \dots, 0, \Omega_1, \dots, \Omega_{n-r}) .$$

Le equazioni del moto per le variabili  $I, \alpha, \beta$  sono allora del tipo

$$\dot{I} = \varepsilon F'(I, \alpha, \beta) , \quad \dot{\alpha} = \varepsilon G'(I, \alpha, \beta) , \quad \dot{\beta} = \Omega + \varepsilon G''(I, \alpha, \beta) ,$$

con opportune  $F', G', G''$ . Si vede bene che le variabili  $\alpha$  sono lente al pari delle  $I$ , e che le uniche variabili sulle quali può aver senso mediare sono i veri angoli veloci  $\beta$ . Tornando alle variabili originali  $\varphi$ , ugualmente si può dire che in presenza di risonanze è illusorio trattare tutti gli angoli come variabili veloci: alcune loro combinazioni lineari (le  $\alpha = k \cdot \varphi$ ) sono lente.

- **Esercizio.** Si consideri il sistema hamiltoniano

$$H = \omega(I_1 + I_2) - \varepsilon[\cos(\varphi_1 - \varphi_2) + \cos \varphi_2] ,$$

e se ne discuta il comportamento introducendo un angolo lento  $\alpha$  e un angolo veloce  $\beta$ . Si confrontino le soluzioni con quelle del sistema mediato e poi con quelle del sistema “parzialmente mediato”, con media solo su  $\beta$ . Si studi in modo analogo il sistema (3.3).

---

<sup>29)</sup> Nel caso hamiltoniano, la trasformazione si completa canonicamente introducendo nuove azioni  $\mathcal{I} = A^{-T}I$ .

Tutti questi esempi mostrano con chiarezza che le risonanze giocano un ruolo cruciale nella dinamica di un sistema del tipo (3.1). E si intuisce facilmente che la questione diviene particolarmente delicata, e insidiosa, nel caso strettamente anisocrono oggetto della proposizione 2.8.

A dispetto di questa difficoltà la teoria delle perturbazioni, in particolare per sistemi hamiltoniani, ha ottenuto (e continua ad ottenere) risultati di grande rilievo, nel tentativo ambizioso, e largamente incompiuto, di produrre risultati che siano contemporaneamente rigorosi matematicamente e interessanti fisicamente. Nel seguito esporremo solo alcuni dei metodi impiegati e dei risultati conseguiti: ci accontenteremo di assaggiare qua e là, senza pretesa di completezza, rinunciando anche ad alcune dimostrazioni. Una sintesi bella e completa fino al 1985 (ma i risultati continuano a susseguirsi) si trova in *Dynamical Systems III*, V.I. Arnol'd e A.I. Neishtadt editori (*Enciclopedia of Mathematical Sciences*, Springer 1988).

#### 4. Il “teorema di non esistenza” di Poincaré<sup>30)</sup>

##### 4.1. Una elementare ostruzione all'integrabilità.

Consideriamo un sistema hamiltoniano quasi integrabile,

$$H_\varepsilon(I, \varphi) = h(I) + \varepsilon f(I, \varphi) , \quad \begin{array}{l} I \in \mathcal{B} \subset \mathbb{R}^n, \\ \varphi \in \mathbb{T}^n . \end{array} \quad (4.1)$$

Una domanda spontanea è se, per  $\varepsilon$  piccolo, non esista una trasformazione canonica

$$(I, \varphi) = \mathcal{C}_\varepsilon(\tilde{I}, \tilde{\varphi}) \quad (4.2)$$

prossima all'identità:

$$\mathcal{C}_\varepsilon(\tilde{I}, \tilde{\varphi}) = (\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon) ,$$

che rimuove la dipendenza dagli angoli: ovvero tale che, denotata con  $\tilde{H}_\varepsilon := H_\varepsilon \circ \mathcal{C}_\varepsilon$  la nuova hamiltoniana, risulti

$$\begin{aligned} \tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) &= \tilde{h}_\varepsilon(\tilde{I}) \\ &= h(\tilde{I}) + \varepsilon g(\tilde{I}, \varepsilon) . \end{aligned} \quad (4.3)$$

Se così fosse l'hamiltoniana  $H_\varepsilon$ , a dispetto della perturbazione, sarebbe integrabile (e se ciò accadesse per una perturbazione in qualche senso generica, si potrebbe dire che

---

<sup>30)</sup> Virgolette indispensabili: l'esistenza di Poincaré è fuori discussione.



l'integrabilità è una proprietà robusta, o “strutturalmente stabile”).<sup>31)</sup> Come abbiamo già visto (proposizione 1.6) tutte le trasformazioni canoniche  $\varepsilon$ -prossime all'identità si possono scrivere tramite una funzione generatrice in variabili miste  $S(\tilde{I}, \varphi) = \tilde{I} \cdot \varphi + \varepsilon W(\tilde{I}, \varphi, \varepsilon)$ , e restano definite da

$$I = \tilde{I} + \varepsilon \frac{\partial W}{\partial \varphi}(\tilde{I}, \varphi, \varepsilon), \quad \tilde{\varphi} = \varphi + \varepsilon \frac{\partial W}{\partial I}(\tilde{I}, \varphi, \varepsilon); \quad (4.4)$$

per  $\varepsilon$  piccolo (e  $W$  regolare) le (4.4) sono invertibili, e dunque l'esistenza della (4.2) è assicurata. La domanda allora è se esista o meno una funzione  $W$ , tale che la nuova hamiltoniana  $\tilde{H}_\varepsilon$  abbia la forma (4.3).

Una pretesa di gran lunga più debole, ma che ugualmente prospetta una eventualità interessante, è che per effetto di un cambiamento di variabili prossimo all'identità del tipo (4.2) risulti

$$\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) = \tilde{h}(\tilde{I}) + \varepsilon \mathcal{R}(\tilde{I}, \tilde{\varphi}, \varepsilon), \quad (4.5)$$

con il resto  $\mathcal{R}$  non nullo ma piccolo con  $\varepsilon$ , ad esempio  $\mathcal{R} = \mathcal{O}(\varepsilon^r)$ ,  $r \geq 1$ . Se ciò avvenisse, allora evidentemente risulterebbe

$$\tilde{I}(t) - \tilde{I}(0) = \mathcal{O}(\varepsilon) \quad \text{per } |t| \leq \varepsilon^{-r},$$

e dal momento che (per ogni  $t$ ) si ha  $\tilde{I} - I \sim \varepsilon$ , seguirebbe anche

$$I(t) - I(0) = \mathcal{O}(\varepsilon) \quad \text{per } |t| \leq \varepsilon^{-r},$$

con sostanziale miglioramento rispetto alla stima di ordine zero

$$|I(t) - I(0)| \leq C \varepsilon |t|, \quad C > 0.$$

Si osservi che per  $r = 1$  questa è proprio la conclusione prospettata dal principio della media, nel caso hamiltoniano.

Si dimostra immediatamente che “in generale” non è così, precisamente

**Proposizione 4.1** (Poincaré). *Le tre seguenti proprietà non sono tra loro compatibili:*

i. *l'hamiltoniana imperturbata  $h$  è non degenere, precisamente in un aperto  $U \in \mathbb{R}^n$  risulta*

$$\det \left( \frac{\partial^2 h}{\partial I_i \partial I_j} \right) \neq 0;$$

ii. *la perturbazione  $f$  è generica, precisamente, scritta la sua serie di Fourier*

$$f(I, \varphi) = \sum_{k \in \mathbb{Z}^n} \hat{f}_k(I) e^{ik \cdot \varphi},$$

---

<sup>31)</sup> Questo problema era sentito come particolarmente cruciale a fine '800. Poincaré lo chiama “*Problème général de la dynamique.*” Poincaré fu il primo a comprendere che la linea di pensiero mirante alla soluzione per quadrature delle equazioni del moto era confinata a una classe di sistemi in qualche modo eccezionali, e che genericamente la dinamica è sostanzialmente più complicata di quella ammessa in un sistema integrabile.

per ogni  $k \in \mathbb{Z}^n$ ,  $k \neq 0$ , esiste  $k'$  parallelo a  $k$  tale che  $\hat{f}_{k'}(I) \neq 0$  in  $U$ ;

- iii. il sistema è "integrabile fino al primo ordine in  $\varepsilon$ " in  $U \times \mathbb{T}^n$ , precisamente esiste una funzione regolare  $\mathcal{W}(\tilde{I}, \varphi)$ , definita per  $(\tilde{I}, \varphi) \in U \times \mathbb{T}^n$ , tale che la trasformazione canonica  $(I, \varphi) = \mathcal{C}_\varepsilon(\tilde{I}, \tilde{\varphi})$  definita dalle (4.4) muta  $H_\varepsilon$  in  $\tilde{H}_\varepsilon = H_\varepsilon \circ \mathcal{C}_\varepsilon$  della forma

$$\tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) = \tilde{h}_\varepsilon(\tilde{I}) + \mathcal{O}(\varepsilon^2). \quad (4.6)$$

**Dimostrazione.** Dalla (4.4) si ottiene, per inversione della seconda equazione e sostituzione nella prima,

$$I = \tilde{I} + \varepsilon \frac{\partial W}{\partial \varphi}(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^2), \quad \varphi = \tilde{\varphi} + \mathcal{O}(\varepsilon),$$

da cui, per sostituzione in  $H_\varepsilon$ , si ottiene

$$\begin{aligned} \tilde{H}_\varepsilon(\tilde{I}, \tilde{\varphi}) &= h(\tilde{I} + \varepsilon \frac{\partial W}{\partial \varphi}(\tilde{I}, \tilde{\varphi}) + \mathcal{O}(\varepsilon^2)) + \varepsilon f(\tilde{I} + \mathcal{O}(\varepsilon), \tilde{\varphi} + \mathcal{O}(\varepsilon)) \\ &= h(\tilde{I}) + \varepsilon \left[ \omega(\tilde{I}) \cdot \frac{\partial W}{\partial \varphi}(\tilde{I}, \tilde{\varphi}) + f(\tilde{I}, \tilde{\varphi}) \right] + \mathcal{O}(\varepsilon^2) \end{aligned}$$

(si è usato  $\omega = \frac{\partial h}{\partial I}$ ). Lo scopo è raggiunto (la proprietà iii. è soddisfatta) se il contenuto della parentesi quadrata è indipendente dagli angoli, ovvero se esiste una funzione  $g : U \rightarrow \mathbb{R}$  tale che

$$\omega(\tilde{I}) \cdot \frac{\partial W}{\partial \varphi}(\tilde{I}, \tilde{\varphi}) + f(\tilde{I}, \tilde{\varphi}) = g(\tilde{I}) \quad (4.7)$$

in  $U \times \mathbb{T}^n$ . La (4.7) va vista come un'equazione per  $W$  e  $g$ , con  $\omega$  e  $f$  assegnate.

Ora si vede immediatamente che, nelle ipotesi i. e ii., la 4.5 non ammette soluzione: infatti, sviluppando anche  $W$  in serie di Fourier degli angoli,

$$W(\tilde{I}, \tilde{\varphi}) = \sum_{k \in \mathbb{Z}^n} \widehat{W}_k(\tilde{I}) e^{ik \cdot \tilde{\varphi}}, \quad (4.8)$$

si ottiene immediatamente  $\hat{f}_0(\tilde{I}) = g(\tilde{I})$ , e

$$i(k \cdot \omega(\tilde{I})) \widehat{W}_k(\tilde{I}) + \hat{f}_k(\tilde{I}) = 0 \quad \text{per ogni } k \neq 0.$$

L'equazione relativa a  $k = 0$  non crea problemi: semplicemente fissa  $g$ , ponendola uguale alla media di  $f$ . Le equazioni per  $k \neq 0$  dovrebbero determinare ciascuna uno dei coefficienti  $\widehat{W}_k$ , ma non sono tutte risolubili: per la proprietà i. (si ricordi la proposizione 2.8), in un insieme  $V$  denso in  $U$  vi sono relazioni di risonanza, ovvero per ogni  $\tilde{I} \in V$  esiste  $k \in \mathbb{Z}^2$  tale che  $k \cdot \omega(\tilde{I}) = 0$ ; evidentemente si ha anche

$$k' \cdot \omega(\tilde{I}) = 0 \quad \text{per ogni } k' \parallel k,$$

ma per la ii., uno almeno dei coefficienti  $\hat{f}_{k'}(\tilde{I})$  non è nullo, e questo basta. Detto più brutalmente: se si cerca di risolvere la (4.8) ponendo

$$\widehat{W}_k(\tilde{I}) = \frac{\hat{f}_k(\tilde{I})}{ik \cdot \omega(\tilde{I})},$$

in un insieme denso almeno qualcuno dei denominatori si annulla, senza che si annulli il corrispondente numeratore. ■

- **Osservazione.** La (4.7), che riscriviamo togliendo la fastidiosa tilde:

$$\omega(I) \cdot \frac{\partial W}{\partial \varphi}(I, \varphi) + f(I, \varphi) = g(I) , \quad (4.9)$$

deve considerarsi *l'equazione fondamentale della teoria hamiltoniana delle perturbazioni*.<sup>32)</sup> Nella sostanza, qualunque sia lo schema perturbativo che si sceglie, ci si trova di fronte a questa equazione, che dunque racchiude la difficoltà essenziale, l'ostruzione fondamentale per l'integrabilità. Come vedremo, i diversi metodi perturbativi si possono tutti riguardare come strategie possibili di fronte a questa equazione. Corrispondentemente, la proposizione 4.1 mostra, si può dire, che una strategia ingenua non esiste: almeno una delle tre affermazioni che appaiono nell'enunciato andrà indebolita.

#### 4.2. Non esistenza di integrali primi

La difficoltà sopra esposta entra in particolare in un sottile teorema dimostrato da Poincaré, riguardante la non esistenza “generica” di integrali del moto indipendenti dall'hamiltoniana. Un possibile enunciato è il seguente:

**Proposizione 4.2** (Poincaré). *Sia  $H_\varepsilon(I, \varphi) = h(I) + \varepsilon f(I, \varphi)$  analitica in  $U \times \mathbb{T}^n$ ,  $U \subset \mathbb{R}^n$  aperto, e siano soddisfatte le ipotesi i. e ii. della proposizione 4.1. Se la funzione  $F_\varepsilon(I, \varphi)$  è a sua volta analitica (anche in  $\varepsilon$ ), ed è un integrale del moto di  $H_\varepsilon$ , allora essa è funzione dell'hamiltoniana: precisamente esiste  $\mathcal{F}_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$ , tale che*

$$F_\varepsilon(I, \varphi) = \mathcal{F}_\varepsilon(H_\varepsilon(I, \varphi)) .$$

Omettiamo la dimostrazione. Il senso del teorema è quello di mostrare che, almeno in classe analitica, non solo l'integrabilità, ma anche l'esistenza di un singolo integrale del moto (con tutte le conseguenze che ne seguirebbero sul confinamento dei moti) è da considerarsi eccezionale. La maggior debolezza della proposizione è la richiesta che  $F_\varepsilon$  sia analitica (in particolare, analitica in  $\varepsilon$ ). Una ulteriore generalizzazione fu compiuta da Fermi:

**Proposizione 4.3** (Fermi). *Nelle ipotesi della proposizione 4.2, se in più si assume che il numero di gradi di libertà  $n$  sia maggiore di due, non esistono superfici  $n - 1$  dimensionali invarianti della forma*

$$F_\varepsilon(I, \varphi) = 0 ,$$

con  $F_\varepsilon$  analitica (anche in  $\varepsilon$ ), diverse dalle superfici di energia costante.

La generalizzazione consiste nel fatto che, mentre nella proposizione precedente si nega l'esistenza di un integrale del moto, e con esso di una foliazione dello spazio delle fasi in fogli invarianti del tipo

$$F_\varepsilon(I, \varphi) = c ,$$

---

<sup>32)</sup> Questa equazione altro non è che l'equazione di Hamilton–Jacobi, al primo ordine in  $\varepsilon$ .

per  $c$  in un qualche intervallo, qui si nega l'esistenza anche di un singolo foglio. La proposizione, di cui omettiamo la dimostrazione, interessava a Fermi in vista del problema ergodico: si comprende infatti che anche un solo un foglio invariante, indipendente dalle superfici di energia costante, sarebbe in grado di ritagliare su di esse insiemi invarianti, e dunque costituirebbe una ostruzione all'ergodicità. Fermi stesso credette ingenuamente<sup>33)</sup> di poter dimostrare, sulla base di questa proposizione, che genericamente i sistemi hamiltoniani (con almeno tre gradi di libertà) sono ergodici.<sup>34)</sup>

## 5. Il metodo di Lie. Piccoli divisori e condizione diofantea.

### 5.1. Il metodo di Lie

In questo paragrafo introdurremo il metodo che sarà usato nel seguito per costruire trasformazioni canoniche prossime all'identità, noto come *metodo di Lie*. La difficoltà espressa dal teorema di non esistenza sopra dimostrato è una difficoltà di sostanza, e dunque la ritroveremo. Ciononostante il metodo di Lie presenta diversi vantaggi rispetto al metodo delle funzioni generatrici in variabili miste, ed è di uso sempre più comune. Benché il nostro interesse sia rivolto esclusivamente al caso hamiltoniano, per meglio comprendere il metodo vale la pena di collocarsi inizialmente in un contesto più generale.

L'idea è semplicissima: se  $\mathcal{D}$  è un aperto di  $\mathbb{R}^l$  ( $l = 2n$  nel caso hamiltoniano), ogni equazione differenziale del tipo

$$\dot{x} = \varepsilon X(x), \quad x \in \mathcal{D},$$

definisce implicitamente tramite la sua mappa al tempo uno  $\mathcal{C}_\varepsilon = \Phi_{\varepsilon X}^1$  un cambiamento di variabili  $x = \mathcal{C}_\varepsilon(\tilde{x})$  piccolo con  $\varepsilon$ :

$$x = \tilde{x} + \mathcal{O}(\varepsilon), \quad \mathcal{C}_\varepsilon : \mathcal{D}_\varepsilon \rightarrow \mathcal{C}_\varepsilon(\mathcal{D}_\varepsilon) \subset \mathcal{D}$$

(il dominio  $\mathcal{D}_\varepsilon$  ove si può prendere il dato iniziale dovrà essere un po' ridotto rispetto a  $\mathcal{D}$ , per far sì che nel tempo uno la soluzione non esca da  $\mathcal{D}$ ; questo problema sarà affrontato più avanti in modo quantitativo). In modo del tutto equivalente si potrebbe prendere il campo vettoriale grande, cioè non proporzionale a  $\varepsilon$ , e considerare la mappa al tempo  $\varepsilon$  anziché al tempo uno: si ha infatti  $\Phi_{\varepsilon X}^t = \Phi_X^{\varepsilon t}$ .

Sia ora  $F$  una qualsiasi funzione  $\mathcal{D} \rightarrow \mathbb{R}$ . Per effetto del cambiamento di variabili la nuova funzione sarà  $F' = F \circ \mathcal{C}_\varepsilon$ , e poiché nel corso del tempo si ha

$$\frac{dF}{dt} = L_{\varepsilon X} F = \varepsilon L_X F, \quad \frac{d^k F}{dt^k} = \varepsilon^k L_X^k F,$$

<sup>33)</sup> L'elemento sottovalutato da Fermi (siamo nel 1921, decisamente agli albori della teoria ergodica) è il fatto che, tipicamente, gli eventuali insiemi invarianti di un sistema dinamico non hanno per frontiera una varietà (né tantomeno una varietà analitica), ma insiemi assai più complicati.

<sup>34)</sup> Più precisamente, "quasi ergodici" nel senso di Ehrenfest (un'orbita "generica" è densa).

risulterà a  $t$  qualsiasi (convergenza a parte)

$$F \circ \Phi_{\varepsilon X}^t = \sum_{k=0}^{\infty} \frac{(\varepsilon t)^k}{k!} L_X^k F = e^{\varepsilon t L_X} F ;$$

in particolare, per  $t = 1$ ,

$$F \circ \mathcal{C}_\varepsilon = e^{\varepsilon L_X} F .$$

Per evitare problemi di convergenza della serie (peraltro risolvibili senza troppa difficoltà, s'intende nel caso analitico) troncheremo lo sviluppo a ordini bassi, scrivendo il resto in forma opportuna: precisamente scriveremo

$$F \circ \mathcal{C}_\varepsilon = F + \varepsilon R_1[X, F] , \quad F \circ \mathcal{C}_\varepsilon = F + \varepsilon L_X F + \varepsilon^2 R_2[X, F] , \quad (5.1)$$

con

$$R_1[X, F] := \int_0^1 (L_X F) \circ \Phi_{\varepsilon X}^t dt , \quad R_2[X, F] := \int_0^1 dt \int_0^t (L_X^2 F) \circ \Phi_{\varepsilon X}^t dt' \quad (5.2)$$

(della stima dei resti ci occuperemo più avanti).

Se  $X$  è un campo vettoriale hamiltoniano, corrispondente a una qualche funzione di Hamilton  $\chi$ , allora (proposizione 1.7) la trasformazione  $\mathcal{C}$  è canonica. Con notazione adattata al caso hamiltoniano le (5.1), (5.2) si riscrivono

$$\begin{aligned} F \circ \mathcal{C}_\varepsilon &= F + \varepsilon R_1[\chi, F] , & F \circ \mathcal{C}_\varepsilon &= F + \varepsilon \{F, \chi\} + \varepsilon^2 R_2[\chi, F] \\ R_1[\chi, F] &:= \int_0^1 \{F, \chi\} \circ \Phi_{\varepsilon \chi}^t dt & R_2[\chi, F] &:= \int_0^1 dt \int_0^t \{\{F, \chi\}, \chi\} \circ \Phi_{\varepsilon \chi}^t dt' . \end{aligned} \quad (5.3)$$

Consideriamo ora una hamiltoniana quasi integrabile

$$H(I, \varphi) = h(I) + \varepsilon f(I, \varphi) ,$$

e cerchiamo  $\chi$  in modo tale che la nuova hamiltoniana  $H' = H \circ \mathcal{C}$  sia integrabile a meno di termini di ordine  $\varepsilon^2$ . Applicando la seconda delle (5.3) a  $h$  e la prima a  $\varepsilon f$ , si trova subito

$$H' = h + \varepsilon (f + \{h, \chi\}) + \varepsilon^2 (R_2[\chi, h] + R_1[\chi, f]) .$$

Riusciamo nel nostro proposito se troviamo  $\chi$  tale che la quantità  $f + \{h, \chi\}$  sia una funzione delle sole azioni: e dunque, tenendo conto della forma di  $h$ , se risulta

$$-\omega(I) \cdot \frac{\partial \chi}{\partial \varphi}(I, \varphi) + f(I, \varphi) = g(I)$$

(abbiamo mantenuto i vecchi nomi  $(I, \varphi)$  alle variabili). Questa è la stessa equazione incontrata nel paragrafo precedente (con  $\chi = -W$ ). Come si è già visto, è obbligatorio

porre  $g = \langle f \rangle_\varphi$  (cosa peraltro evidente se si osserva che  $\omega \cdot \frac{\partial \chi}{\partial \varphi}$  ha necessariamente media nulla sugli angoli), e dunque l'equazione prende la forma

$$\omega(I) \cdot \frac{\partial \chi}{\partial \varphi}(I, \varphi) = f(I, \varphi) - \langle f \rangle_\varphi . \quad (5.4)$$

La nuova hamiltoniana corrispondentemente avrà la forma

$$H'(I, \varphi) = h(I) + \varepsilon \langle f \rangle_\varphi + \varepsilon^2 f'(I, \varphi) , \quad f' = R_2[\chi, h] + R_1[\chi, f] . \quad (5.5)$$

## 5.2. Studio preliminare della (5.4) nel caso isocrono

Il caso più semplice in cui la (5.4) è risolubile è il cosiddetto caso di Birkhoff, o caso isocrono finito nonrisonante, in cui:

- i)  $\omega$  è costante ;
- ii) la perturbazione  $f$  ha un numero finito di componenti di Fourier, diciamo

$$f(I, \varphi) = \sum_{|k| \leq K} \hat{f}(I) e^{ik \cdot \varphi} ;$$

- iii)  $\omega$  non risuona con i vettori  $k$  presenti nello sviluppo di  $f$ .

L'interesse di Birkhoff è dovuto al fatto che queste ipotesi sono naturali per il problema delle piccole oscillazioni attorno a una configurazione di equilibrio.<sup>35)</sup>

Entrambe le ipotesi violano le ipotesi del teorema di Poincaré, dunque non c'è da meravigliarsi che in questo caso la (5.5) si risolva. La soluzione si trova subito in forma di serie di Fourier, precisamente

$$\chi(I, \varphi) = \sum_{|k| \leq K} \hat{\chi}(I) e^{ik \cdot \varphi} , \quad \hat{\chi}(I) = \frac{\hat{f}(I)}{ik \cdot \omega} .$$

I termini sono in numero finito, e per le ipotesi fatte i denominatori sono sempre diversi da zero.

Un secondo caso interessante è ancora quello isocrono, senza però l'ipotesi ii) su  $f$ . L'ipotesi iii) sembrerebbe divenire naturalmente l'ipotesi che  $\omega$  sia non risonante: ma un minimo di riflessione mostra che ciò non basta, perché tale condizione è sufficiente solo a garantire l'esistenza dei singoli coefficienti di Fourier  $\hat{\chi}_k$  e non la convergenza della serie

$$\chi(I, \varphi) = \sum_{k \in \mathbb{Z}^n} \hat{\chi}(I) e^{ik \cdot \varphi} , \quad \hat{\chi}(I) = \frac{\hat{f}(I)}{ik \cdot \omega} \quad (5.6)$$

che formalmente risolve la (5.4). Infatti per  $k$  (grandi) opportuni i denominatori, pur non nulli, possono essere arbitrariamente piccoli, precludendo la convergenza della serie e a

---

<sup>35)</sup> Più precisamente, Birkhoff si confrontava a ogni ordine perturbativo con un problema sostanzialmente equivalente a questo, scritto però in variabili cartesiane anziché di azione–angolo.

maggior ragione la convergenza a una funzione regolare. Sarà necessario, per la convergenza:

- i) chiedere che i coefficienti  $\hat{f}_k$  decrescano abbastanza rapidamente per  $|k| \rightarrow \infty$ ;
- ii) chiedere che il denominatore  $|k \cdot \omega|$  non cresca troppo in fretta al crescere di  $|k|$ , diciamo che risulti

$$|k \cdot \omega| \geq \xi(|k|) , \quad (5.7)$$

con  $\xi \rightarrow 0$  per  $|k| \rightarrow \infty$  “in modo opportuno”.

Qui e nel seguito  $|k|$  denota la norma, particolarmente comoda per vettori  $k \in \mathbb{Z}^n$ ,

$$|k| = |k_1| + \dots + |k_n| .$$

Per quanto riguarda la prima condizione, è ben noto che più una funzione  $F : \mathbb{T}^n \rightarrow \mathbb{R}$  è regolare, più i coefficienti di Fourier

$$\hat{F}_k = (2\pi)^{-n} \int_0^{2\pi} F(\varphi) e^{-ik \cdot \varphi} d\varphi , \quad k \in \mathbb{Z}^n , \quad (5.8)$$

decadono rapidamente al crescere di  $|k|$ . Se  $F$  è di classe  $C^s$ , allora  $\hat{F}_k$  decresce come  $|k|^{-(s+1)}$ ; se è di classe  $C^\infty$  allora  $\hat{F}_k$  decresce più rapidamente di ogni potenza inversa di  $|k|$ ; infine se  $F$  è analitica  $\hat{F}_k$  decresce esponenzialmente con  $|k|$ . A noi interesserà il caso analitico, precisamente il caso di funzioni  $F$  analitiche e limitate in una striscia complessa della forma

$$\mathcal{S}_\sigma := \{ \varphi \in \mathbb{C}^n : |\operatorname{Im} \varphi_j| < \sigma, j = 1, \dots, n \} , \quad \sigma > 0 .$$

Denotiamo

$$\|F\|_\sigma = \sup_{\varphi \in \mathcal{S}_\sigma} |F(\varphi)| .$$

Un risultato classico è che

**Lemma 5.1.** *Se  $F$  è una funzione analitica:  $\mathcal{S}_\sigma \rightarrow \mathbb{C}$ , periodica di periodo reale  $2\pi$  in ciascuna variabile, allora si ha*

$$|\hat{F}_k| \leq \|F\|_\sigma e^{-|k|\sigma} . \quad (5.9)$$

*Viceversa se i coefficienti  $\hat{F}_k$ ,  $k \in \mathbb{Z}^n$ , decadono esponenzialmente con  $|k|$ ,*

$$|\hat{F}_k| \leq \mathcal{F} e^{-|k|\sigma} ,$$

*allora la serie di Fourier*

$$F(\varphi) = \sum_{k \in \mathbb{Z}^n} \hat{F}_k e^{ik \cdot \varphi} \quad (5.10)$$

*converge a una funzione analitica  $F : \mathcal{S}_\sigma \rightarrow \mathbb{C}$ , e per ogni  $\delta$  positivo,  $\delta < \min(\sigma, 1)$ , vale la stima*

$$\|F\|_{\sigma-\delta} < 2^{2n} \delta^{-n} \mathcal{F} . \quad (5.11)$$

- **Osservazione.** La condizione  $\delta < 1$  non ha in realtà alcuna rilevanza, serve solo a semplificare un po' l'espressione del risultato.

**Dimostrazione.** La prima parte segue dall'espressione (5.8) di  $F_k$ : l'integrazione in ciascuna variabile  $\varphi_i$  da 0 a  $2\pi$  si fa sul cammino  $\text{Im } \varphi = \pm\sigma$ , con segno opposto al segno di  $k_i$  (i tratti "verticali" in 0 e  $2\pi$  si compensano), e la (5.9) è immediata. Per la seconda parte, mostriamo che la serie (5.10) converge uniformemente in ogni striscia  $\mathcal{S}_{\sigma-\delta}$ ,  $0 < \delta < \sigma$ , e che per  $\delta < 1$  vale precisamente la (5.11). Per questo osserviamo che per  $\varphi \in \mathcal{S}_{\sigma-\delta}$  si ha

$$|F(\varphi)| \leq \sum_{k \in \mathbb{Z}^n} |F_k| e^{|k|(\sigma-\delta)} \leq \mathcal{F} \sum_{k \in \mathbb{Z}^n} e^{-|k|\delta} = \mathcal{F} \left( \sum_{s \in \mathbb{Z}} e^{-|s|\delta} \right)^n,$$

e d'altra parte

$$\sum_{s \in \mathbb{Z}} e^{-|s|\delta} < 2 \sum_{s=0}^{\infty} e^{-s\delta} = \frac{2}{1 - e^{-\delta}}.$$

Per  $\delta < 1$  il denominatore è maggiore di  $\frac{1}{2}\delta$ , e la conclusione è immediata. ■

Veniamo ora allo studio dei piccoli denominatori, ovvero alla condizione (5.7). La scelta di  $\xi$  è delicata: se va a zero troppo rapidamente allora è in pericolo la convergenza o la regolarità della serie. Se va a zero troppo lentamente le  $\omega$  che la soddisfano sono poche (o non ne esistono affatto).

Una scelta buona<sup>36)</sup> risulta essere la seguente:

$$|k \cdot \omega| \geq \gamma |k|^{-\tau} \quad \forall k \in \mathbb{Z}^n \setminus \{0\}, \quad \gamma > 0, \quad \tau > n - 1. \quad (5.12)$$

Questa condizione (più in generale, una condizione aritmetica della forma (5.7)) è nota in letteratura come *condizione diofantea*.<sup>37)</sup> Si sa dalla teoria dei numeri che l'insieme delle  $\omega$  che passano la condizione (5.12), diciamo l'insieme

$$\Omega_{\gamma,\tau} = \{\omega \in \Omega : |k \cdot \omega| \geq \gamma |k|^{-\tau} \forall k \in \mathbb{Z}^n \setminus \{0\}\}, \quad \Omega := \{\omega \in \mathbb{R}^n : \|\omega\| < L\},$$

- è vuoto se  $\tau < n - 1$ ;
- è non numerabile ma di misura zero per  $\tau = n - 1$ ;
- ha misura  $\gamma$ -prossima a  $\text{mes } \Omega$  per  $\tau > n - 1$ .

Nell'appendice D dimostriamo l'ultima proprietà (che è quella che più ci interessa), e precisamente che

**Proposizione 5.2.** *Se  $\tau > n - 1$  allora risulta*

$$\text{mes}(\Omega \setminus \Omega_{\gamma,\tau}) < C(\tau, n) \gamma L^{-1} \text{Vol}(\Omega), \quad C(\tau, n) = n^{\frac{3}{2}} \sum_{s=1}^{\infty} s^{n-\tau-2} < \infty. \quad (5.13)$$

<sup>36)</sup> Siegel (un teorico dei numeri interessato anche a sistemi hamiltoniani e Meccanica Celeste), 1942.

<sup>37)</sup> Si usa anche lo sgradevole inglesismo "diofantino" al posto di diofanteo. Diofanto (III secolo d. C.) è autore di un'opera di aritmetica di cui una parte (una collezione di problemi a soluzione intera o razionale) si è conservata. L'opera ha molto stimolato, al tempo di Fermat, il sorgere della moderna teoria dei numeri. Fermat ne ebbe una copia, sul margine della quale annotava i teoremi che via via scopri, compreso il famoso "ultimo teorema".



In particolare per  $\tau = n$  si ha  $C < 2n^{\frac{3}{2}}$ .

- **Osservazione.** L'insieme  $\Omega_{\gamma,\tau}$ , anche per  $\tau > n - 1$  e  $\gamma$  piccolo, ha interno vuoto: con evidenza infatti  $\Omega \setminus \Omega_{\gamma,\tau}$  contiene l'insieme dei punti risonanti, che è denso in  $\Omega$ . Anzi, dalla definizione si vede che  $\Omega \setminus \Omega_{\gamma,\tau}$  è un aperto denso, benché di misura prossima a zero per  $\gamma$  piccolo. L'uso della misura, e precisamente della misura di Lebesgue, qui è essenziale.
- **Osservazione.** Per  $n = 2$  stiamo considerando l'insieme delle  $\omega$  tali che il rapporto  $\omega_2/\omega_1$  è male approssimato da ogni razionale  $p/q$ , precisamente

$$\left| \frac{\omega_2}{\omega_1} - \frac{p}{q} \right| \geq \frac{1}{|q| |\omega_1|} \frac{\gamma}{(|p| + |q|)^\tau}$$

(basta porre  $p = k_1$ ,  $q = -k_2$ ). Ci si riporta così al problema delle approssimazioni razionali dei reali. Il problema è ben studiato, e lo strumento fondamentale è la *frazione continua*. Molto poco si sa invece del caso  $n \geq 3$ .

Mettendo insieme la condizione diofantea e le proprietà di decadimento dei coefficienti di Fourier di funzioni analitiche si deduce facilmente che la (5.4) si risolve, e che la soluzione  $\chi$  è a sua volta analitica in  $\mathcal{D}_\sigma$ . A livello qualitativo la cosa è evidente: a numeratore nella (5.6) i coefficienti decadono esponenzialmente con  $|k|$ , mentre i denominatori crescono solo a potenza. Stime quantitative sulla (5.5) richiedono scelte precise di domini e norme, e preferiamo rinviarle al prossimo paragrafo 6.2 quando dovremo risolvere l'equazione all'interno di un problema interessante.

### 5.3. Il metodo di Lie: alcune stime

Introduciamo qui, per alleggerire il prossimo paragrafo, alcune stime di base che intervengono nel metodo di Lie introdotto sopra, adattate al caso analitico che più ci interessa. Anche qui, per ragioni di chiarezza, conviene tenersi in un ambito più generale di quello hamiltoniano.

Sia dunque  $\mathcal{D}$  un dominio reale,  $\mathcal{D} \subset \mathbb{R}^l$ ; in corrispondenza all' $n$ -pla positiva  $\rho = (\rho_1, \dots, \rho_l)$  introduciamo il dominio esteso

$$\mathcal{D}_\rho = \bigcup_{x \in \mathcal{D}} \Delta_\rho(x) \tag{5.14}$$

$$\Delta_\rho(x) := \{x' \in \mathbb{C}^l : |x'_j - x_j| \leq \rho_j, j = 1, \dots, l\}.$$

Interessano funzioni analitiche  $F : \mathcal{D}_\sigma \rightarrow \mathbb{C}$ , con  $\sigma = (\sigma_1, \dots, \sigma_l) \leq \rho$  (la disuguaglianza si intende valere separatamente su ciascun elemento). Tutte le funzioni di questo tipo saranno implicitamente supposte reali per variabile reale. Se  $F : \mathcal{D}_\sigma \rightarrow \mathbb{C}$  denoteremo

$$\|F\|_\sigma = \sup_{x \in \mathcal{D}_\sigma} |F(x)|. \tag{5.15}$$

Avremo anche a che fare con funzioni analitiche a valori in  $\mathbb{C}^n$ , come i campi vettoriali  $X : \mathcal{D}_\sigma \rightarrow \mathbb{C}^n$  o i loro flussi. Per esse sarà naturale fare riferimento a una norma che include anche il massimo sulle componenti, e precisamente

$$\|X\|_\sigma = \max_{1 \leq j \leq l} \frac{\|X_j\|_\sigma}{\rho_j}. \tag{5.16}$$

La divisione per  $\rho_j$  (attenzione: non per  $\sigma_j$ ) si rivela utile, e tra l'altro rende omogenee dimensionalmente le quantità tra le quali si prende il massimo, con grandi benefici.<sup>38)</sup>

- **Osservazione.** Le norme qui introdotte sono norme naturali per funzioni analitiche, dal momento che, grazie alla *disuguaglianza di Cauchy*, la norma di una funzione domina anche le sue derivate di ogni ordine. Nel caso semplice di una funzione di una sola variabile complessa  $x \in \mathcal{D}_\sigma$ , ricordiamo, la disuguaglianza di Cauchy<sup>39)</sup> si scrive

$$|F^{(s)}(x)| \leq \frac{s!}{\delta^s} \sup_{z \in \Delta_\delta(x)} |F(z)|, \quad x \in \mathcal{D}_{\sigma-\delta},$$

e dunque in particolare

$$\|F^{(s)}\|_{\sigma-\delta} \leq \frac{s!}{\delta^s} \|F\|_\sigma; \quad \text{per } x \in \mathcal{D} \quad |F^{(s)}(x)| \leq \frac{s!}{\sigma^s} \|F\|_\sigma.$$

Disuguaglianze della forma

$$|F^{(s)}(x)| \leq (\text{cost}) \frac{s!}{\delta^s}$$

sono caratteristiche delle funzioni analitiche (contengono l'informazione che  $F$  è analitica in un disco di raggio  $\delta$  attorno a  $x$ ).

Consideriamo dunque l'equazione differenziale

$$\dot{x} = X(x), \quad x \in \mathcal{D}_\sigma,$$

e sia  $\Phi_X^t(x)$  la soluzione al tempo (complesso)  $t$  con dato iniziale  $x \in \mathcal{D}_\sigma$ . Vale il seguente elementare

**Lemma 5.3.** *Se il campo vettoriale  $X$  è analitico in  $\mathcal{D}_\sigma$ , e soddisfa la disuguaglianza*

$$\|X\|_\sigma \leq \theta, \tag{5.17}$$

*allora  $\Phi_X^t$  per  $|t| \leq 1$  è una mappa analitica da  $\mathcal{D}_{\sigma-\theta\rho}$  a  $\mathcal{D}_\sigma$ , e si ha*

$$\|\Phi_X^t - \text{Id}\|_{\sigma-\theta\rho} \leq |t| \|X\|_\sigma. \tag{5.18}$$

**Dimostrazione.** Portando in forma integrale l'equazione differenziale, e usando la notazione abbreviata  $x_i(t) := (\Phi_X^t(x))_i$ , si ottiene

$$x_i(t) - x_i(0) = \int_0^t X_i(\Phi_X^{t'}(x)) dt', \quad i = 1, \dots, n;$$

<sup>38)</sup> Nel caso hamiltoniano, ma in pratica in ogni problema fisico, le diverse variabili hanno dimensione fisica diversa. Rispettare le dimensioni fisiche non è un vezzo ma un fatto essenziale anche dal punto di vista matematico, perché rende la teoria invariante per cambiamento delle unità di misura, e fa comparire le sole costanti essenziali di ogni problema che sono quelle adimensionali.

<sup>39)</sup> Conseguenza immediata della formula integrale di Cauchy  $F^{(s)}(x) = \frac{s!}{2\pi i} \oint \frac{F(z) dz}{(z-x)^{s+1}}$ .

come cammino di integrazione nel piano complesso  $t$  si può prendere il segmento da zero a  $t$ . Si vede subito allora che se  $|t|$  è abbastanza piccolo, allora  $\Phi_X^{t'}(x)$ , per  $x \in \mathcal{D}_{\sigma-\theta\rho}$ , non esce da  $\mathcal{D}_\sigma$ , e si ha la stima

$$|x_i(t) - x_i(0)| \leq |t| \|X_i\|_\sigma \leq |t| \|X\|_\sigma \rho_i, \quad i = 1, \dots, n, \quad (5.19)$$

ovvero la (5.18). Dalla (5.19) assieme alla (5.17) si vede poi immediatamente che  $\Phi_X^t(x)$  non può uscire<sup>40)</sup> da  $\mathcal{D}_\sigma$ , ove vale la (5.17) stessa, per  $|t| < 1$ . L'analiticità di  $\Phi_X^t(x)$  come funzione del dato iniziale  $x$  è un fatto generale per le equazioni differenziali analitiche. ■

Infine, in vista della stima dei resti (5.1), serve una stima della derivata di Lie  $L_X$ .

**Lemma 5.4.** *Se  $F$  è una funzione analitica:  $\mathcal{D}_\sigma \rightarrow \mathbb{C}$ , e il campo vettoriale  $X$  è analitico in  $\mathcal{D}_{\sigma-\theta\rho}$ , allora si ha*

$$\|L_X F\|_{\sigma-\theta\rho} \leq \frac{1}{\theta} \|X\|_{\sigma-\theta\rho} \|F\|_\sigma. \quad (5.20)$$

**Dimostrazione.** La stima (5.20) si ottiene dalla disuguaglianza di Cauchy, scrivendo la derivata di Lie  $L_X F$  come derivata direzionale: posto infatti  $\mathcal{F}_x(t) = F(x + tX(x))$ , si ha evidentemente

$$(L_X F)(x) = \mathcal{F}'_x(0)$$

(l'apice denota qui la derivata rispetto a  $t$ ). Ma per  $x \in \mathcal{D}_{\sigma-\theta\rho}$ , e

$$|t| \leq \tau := \theta \|X\|_{\sigma-\theta\rho}^{-1},$$

si ha  $x + tX(x) \in \mathcal{D}_\sigma$ , ove  $F$  è stimata. Segue allora

$$|\mathcal{F}'_x(0)| \leq \frac{1}{\tau} \sup_{|t| \leq \tau} |\mathcal{F}(t)| \leq \frac{1}{\tau} \|F\|_\sigma,$$

e la (5.20) è immediata.<sup>41)</sup> ■

<sup>40)</sup> Una dimostrazione formale di questo punto, che può apparire una petizione di principio, si fa per assurdo: se si ammette che  $\Phi_X^t(x)$  sia sul bordo di  $\mathcal{D}_\sigma$  per  $|t| = \tau < 1$ , e che invece risulti  $\Phi_X^t(x) \in \mathcal{D}_\sigma$  “prima”, cioè per  $|t| < \tau$ , usando la (5.17) si cade immediatamente in una contraddizione.

<sup>41)</sup> Un uso più ingenuo della stima di Cauchy per ciascuna delle  $l$  derivate parziali che appaiono in  $L_X F$  condurrebbe a una stima analoga, con in più uno spurio fattore  $l$ .

## 6. Sistemi isocroni: stime esponenziali

### 6.1. Formulazione del risultato

Consideriamo un sistema di oscillatori armonici debolmente accoppiati scritto in variabili di azione–angolo,

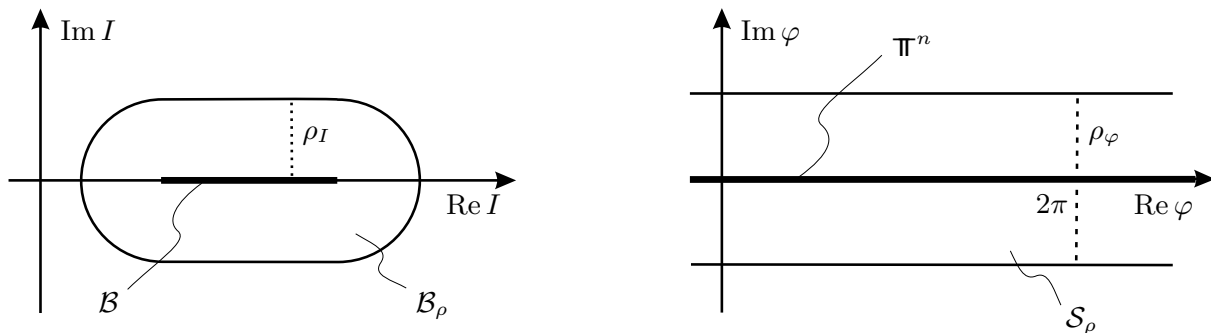
$$H(I, \varphi; \varepsilon) = \omega \cdot I + \varepsilon f(I, \varphi; \varepsilon), \quad I = (I_1, \dots, I_n) \in \mathcal{B} \subset \mathbb{R}^n$$

$$\varphi = (\varphi_1, \dots, \varphi_n) \in \mathbb{T}^n.$$

Per estendere nel complesso il dominio reale  $\mathcal{D} = \mathcal{B} \times \mathbb{T}^n$  introduciamo per semplicità la sola coppia<sup>42)</sup>  $\rho = (\rho_I, \rho_\varphi)$ , con  $\rho_I, \rho_\varphi > 0$ , e con riferimento alla definizione (5.14) di  $\Delta_\rho$  poniamo

$$\mathcal{B}_\rho = \bigcup_{I \in \mathcal{B}} \Delta_\rho(I), \quad \mathcal{S}_\rho = \{\varphi \in \mathbb{C}^n : |\operatorname{Im} \varphi_i| < \rho_\varphi, i = 1, \dots, n\}, \quad \mathcal{D}_\rho = \mathcal{B}_\rho \times \mathcal{S}_\rho.$$

La figura 6.1 illustra i domini  $\mathcal{B}_\rho$  e  $\mathcal{S}_\rho$  per  $n = 1$ .



**Figura 6.1.** I domini  $\mathcal{B}_\rho$  e  $\mathcal{S}_\rho$  per  $n = 1$ .

L’hamiltoniana sarà supposta analitica (e limitata) in  $\mathcal{D}_\rho$ ; più in generale, avremo a che fare con funzioni analitiche e limitate in insiemi del tipo  $\mathcal{D}_\sigma \subset \mathcal{D}_\rho$ ,  $\sigma = (\sigma_I, \sigma_\varphi) \leq \rho$  (come sopra, la disuguaglianza si intende valere separatamente sulle diverse componenti). Tutte le funzioni saranno implicitamente supposte reali per variabile reale. In accordo con la (5.15) porremo<sup>43)</sup>

$$\|F\|_\sigma = \sup_{(I, \varphi) \in \mathcal{D}_\sigma} |F(I, \varphi)|;$$

<sup>42)</sup> Non vi sarebbe difficoltà alcuna a introdurre diverse estensioni complesse per i diversi gradi di libertà, ponendo  $\rho = (\rho_{I_1}, \dots, \rho_{I_n}, \rho_{\varphi_1}, \dots, \rho_{\varphi_n})$ .

<sup>43)</sup> L’uso della sup–norma non è ottimale (norme che facciano esplicito riferimento alle componenti di Fourier delle funzioni danno risultati migliori), ma dà luogo agli sviluppi più semplici.

se  $F$  dipende parametricamente anche da  $\varepsilon$ , sarà sottinteso il sup anche in  $\varepsilon$  in un fissato intervallo  $(-\bar{\varepsilon}, \bar{\varepsilon})$ . Per quanto riguarda i campi vettoriali

$$X = (X_I, X_\varphi) = (X_{I_1}, \dots, X_{I_n}, X_{\varphi_1}, \dots, X_{\varphi_n}) ,$$

o più in generale per le funzioni a valori in  $\mathbb{C}^{2n}$  come le trasformazioni canoniche, in analogia alla (5.16) porremo

$$\|X\|_\sigma = \max_{i=1, \dots, n} \max \left( \frac{\|X_{I_i}\|_\sigma}{\rho_I}, \frac{\|X_{\varphi_i}\|_\sigma}{\rho_\varphi} \right) . \quad (6.1)$$

La proposizione seguente è un adattamento del teorema di Nekhoroshev al caso (più semplice, ma non compreso nel teorema) degli oscillatori armonici con frequenze  $\omega = (\omega_1, \dots, \omega_n)$  diofantee, precisamente soddisfacenti la condizione (5.12); per semplicità in questa condizione fisseremo  $\tau = n$ . La prima enunciazione della proposizione è dovuta a G. Gallavotti (1984).

**Proposizione 6.1.** *Sia data l'hamiltoniana*

$$H(I, \varphi; \varepsilon) = \omega \cdot I + \varepsilon f(I, \varphi; \varepsilon) , \quad I = (I_1, \dots, I_n) \in \mathcal{B} \subset \mathbb{R}^n \\ \varphi = (\varphi_1, \dots, \varphi_n) \in \mathbb{T}^n .$$

Posto  $\mathcal{D} = \mathcal{B} \times \mathbb{T}^n$ ,  $\rho = (\rho_I, \rho_\varphi)$ , e con riferimento alle estensioni complesse e alle altre notazioni sopra introdotte, facciamo le seguenti ipotesi:

- a)  $f$  è analitica e limitata in  $\mathcal{D}_\rho$ ;
- b) l' $n$ -pla delle frequenze soddisfa la relazione diofantea

$$|k \cdot \omega| \geq \gamma |k|^{-n} \quad \forall k \in \mathbb{Z}^n \setminus \{0\} \quad (6.2)$$

per un opportuno  $\gamma > 0$ .

Allora, posto<sup>44)</sup>

$$E = \gamma \rho_\varphi^{2n+1} \rho_I , \quad \varepsilon_0 = C^{-1} E \|f\|_\rho^{-1} ,$$

ove  $C$  è una costante opportuna dipendente solo da  $n$  (un valore possibile per  $C$  è  $2^{10n+6} n^n e^{-n+1}$ ),

- i) esiste una trasformazione canonica  $(I, \varphi) = \mathcal{C}(I', \varphi')$ ,  $\mathcal{C} : \mathcal{D}_{\frac{1}{2}\rho} \rightarrow \mathcal{C}(\mathcal{D}_{\frac{1}{2}\rho}) \subset \mathcal{D}_\rho$ , che dà alla nuova hamiltoniana  $H' = H \circ \mathcal{C}$  la forma

$$H'(I', \varphi'; \varepsilon) = \omega \cdot I' + \varepsilon g(I'; \varepsilon) + \varepsilon e^{-[(\varepsilon_0/\varepsilon)^{1/m}]} f'(I', \varphi'; \varepsilon) , \quad m = 2n + 2 , \quad (6.3)$$

ove  $[\cdot]$  denota la parte intera, con  $g, f'$  limitate da

$$\|g\|_{\frac{1}{2}\rho} \leq 2 \|f\|_\rho , \quad \|f'\|_{\frac{1}{2}\rho} \leq \|f\|_\rho ; \quad (6.4)$$

<sup>44)</sup> La costante  $E$  ha le dimensioni di un'energia, e viene ad essere un naturale termine di confronto per la perturbazione  $f$ .

ii) la trasformazione  $\mathcal{C}$  è piccola con  $\varepsilon$ , precisamente si ha<sup>45)</sup>

$$\|\mathcal{C} - \text{Id}\|_{\frac{1}{4}\rho} < \frac{\varepsilon}{\varepsilon_0} . \quad (6.5)$$

- **Osservazione.** Non appare esplicitamente tra le ipotesi un requisito di piccolezza per il parametro perturbativo  $\varepsilon$ , del tipo  $\varepsilon < \varepsilon_0$ , tipico di tanti teoremi in teoria delle perturbazioni. Formalmente, per  $\varepsilon \geq \varepsilon_0$ , la proposizione resta vera — basta prendere  $\mathcal{C} = \text{Id}$ ,  $H' = H$  — ma ovviamente non dice più nulla. Diciamo che  $\varepsilon$  deve essere piccolo perché il risultato sia interessante.
- **Osservazione.** Le costanti che compaiono in questo enunciato non sono ottimali. In particolare per  $m$ , che tra esse è la più rilevante, con un po' di lavoro in più e una scelta delle norme adattata al problema si otterrebbe il valore  $n + 1$  al posto di  $2n + 2$ .

La strategia della dimostrazione sarà quella di costruire la trasformazione canonica  $\mathcal{C}$  come composizione di  $r$  trasformazioni canoniche elementari  $\mathcal{W}_0, \dots, \mathcal{W}_{r-1}$ ,

$$\mathcal{C} = \mathcal{W}_{r-1} \circ \dots \circ \mathcal{W}_0 ,$$

con  $r$  da determinarsi in modo opportuno; dopo  $s$  passi,  $s \leq r$ , l'hamiltoniana dovrà avere la forma “integrabile fino all'ordine  $s$ ”

$$H_s(I, \varphi; \varepsilon) = \omega \cdot I + \varepsilon g_s(I; \varepsilon) + \varepsilon^{s+1} f_s(I, \varphi; \varepsilon) .$$

La difficoltà della costruzione sta nella necessità di eliminare progressivamente gli angoli dall'hamiltoniana, e contemporaneamente<sup>46)</sup> “tenere sotto controllo” le funzioni  $g_s$  e  $f_s$ . Ciò non riesce nel dominio iniziale  $\mathcal{D}_\rho$ , ma in domini  $\mathcal{D}_{\rho_s} \subset \mathcal{D}_\rho$  con l'estensione complessa progressivamente ridotta, precisamente

$$\rho_s = \rho(1 - s\eta) < \rho , \quad \eta = \frac{1}{2r} ;$$

la scelta di  $\eta$  è dovuta alla necessità di finire, dopo  $r$  passi, in  $\mathcal{D}_{\frac{1}{2}\rho}$ .

Procederemo allora così: nel prossimo paragrafo studiamo in dettaglio il singolo passo, ovvero la trasformazione canonica elementare

$$\mathcal{W}_s : \mathcal{D}_{\rho_{s+1}} \rightarrow \mathcal{D}_{\rho_s} , \quad s = 0, \dots, r - 1 ;$$

nel paragrafo successivo mettiamo insieme i diversi passi e concludiamo la dimostrazione.

<sup>45)</sup> Ovvero, posto  $\mathcal{C}(I', \varphi') = (I', \varphi') + \alpha(I', \varphi')$ ,  $\alpha = (\alpha^{I_1}, \dots, \alpha^{I_n}, \alpha^{\varphi_1}, \dots, \alpha^{\varphi_n})$ , si ha

$$\|\alpha^{I_i}\|_{\frac{1}{4}\rho} < \frac{\varepsilon}{\varepsilon_0} \rho_I , \quad \|\alpha^{\varphi_i}\|_{\frac{1}{4}\rho} < \frac{\varepsilon}{\varepsilon_0} \rho_\varphi , \quad i = 1, \dots, n .$$

<sup>46)</sup> Come si vedrà, un controllo anche povero di  $f_s$ , con la norma di  $f_s$  che cresce nientemeno che come  $s^{m_s}$ ,  $m > 0$ , sarà sufficiente a dedurre, alla fine, la stima esponenziale (6.3).

6.2. La trasformazione canonica elementare  $\mathcal{W}_s : \mathcal{D}_{\rho_{s+1}} \rightarrow \mathcal{D}_{\rho_s}$ .

Supponiamo di avere assegnata

$$H_s(I, \varphi; \varepsilon) = \omega \cdot I + \varepsilon g_s(I; \varepsilon) + \varepsilon^{s+1} f_s(I, \varphi; \varepsilon) ,$$

analitica e limitata in  $\mathcal{D}_{\rho_s}$ , ed eseguiamo una trasformazione canonica con il metodo di Lie: ovvero poniamo<sup>47)</sup>  $\mathcal{W}_s = \Phi_\chi = \Phi_\chi^1$ , e cerchiamo una funzione  $\chi$  tale che la nuova hamiltoniana  $H_{s+1} = H_s \circ \Phi_\chi$  abbia la forma voluta. La nuova hamiltoniana ha la forma

$$H_{s+1} = \omega \cdot I + \varepsilon g_s + \varepsilon^{s+1} f_s + \{\omega \cdot I, \chi\} + \mathcal{R}_2[\chi, \omega \cdot I] + \mathcal{R}_1[\chi, \varepsilon g_s + \varepsilon^{s+1} f_s] ,$$

con  $R_1$  e  $R_2$  definiti dalle (5.3), e se vogliamo che la quantità  $\varepsilon^{s+1} f_s + \{\omega \cdot I, \chi\}$  sia indipendente dagli angoli ci troviamo a dover risolvere l'equazione

$$\{\omega \cdot I, \chi\} + \varepsilon^{s+1} f_s = \varepsilon^{s+1} \langle f \rangle_\varphi ,$$

ovvero

$$\omega \cdot \frac{\partial \chi}{\partial \varphi} = \varepsilon^{s+1} (f - \langle f \rangle_\varphi) , \quad (6.6)$$

che è precisamente la (5.4) nel caso isocrono diofanteo. Se  $\chi$  è soluzione si ha poi

$$\begin{aligned} g_{s+1} &= g_s + \varepsilon^s \langle f_s \rangle_\varphi \\ \varepsilon^{s+2} f_{s+1} &= \mathcal{R}_2[\chi, \omega \cdot I] + \mathcal{R}_1[\chi, \varepsilon g_s + \varepsilon^{s+1} f_s] \end{aligned} \quad (6.7)$$

(si osservi che  $\chi$  che risolve la (6.6) è di ordine  $\varepsilon^{s+1}$ ; entrambi i resti allora sono di ordine almeno  $\varepsilon^{s+2}$ , e ciò giustifica la seconda delle (6.7)).

Procediamo allora in due passi:

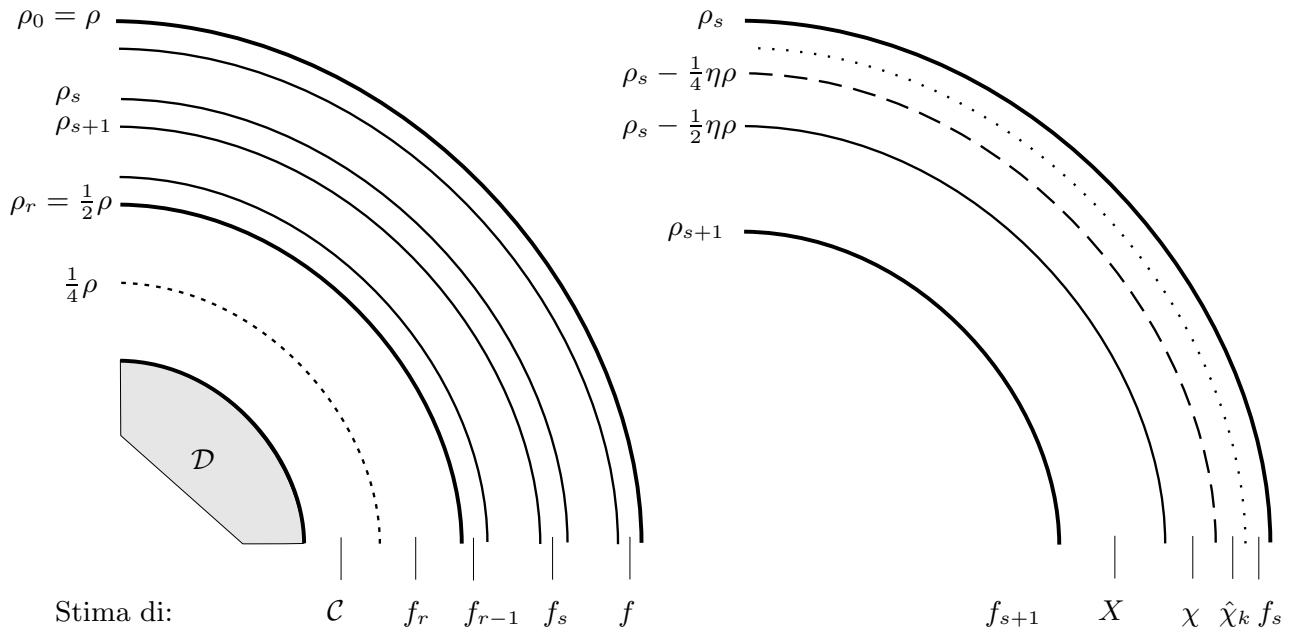
- i) usando il lemma 5.1 sulla serie di Fourier, assieme alla condizione diofantea, mostriamo che l'equazione (6.6) per  $\chi$  effettivamente si risolve, e stimiamo il campo vettoriale  $X$  di  $\chi$  in un generico insieme ridotto del tipo  $\mathcal{D}_{\rho_s - \theta\rho}$  (la stima servirà per  $\theta = \frac{1}{2}\eta$  e anche per  $\theta = \frac{1}{4}\eta$ );
- ii) usando i lemmi 5.3 e 5.4 mostriamo che la trasformazione canonica  $\mathcal{W}_s = \Phi_\chi : \mathcal{D}_{\rho_{s+1}} \rightarrow \mathcal{D}_{\rho_s}$  è ben definita, e stimiamo  $g_{s+1}$ ,  $f_{s+1}$  in termini di  $g_s$  e  $f_s$ , in modo da innescare un procedimento ricorsivo; come sottoprodotto otteniamo poi anche la stima di  $\mathcal{W}_s$  nell'insieme ridotto  $\mathcal{D}_{\rho_{s+1} - \frac{1}{4}\rho}$ , utile per la stima finale (6.5) sulla trasformazione canonica.

Per seguire le dimostrazioni può essere utile la figura 6.2, nella quale sono indicati i domini nei quali sono valide le diverse stime (nella figura si è fissato  $\theta = \frac{1}{2}\eta$ ).

Il primo punto è oggetto del seguente lemma:

---

<sup>47)</sup> l'indice  $s$  è omissso per semplicità da  $\chi$ .



**Figura 6.2.** I diversi domini e le relative stime.

**Lemma 6.2.** *Se  $f_s$  è analitica in  $\mathcal{D}_{\rho_s}$  e  $\omega$  soddisfa la condizione diofantea (6.2), allora esiste  $\chi$  analitica in  $\mathcal{D}_{\rho_s}$  che risolve la (6.6), e per ogni  $\theta$ ,  $0 < \theta < \frac{1}{2}$ , il suo campo vettoriale  $X$  è stimato da*

$$\|X\|_{\rho_s - \theta\rho} < C_0 \varepsilon^{s+1} \theta^{-2n-1} \frac{\|f_s\|_{\rho_s}}{E}, \tag{6.8}$$

con

$$C_0 = 2^{6n+1} n^n e^{-n}, \quad E = \gamma \rho_\varphi^{2n+1} \rho_I.$$

**Dimostrazione.** La (6.6) si risolve per serie di Fourier: posto

$$\chi(I, \varphi) = \varepsilon^{s+1} \sum_{k \in \mathbb{Z}^n} \hat{\chi}_k(I) e^{ik \cdot \varphi}, \quad f_s(I, \varphi) = \sum_{k \in \mathbb{Z}^n} \hat{f}_k(I) e^{ik \cdot \varphi},$$

l'equazione si risolve prendendo<sup>48)</sup>

$$\hat{\chi}_k(I) = \frac{\hat{f}_k(I)}{ik \cdot \omega}, \quad k \in \mathbb{Z}^n \setminus \{0\}; \quad \hat{\chi}_0(I) = 0.$$

Usando il lemma 5.1 non è difficile dimostrare che la serie per  $\chi$  converge, e vale la stima

$$\|\chi\|_{\rho_s - \frac{1}{2}\theta\rho} \leq (2^6 n/e)^n \varepsilon^{s+1} \gamma^{-1} \rho_\varphi^{-2n} \theta^{-2n} \|f_s\|_{\rho_s}. \tag{6.9}$$

<sup>48)</sup>  $\hat{\chi}_0$  è evidentemente arbitrario; scegliendolo in modo diverso si ottengono trasformazioni canoniche che differiscono solo per uno sfasamento negli angoli.



Si procede nel modo seguente: per l'analiticità di  $f_s$  in  $\mathcal{D}_{\rho_s}$  si ha

$$|\hat{\chi}_k(I)| \leq \frac{\|f_s\|_{\rho_s} e^{-|k|\rho_{s,\varphi}}}{|k \cdot \omega|}, \quad \rho_{s,\varphi} := (\rho_s)_\varphi = \text{componente } \varphi \text{ di } \rho_s.$$

Una parte dell'esponenziale si adopera per compensare il piccolo denominatore, che per la condizione diofantea è stimato da  $\gamma|k|^{-n}$ : precisamente, si scrive

$$|\hat{\chi}_k(I)| \leq (|k|^n e^{-\frac{1}{4}|k|\theta\rho_\varphi}) \gamma^{-1} \|f_s\|_{\rho_s} e^{-|k|(\rho_{s,\varphi} - \frac{1}{4}\theta\rho_\varphi)},$$

e si usa la disuguaglianza, di facile verifica,<sup>49)</sup>

$$|k|^n e^{-|k|\frac{1}{4}\theta\rho_\varphi} \leq (4n/e)^n \theta^{-n} \rho_\varphi^{-n};$$

segue allora per le componenti di Fourier di  $\chi$  una stima della forma richiesta nel lemma 5.1, precisamente

$$\begin{aligned} |\hat{\chi}_k(I)| &\leq \mathcal{F} e^{-|k|\sigma}, \quad \sigma = \rho_{s,\varphi} - \frac{1}{4}\theta\rho_\varphi \\ \mathcal{F} &= (4n/e)^n \gamma^{-1} \theta^{-n} \rho_\varphi^{-n} \|f_s\|_{\rho_s}. \end{aligned}$$

Usando il lemma 5.1, con  $\delta = \frac{1}{4}\theta\rho_\varphi$  (cosicché  $\sigma - \delta = \rho_{s,\varphi} - \frac{1}{2}\theta\rho_\varphi$ ), la (6.9) è immediata.<sup>50)</sup> Da questa disuguaglianza, con una semplice stima di Cauchy per  $X_I = -\frac{\partial\chi}{\partial\varphi}$ ,  $X_\varphi = \frac{\partial\chi}{\partial I}$  (con riduzione del dominio di  $\frac{1}{2}\theta\rho$ ) si trova subito

$$\|X\|_{\rho_s - \theta\rho} \leq 2(2^6 n/e)^n \varepsilon^{s+1} \gamma^{-1} \rho_\varphi^{-2n-1} \rho_I^{-1} \theta^{-2n-1} \|f_s\|_{\rho_s},$$

ovvero la (6.8); non si dimentichino i denominatori a secondo membro della (6.1). ■

Il secondo punto è oggetto del seguente “lemma iterativo”, vero cuore del teorema:

**Lemma 6.3.** *Se l'hamiltoniana*

$$H_s(I, \varphi; \varepsilon) = \omega \cdot I + \varepsilon g_s(I; \varepsilon) + \varepsilon^{s+1} f_s(I, \varphi; \varepsilon) \tag{6.10}$$

*è analitica e limitata in  $\mathcal{D}_{\rho_s}$ , ed è soddisfatta la condizione (su  $\varepsilon$ )*

$$C_1 \varepsilon^{s+1} \eta^{-2n-2} E^{-1} \|f_s\|_{\rho_s} < 1, \quad C_1 = 2^{2n+2} C_0, \tag{6.11}$$

*allora:*

*i) esiste la trasformazione canonica elementare  $(I, \varphi) = \mathcal{W}_s(I', \varphi')$ ,*

$$\mathcal{W}_s : \mathcal{D}_{\rho_{s+1}} \rightarrow \mathcal{D}_{\rho_s}, \tag{6.12}$$

<sup>49)</sup> In generale:  $x^\tau e^{-\delta x} \leq (\tau/e)^\tau \delta^{-\tau}$ , per ogni  $\tau, \delta > 0$  (si prende il max su  $x$ ).

<sup>50)</sup> Per semplicità di notazione, nella (6.9) si è ridotto in proporzione anche il dominio delle variabili  $I$ , ma non sarebbe stato necessario.

che muta  $H_s$  in

$$H_{s+1}(I', \varphi'; \varepsilon) = \omega \cdot I' + \varepsilon g_{s+1}(I'; \varepsilon) + \varepsilon^{s+2} f_{s+1}(I', \varphi'; \varepsilon) ,$$

analitica e limitata in  $\mathcal{D}_{\rho_{s+1}}$ ;

ii) valgono le stime

$$\|g_{s+1}\|_{\rho_{s+1}} \leq \|g_s\|_{\rho_s} + \varepsilon^s \|f_s\|_{\rho_s} \quad (6.13a)$$

$$\|f_{s+1}\|_{\rho_{s+1}} \leq C_1 \eta^{-2n-2} (\|g_s\|_{\rho_s} + 2\varepsilon^s \|f_s\|_{\rho_s}) E^{-1} \|f_s\|_{\rho_s} ; \quad (6.13b)$$

iii) nell'insieme ridotto  $\mathcal{D}_{\rho_s - \frac{1}{4}\rho}$  la trasformazione è  $\varepsilon^{s+1}$ -prossima all'identità, precisamente si ha

$$\|W_s - \text{Id}\|_{\rho_{s+1} - \frac{1}{4}\rho} \leq C_2 \varepsilon^{s+1} E^{-1} \|f_s\|_{\rho_s} , \quad C_2 = 2^{4n+2} C_0 . \quad (6.14)$$

**Dimostrazione.** Sia  $\chi$  la funzione costruita nel lemma 6.2. Per  $\theta = \frac{1}{2}\eta$  la stima (6.8) di  $X$  diventa

$$\|X\|_{\rho_s - \frac{1}{2}\eta\rho} \leq 2^{2n+1} C_0 \varepsilon^{s+1} \eta^{-2n-1} E^{-1} \|f_s\|_{\rho_s} ; \quad (6.15)$$

usando l'ipotesi (6.11) si ha allora

$$\|X\|_{\rho_s - \frac{1}{2}\eta\rho} \leq \frac{1}{2}\eta ,$$

perciò si può usare il lemma 5.3, con  $\sigma = \rho_s - \frac{1}{2}\theta\rho$  e  $\theta = \frac{1}{2}\eta$ , per garantire l'esistenza della trasformazione  $W_s = \Phi_\chi : \mathcal{D}_{\rho_{s+1}} \rightarrow \mathcal{D}_{\rho_s - \frac{1}{2}\eta\rho} \subset \mathcal{D}_{\rho_s}$ . La scelta di  $\chi$  è proprio quella che garantisce per  $H_{s+1}$  la forma (6.12), con  $g_s$  e  $f_s$  date dalle (6.7); mostriamo che da queste espressioni seguono le stime (6.13). La prima di esse, relativa a  $g_{s+1}$ , è banale conseguenza della prima delle (6.7). Per la seconda, relativa a  $f_{s+1}$ , ci serve invece la stima dei resti. Per  $\mathcal{R}_1$  si trova subito

$$\|\mathcal{R}_1[\chi, \varepsilon g_s + \varepsilon^{s+1} f_s]\|_{\rho_{s+1}} \leq \|L_\chi(\varepsilon g_s + \varepsilon^{s+1} f_s)\|_{\rho_s - \frac{1}{2}\eta\rho} ,$$

e applicando il lemma 5.4 (con  $\sigma = \rho_s$ ,  $\theta = \frac{1}{2}\eta$ )

$$\|\mathcal{R}_1[\chi, \varepsilon g_s + \varepsilon^{s+1} f_s]\|_{\rho_{s+1}} \leq 2\eta^{-1} \|X\|_{\sigma - \frac{1}{2}\eta\rho} (\varepsilon \|g_s\|_{\rho_s} + \varepsilon^{s+1} \|f_s\|_{\rho_s}) .$$

Per  $\mathcal{R}_2$  invece osserviamo preliminarmente che per la scelta stessa di  $\chi$  si ha

$$L_\chi^2(\omega \cdot I) = \varepsilon^{s+1} L_\chi(f - \langle f \rangle_\varphi) ;$$

segue subito (osservando che il dominio dell'integrale doppio ha area  $\frac{1}{2}$ )

$$\|\mathcal{R}_2[\chi, \omega \cdot I]\|_{\rho_{s+1}} \leq \frac{1}{2} \varepsilon^{s+1} \|L_X(f - \langle f \rangle_\varphi)\|_{\rho_s - \frac{1}{2}\eta\rho} ,$$

e per il lemma 5.4

$$\begin{aligned} \|\mathcal{R}_2[\chi, \omega \cdot I]\|_{\rho_{s+1}} &\leq \eta^{-1} \varepsilon^{s+1} \|X\|_{\sigma - \frac{1}{2}\eta\rho} \|f_s - \langle f_s \rangle_\varphi\|_{\rho_s} \\ &\leq 2\eta^{-1} \varepsilon^{s+1} \|X\|_{\sigma - \frac{1}{2}\eta\rho} \|f\|_{\rho_s} . \end{aligned}$$

Mettendo insieme le due stime, e usando la (6.15), segue facilmente la (6.13b).

Infine, per la stima (6.14) sulla trasformazione canonica nell'insieme ridotto  $\mathcal{D}_{\rho_{s+1} - \frac{1}{4}\rho}$ , basta usare la (5.18) con  $|t| = 1$ ,  $\sigma = \rho_{s+1}$ ,  $\theta = \frac{1}{4}$ , e poi la (6.8), ancora con  $\theta = \frac{1}{4}$ , per stimare  $X$ . ■

- **Osservazione.** Ci si può chiedere perchè, per stimare la trasformazione canonica, è stata necessaria la “grossa” riduzione del dominio di  $\frac{1}{4}\rho$ : ovviamente si poteva ottenere (usando nell'ultimo punto  $\theta = \frac{1}{2}\eta$ , anziché  $\theta = \frac{1}{4}\rho$ )

$$\|\mathcal{W}_s - \text{Id}\|_{\rho_{s+1}} \leq C'_2 \varepsilon^{s+1} \eta^{-2n-1} E^{-1} \|f_s\|_{\rho_s} , \quad C'_2 = 2^{2n+1} C_0 .$$

La presenza del fattore  $\eta^{-2n-1}$ , con  $\eta = \frac{1}{2r}$ , rende questa stima *molto* peggiore della (6.14): come si vedrà, se cercassimo di impiegare questa al posto della (6.14) nella dimostrazione del punto ii) della proposizione (la disuguaglianza (6.5)), perderemmo il prezioso fattore  $\varepsilon$  a numeratore.

### 6.3. Iterazione e conclusione della dimostrazione.

Poniamo ora  $H_0 = H$ , ovvero  $g_0 = 0$ ,  $f_0 = f$ ; dalle rispettive definizioni risulta anche  $\rho_0 = \rho$ ,  $\mathcal{D}_{\rho_0} = \mathcal{D}_\rho$ . Usando iterativamente il lemma 6.3, mostriamo che se risulta

$$\varepsilon \leq \frac{1}{2\Lambda} , \quad (6.16)$$

con

$$\Lambda = 2C_1 \eta^{-2n-2} E^{-1} \|f\|_\rho , \quad \eta = \frac{1}{2r} , \quad (6.17)$$

allora si possono eseguire  $r$  trasformazioni canoniche elementari, ottenendo  $H_s$ ,  $0 \leq s \leq r$ , della forma (6.10), con  $g_s$  e  $f_s$  analitiche in  $\mathcal{D}_{\rho_s}$ , ed ivi soddisfacenti le stime

$$\|g_s\|_{\rho_s} \leq \|f\|_\rho \sum_{j=0}^{s-1} 2^{-j} , \quad \|f_s\|_{\rho_s} \leq \Lambda^s \|f\|_\rho . \quad (6.18)$$

Per  $s = 0$  l'affermazione è banalmente vera; procedendo per induzione, supponiamola vera per un certo  $s$ ,  $0 \leq s < r$ , e mostriamola vera per  $s + 1$ . Per questo osserviamo innanzitutto che è soddisfatta la condizione (6.11) per l'esistenza della trasformazione canonica  $\mathcal{W}_s : \mathcal{D}_{\rho_{s+1}} \rightarrow \mathcal{D}_{\rho_s}$ ; infatti, dalle (6.16) e (6.18) si ottiene, come richiesto,

$$C_1 \varepsilon^{s+1} \eta^{-2n-2} E^{-1} \|f_s\|_{\rho_s} \leq \frac{1}{2} (\Lambda \varepsilon)^{s+1} < 1 .$$

Dalla (6.13a), usando l'ipotesi ricorrente, segue allora

$$\|g_{s+1}\|_{\rho_{s+1}} \leq \|f\|_\rho \left( \sum_{j=0}^{s-1} 2^{-j} + \varepsilon^s \Lambda^s \right) , \quad \varepsilon^s \Lambda^s \leq 2^{-s} ,$$

e la prima delle (6.18) è immediata. Infine, ancora per l'ipotesi ricorrente, si ha

$$\|g_s\|_{\rho_s} + 2\varepsilon^s \|f_s\|_{\rho_s} \leq \|f\|_{\rho} \left( \sum_{j=0}^{s-1} 2^{-j} + 2^{-(s-1)} \right) = 2 \|f\|_{\rho} ;$$

dalla (6.13b) segue allora

$$\begin{aligned} \|f_{s+1}\|_{\rho_{s+1}} &\leq 2C_1 \eta^{-2n-2} \|f\|_{\rho} E^{-1} \|f_s\|_{\rho_s} \\ &\leq \Lambda \|f_s\|_{\rho_s} , \end{aligned}$$

e anche la seconda delle (6.18) è soddisfatta.

A questo punto possiamo anche stimare la trasformazione canonica complessiva

$$\mathcal{C} = \mathcal{W}_0 \circ \dots \circ \mathcal{W}_{r-1} ,$$

nell'insieme ridotto finale  $\mathcal{D}_{\frac{1}{4}\rho} = \mathcal{D}_{\rho_r - \frac{1}{4}\rho}$ ; molto semplicemente, dalla (6.14) si ha

$$\|\mathcal{W}_s - \text{Id}\|_{\rho_{s+1} - \frac{1}{4}\rho} \leq C_2 \varepsilon^{s+1} \Lambda^s E^{-1} \|f\|_{\rho} \leq C_2 \varepsilon E^{-1} \|f\|_{\rho} 2^{-s} ,$$

e sommando gli spostamenti alle variabili dovuti alle successive trasformazioni,<sup>51)</sup> ovvero sommando su  $s$ , segue immediatamente la stima, *uniforme in  $r$* ,

$$\|\mathcal{C} - \text{Id}\|_{\frac{1}{4}\rho} < 2C_2 \varepsilon E^{-1} \|f\|_{\rho} . \quad (6.19)$$

Va detto che la (6.14) si può applicare perchè, come si verifica immediatamente, si ha

$$\mathcal{W}_s : \mathcal{D}_{\rho_{s+1} - \frac{1}{4}\rho} \rightarrow \mathcal{D}_{\rho_s - \frac{1}{4}\rho}$$

(ovvero, i domini ridotti su cui vale la (6.14) “si incastrano bene”).

- **Osservazione.** Si verifichi che senza la riduzione di  $\frac{1}{4}\rho = \mathcal{O}(1)$  del dominio di analiticità la stima per  $\mathcal{C}$  non è più di ordine  $\varepsilon$ .

Siamo ora assai prossimi alla conclusione della dimostrazione. Posto

$$\begin{aligned} \Lambda &= M r^m , & m &= 2n + 2 , & M &= C_3 E^{-1} \|f\|_{\rho} \\ C_3 &= 2^{2n+3} C_1 = 2^{4n+5} C_0 , \end{aligned}$$

la condizione (6.16) si scrive

$$M \varepsilon r^m \leq \frac{1}{2} , \quad r \leq r_{\max}(\varepsilon) := \left( \frac{1}{2M\varepsilon} \right)^{-\frac{1}{m}} ; \quad (6.20)$$

---

<sup>51)</sup> Equivalentemente, posto  $\mathcal{C}_s = \mathcal{W}_s \circ \dots \circ \mathcal{W}_{r-1}$ , si scrive l'identità

$$\mathcal{C} = \mathcal{C}_0 = \sum_{s=0}^{r-1} (\mathcal{W}_s - \text{Id}) \circ \mathcal{C}_{s+1}$$

(con  $\mathcal{C}_r = \text{Id}$ ), e si usa la disuguaglianza triangolare.

se questa disuguaglianza è soddisfatta, allora possiamo eseguire  $r$  passi perturbativi, ottenendo alla fine per il termine ancora contenente gli angoli, che denotiamo  $\mathcal{R} = \varepsilon^{r+1} f_r$ , la stima

$$\|\mathcal{R}\|_{\frac{1}{2}\rho} = \varepsilon (M \varepsilon r^m)^r \|f\|_{\rho} . \quad (6.21)$$

L'idea che fa passare da questa stima apparentemente di ordine  $r$  finito alla stima esponenziale – cioè a un andamento che decresce più rapidamente di qualunque potenza di  $\varepsilon$  – è quello di scegliere  $r$ , che fino ad ora è stato un parametro libero, dipendente invece da  $\varepsilon$ , e crescente per  $\varepsilon$  piccolo. La cosa più semplice, anche se non ottimale, è quella di prendere  $r = r(\varepsilon)$  in modo che sia

$$M \varepsilon r^m = e^{-1} ;$$

più precisamente, dovendo  $r$  essere un intero, prendiamo

$$r = \left[ \left( \frac{\varepsilon_0}{\varepsilon} \right)^{\frac{1}{m}} \right] , \quad \varepsilon_0 = \frac{1}{M e} = 2^{-4n-5} e^{-1} C_0^{-1} E / \|f\|_{\rho} \\ = C^{-1} E / \|f\|_{\rho} ,$$

ove  $C = 2^{10n+6} n^n e^{-n+1}$  è la costante numerica che compare nell'enunciato della proposizione. Si osservi che per questa scelta di  $r$  la (6.20) è soddisfatta. Si ottiene allora immediatamente per  $\mathcal{R}$  la stima esponenziale cercata,

$$\|\mathcal{R}\|_{\frac{1}{2}\rho} = \varepsilon e^{-[(\varepsilon_0/\varepsilon)^{1/m}]} \|f\|_{\rho} .$$

La dimostrazione della proposizione è così conclusa: l'hamiltoniana  $H' = H_r = H \circ \mathcal{C}$  ha proprio la forma (6.3), con  $g = g_r$  e  $\varepsilon e^{[(\varepsilon_0/\varepsilon)^{1/m}]} f' = \mathcal{R}$ ; le (6.4) sono soddisfatte, ed è anche immediato riconoscere nella (6.19) la stima cercata (6.5) per la trasformazione canonica  $\mathcal{C}$  (si è qui usata la disuguaglianza rozza  $C_2 < C$ ).

- **Osservazione.** Un procedimento migliore, per questa parte finale, è quello di scegliere per  $r(\varepsilon)$  il valore ottimale, ovvero quello che dà il minimo di  $\mathcal{R}$ . Derivando rispetto a  $r$  il logaritmo di  $\mathcal{R}$  si trova immediatamente il valore ottimale

$$r^* = \frac{1}{e} (M \varepsilon)^{-1/m} ,$$

di cui poi va presa la parte intera. Il risultato non è molto diverso dal precedente, ma il metodo è concettualmente migliore.<sup>52)</sup>

---

<sup>52)</sup> Questo modo di procedere è tipico degli sviluppi in serie asintotici, introdotti da Poincaré alla fine del secolo scorso. Vale la pena di osservare che Poincaré aveva già tutti gli elementi per dimostrare un teorema come questo (o come altri teoremi di teoria delle perturbazioni, come il teorema KAM), *tranne l'interesse per il risultato*: senza la misura di Lebesgue, l'insieme delle  $\omega$  che soddisfano la condizione diofantea appare bizzarro e poco significativo.

## 7. Sistemi non isocroni: un passo perturbativo

Questo paragrafo è dedicato allo studio perturbativo dei sistemi non isocroni, precisamente del tipo

$$H(I, \varphi) = h(I) + \varepsilon f(I, \varphi), \quad I \in \mathcal{B} \subset \mathbb{R}^n, \quad \varphi \in \mathbb{T}^n \quad (7.1)$$

con

$$\det \left( \frac{\partial^2 h}{\partial I_i \partial I_j} \right) \neq 0 \quad \text{in } \mathcal{B}. \quad (7.2)$$

Studieremo per prima cosa un modello particolare, facile da seguire anche grazie al suo significato fisico, e poi procederemo generalizzando progressivamente (per quanto si può) le ipotesi.

### 7.1. Il modello dei rotatori

Consideriamo allora il sistema hamiltoniano

$$H(I, \varphi) = \frac{1}{2} \sum_{i=1}^n I_i^2 - \varepsilon \sum_{i=1}^n \cos(\varphi_i - \varphi_{i-1}), \quad \varphi_0 := \varphi_n; \quad (7.3)$$

si tratta di una catena di rotatori, con un piccolo potenziale di accoppiamento che è minimo quando  $\varphi_1 = \varphi_2 = \dots = \varphi_n$ .<sup>53)</sup> La condizione di genericità di  $f$  che appare nel teorema di Poincaré 4.1 (punto ii.) è con evidenza violata: nella serie di Fourier compaiono solo le armoniche del tipo

$$k = \pm(0, \dots, 0, 1, -1, 0, \dots, 0). \quad (7.4)$$

Il fatto che la perturbazione sia a media nulla, e indipendente dalle  $I$ , semplifica lo studio ma non la sostanza; il fatto che l'interazione sia locale (ciascun rotatore interagisce solo con i vicini, ovvero sono presenti solo i  $k$  della forma (7.4)) aiuta a leggere i risultati, ma ancora non tocca la sostanza (è invece rilevante il fatto che  $f$  contenga un numero finito di componenti di Fourier). Grazie alla forma particolarmente semplice di  $h$ , le frequenze coincidono con le ampiezze:  $\omega(I) = I$ .

Eseguiamo un passo perturbativo (ad esempio) col metodo di Lie, ponendo  $\chi(I, \varphi) = \varepsilon \tilde{\chi}(I, \varphi)$ ; la nuova hamiltoniana si trova subito essere

$$H' = h + \varepsilon(f + \{h, \tilde{\chi}\}) + \varepsilon^2 f', \quad (7.5)$$

con

$$f' = \frac{1}{2} \{ \{h, \tilde{\chi}\}, \tilde{\chi} \} + \{f, \tilde{\chi}\} + \mathcal{O}(\varepsilon). \quad (7.6)$$

Si elimina allora la dipendenza dagli angoli all'ordine  $\varepsilon$  se  $\{h, \tilde{\chi}\} + f = 0$ , ovvero se

$$I \cdot \frac{\partial \tilde{\chi}}{\partial \varphi} = - \sum_{i=1}^n \cos(\varphi_i - \varphi_{i-1}),$$

---

<sup>53)</sup> “Modello XY” a una dimensione, nel linguaggio della meccanica statistica.

e questa equazione è risolta da

$$\tilde{\chi}(I, \varphi) = - \sum_{i=1}^n \frac{\sin(\varphi_i - \varphi_{i-1})}{I_i - I_{i-1}} . \quad (7.7)$$

Si osservi che al numeratore appaiono le armoniche di  $f$ , mentre i denominatori hanno la forma  $k \cdot I$  con  $k$  del tipo (7.4). Ovviamente  $\tilde{\chi}$  non è definita sulle risonanze  $I_i - I_{i-1} = 0$ . Quanto al resto  $f'$ , dalla (7.6) si trova, grazie a  $\{h, \tilde{\chi}\} = -f$ ,  $f' = \frac{1}{2}\{f, \tilde{\chi}\} + \mathcal{O}(\varepsilon)$ , e a conti fatti

$$f'(I, \varphi) = \frac{1}{2} \sum_{i=1}^n \frac{1 - \cos 2(\varphi_i - \varphi_{i-1})}{(I_i - I_{i-1})^2} + \frac{1}{4} \sum_{i=1}^n \left[ \cos(\varphi_{i+1} - \varphi_{i-1}) - \cos(\varphi_{i+1} - 2\varphi_i + \varphi_{i-1}) \right] \left[ \frac{1}{(I_{i+1} - I_i)^2} - \frac{1}{(I_i - I_{i-1})^2} \right] + \mathcal{O}(\varepsilon) .$$

I fatti significativi sono che

- compaiono a denominatore i quadrati dei denominatori presenti in  $\tilde{\chi}$ ;
- compaiono nuove armoniche, precisamente le somme dei  $k$  già presenti in  $f$ , ma l'interazione mantiene la forma di polinomio di Fourier.

L'analisi di  $f'$  mostra che se siamo ben fuori da tutte le risonanze, diciamo  $I_i - I_{i-1} = \mathcal{O}(1)$  per ogni  $i$ , allora abbiamo guadagnato un ordine in  $\varepsilon$ : la nuova perturbazione è  $\mathcal{O}(\varepsilon^2)$ , e modifica le (nuove) azioni solo su scala di tempo non inferiore a  $\varepsilon^{-2}$ . *Se invece ci avviciniamo alle risonanze, fino a  $I_i - I_{i-1} = \mathcal{O}(\sqrt{\varepsilon})$ , allora la procedura perturbativa non porta nessun guadagno:  $f'$  è grande quanto  $f$  (vi sono ovviamente tutte le possibilità intermedie).*

Fissiamo ora l'attenzione su una particolare risonanza, e per fissare le idee prendiamo

$$|I_2 - I_1| \leq \sqrt{\varepsilon} , \quad |I_i - I_{i-1}| \gg \sqrt{\varepsilon} \quad \text{per } i > 2 .$$

Possiamo ancora (utilmente) fare un passo perturbativo in prossimità della risonanza, se rinunciamo ad eliminare completamente la dipendenza della perturbazione dagli angoli, e usiamo, al posto della (7.7),

$$\tilde{\chi}(I, \varphi) = - \sum_{i \neq 2} \frac{\sin(\varphi_i - \varphi_{i-1})}{I_i - I_{i-1}} .$$

Con evidenza la nuova hamiltoniana ha ora la forma

$$H'(I, \varphi) = h(I) + \varepsilon g(\varphi) + \varepsilon^2 f'(I, \varphi) , \quad g(\varphi) = - \cos(\varphi_2 - \varphi_1) ,$$

e la nuova perturbazione è effettivamente di ordine  $\varepsilon^2$ ; la nuova hamiltoniana è dunque una piccola perturbazione di

$$h'(I, \varphi) = \mathcal{H}(I_1, I_2, \varphi_1, \varphi_2) + \frac{1}{2} \sum_{i \geq 3} I_i^2 \quad (7.8)$$

$$\mathcal{H}(I_1, I_2, \varphi_1, \varphi_2) = \frac{1}{2}(I_1^2 + I_2^2) - \varepsilon \cos(\varphi_2 - \varphi_1) .$$

Questa hamiltoniana con evidenza è integrabile: essa infatti rappresenta  $n - 2$  rotatori liberi, e una coppia di rotatori interagenti con sole forze interne, per i quali si conservano l'energia  $\mathcal{H}$  e il momento angolare totale  $I_1 + I_2$ .

- **Esercizio.** Si dimostri che un'hamiltoniana della forma  $h'(I, \varphi) = h(I) + g(I, \varphi)$  è integrabile ogniqualvolta  $g$  dipende da una sola combinazione di angoli,  $g = \hat{g}(I)e^{ik \cdot \varphi} + \text{c.c.}$  [Suggerimento: si trovino  $n - 1$  combinazioni lineari delle azioni che si conservano].

Per i primi due rotatori vale la pena di passare a nuove variabili canoniche  $J, A, \psi, \alpha$  date da

$$J = I_1 + I_2, \quad A = \frac{1}{2}(I_2 - I_1), \quad \psi = \frac{1}{2}(\varphi_1 + \varphi_2), \quad \alpha = \varphi_2 - \varphi_1$$

(i nuovi angoli *non* corrono sul toro standard  $\mathbb{T}^2$  — la periodicità è alterata, la matrice della trasformazione non è intera — ma ciò non avrà conseguenze). L'hamiltoniana  $\mathcal{H}$  è mutata in

$$\mathcal{H}' = \frac{1}{2} \left( \frac{J^2}{2} + 2A^2 \right) - \varepsilon \cos \alpha,$$

e rappresenta un rotatore libero di “massa” doppia (il baricentro) e un pendolo, tra loro disaccoppiati. Il momento totale  $J$  resta costante, e l'angolo  $\psi$  ad esso coniugato avanza uniformemente, mentre  $A$  e  $\alpha$  oscillano; corrispondentemente  $I_1$  e  $I_2$  oscillano attorno a  $\frac{1}{2}J$ , mentre il moto di  $\varphi_1$  e  $\varphi_2$  è una rotazione uniforme, con sovrapposte oscillazioni limitate in controfase. Si osservi che per  $\varepsilon$  piccolo il pendolo ha pulsazione lenta  $\omega = \sqrt{2\varepsilon}$ , e corrispondentemente l'ampiezza delle oscillazioni di  $A$  (delle oscillazioni di  $I_1, I_2$ ) è piccola,  $\mathcal{O}(\sqrt{\varepsilon})$ .

- **Osservazione.** La situazione è un po' diversa dai casi tipici incontrati studiando il principio della media: tuttavia, *grazie al fatto che l'ampiezza delle oscillazioni è limitata da  $\sqrt{\varepsilon}$ , e che la perturbazione  $f'$  muove le azioni solo su tempi  $\mathcal{O}(\varepsilon^{-2})$ , il principio della media resta valido.*
- **Osservazione.** Si confronti la (7.8) con il controesempio (3.3) al principio della media. Il segno positivo dentro a  $h$  è essenziale.
- **Esercizio.** Si studi la situazione analoga per  $h'(I, \varphi) = h(I) + \varepsilon g(I, \varphi)$ , quando  $g$  contiene una sola componente di Fourier.

Consideriamo ora il caso, più interessante, in cui vi siano due risonanze consecutive: per fissare le idee

$$|I_2 - I_1| \simeq \sqrt{\varepsilon}, \quad |I_3 - I_2| \simeq \sqrt{\varepsilon}.$$

Procedendo in analogia al caso precedente, si ottiene subito

$$h'(I, \varphi) = \mathcal{H}(I_1, I_2, I_3, \varphi_1, \varphi_2, \varphi_3) + \frac{1}{2} \sum_{i \geq 4} I_i^2 + \mathcal{O}(\varepsilon^2),$$

questa volta con

$$\mathcal{H} = \frac{1}{2}(I_1^2 + I_2^2 + I_3^2) - \varepsilon[\cos(\varphi_2 - \varphi_1) + \cos(\varphi_3 - \varphi_2)]. \quad (7.9)$$



Questa volta l'hamiltoniana  $\mathcal{H}$  non è integrabile. Per capire, si può anche qui separare il moto medio, o più semplicemente (come si sarebbe potuto fare anche prima) restringersi a dati iniziali<sup>54)</sup> con  $J := I_1 + I_2 + I_3 = 0$  ( $J$  poi si conserva). In tal caso le tre azioni sono tutte  $\mathcal{O}(\sqrt{\varepsilon})$ , ed è spontaneo eseguire il riscaldamento (canonico in senso generalizzato, paragrafo 1.5)

$$I = \sqrt{\varepsilon} \tilde{I}, \quad t = \tau/\sqrt{\varepsilon},$$

che muta  $\mathcal{H}$  nell'hamiltoniana *indipendente da  $\varepsilon$*

$$\tilde{\mathcal{H}} = \frac{1}{2}(\tilde{I}_1^2 + \tilde{I}_2^2 + \tilde{I}_3^2) - [\cos(\varphi_2 - \varphi_1) + \cos(\varphi_3 - \varphi_2)]. \quad (7.10)$$

Questo è un sistema *fortemente accoppiato*, per il quale è ragionevole aspettarsi anche moti caotici (mettendo un piccolo parametro  $\mu$  davanti a uno dei coseni, si potrebbe anche dimostrare l'esistenza di un punto omoclino).<sup>55)</sup> Si osservi tuttavia, ed è un passaggio essenziale, che grazie alla conservazione dell'energia i moti delle  $\tilde{I}$ , per ogni dato iniziale, anche se caotici sono limitati (la parte cinetica  $h$  è definita positiva, la parte potenziale è limitata). Corrispondentemente anche le  $I$  descrivono moti limitati: le  $I$  oscillano, eventualmente in modo caotico, su scala  $\sqrt{\varepsilon}$ , e i tempi caratteristici, diciamo i tempi necessari a variazioni delle azioni  $\mathcal{O}(\sqrt{\varepsilon})$ , sono  $\mathcal{O}(1/\sqrt{\varepsilon})$ . La conservazione dell'energia si poteva anche applicare direttamente all'hamiltoniana (7.9), con energie  $\mathcal{O}(\varepsilon)$ . E' importante osservare che *anche in questo caso il principio della media, benché il sistema non sia integrabile e siano possibili moti caotici, continua a valere*: infatti su scala di tempo  $\mathcal{O}(\varepsilon^{-1})$  la perturbazione  $\mathcal{O}(\varepsilon^2)$  non muove significativamente le azioni, e l'ampiezza delle oscillazioni di tutte le azioni è piccola con  $\varepsilon$ .

#### Studio numerico dell'hamiltoniana dei rotatori.<sup>56)</sup>

I risultati si riferiscono a una catena di otto rotatori, con hamiltoniana (7.3). Le quantità calcolate numericamente comprendono:

(a) La media temporale delle azioni fino al tempo  $T$ ,

$$\overline{I_j}(T) = \frac{1}{T} \int_0^T I_j(t) dt.$$

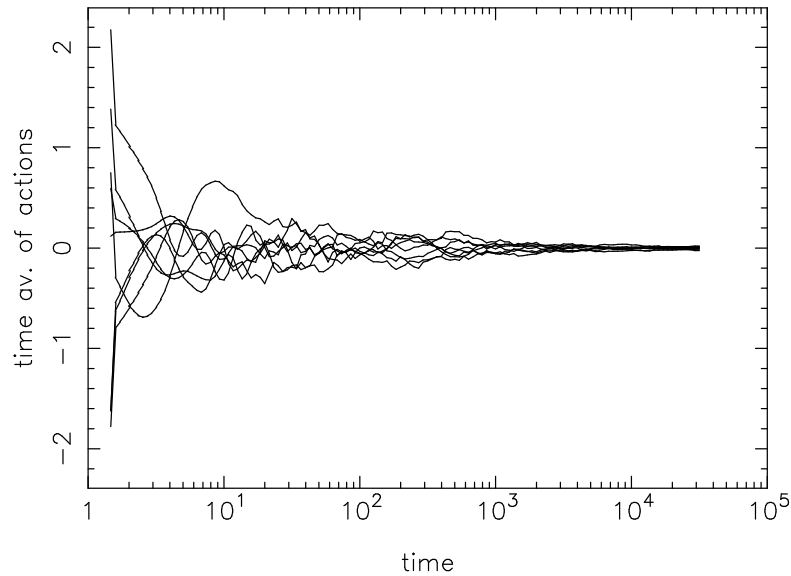
In pratica, se  $s$  denota il "passo di integrazione" impiegato nell'algoritmo numerico, la quantità calcolata numericamente è  $\overline{I_j} = (Ns)^{-1} \sum_{k=0}^{N-1} I_j(ks)$ . La media sul tempo

<sup>54)</sup> Operazione in realtà non restrittiva: se inizialmente  $I_1 + I_2 + I_3 = J^o \neq 0$ , si effettua la traslazione  $I'_j = I_j - \frac{1}{3}J^o$ ,  $\varphi' = \varphi - \frac{1}{3}J^o t$ ; non è difficile vedere che il moto delle  $I', \varphi'$  è ancora descritto dall'hamiltoniana (7.9).

<sup>55)</sup> Vale però la pena di osservare che per energie molto piccole ( $I_j$  prossime a zero, differenze  $\varphi_j - \varphi_{j-1}$  piccole) il sistema (7.10), e con esso il sistema di partenza, diviene nuovamente prossimo a un sistema integrabile, precisamente a un sistema di oscillatori armonici. Per rendersene conto basta sviluppare in serie i coseni, e trattare i termini non quadratici come piccola perturbazione (si ripensi al modello di Fermi–Pasta–Ulam, cap. I par. 7.1). Dunque i moti che si svolgono molto vicino a una risonanza (molto più vicino di  $\sqrt{\varepsilon}$ ) saranno ancora regolari per tempi lunghi (ma con altre azioni).

<sup>56)</sup> I risultati qui riportati sono tratti da: G. Benettin, L. Galgani, A. Giorgilli, *Numerical Investigation on a Chain of Weakly Coupled Rotators, in the Light of Classical Perturbation Theory*. Nuovo Cimento **89B**, 103 (1985).

eps = 0.9



**Figura 7.1.**  $\overline{I_j}(T)$  in funzione di  $T$ , per  $\varepsilon = 0.9$ .

infinito sarà denotata con  $\overline{I_j}^2(\infty)$ ; ovviamente il calcolo numerico deve arrestarsi a un tempo  $T_{\max}$  considerato grande.

- (b) La “funzione di autocorrelazione” normalizzata  $g_j$ , relativa a ciascuna azione  $I_j$ , così definita:

$$G_j(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I_j(t + \tau) I_j(t) dt$$

$$g_j(\tau) = \frac{G_j(\tau) - \overline{I_j}^2(\infty)}{\sigma^2},$$

ove  $\sigma^2 = G_j(0) - \overline{I_j}^2(\infty)$ . Il calcolo,<sup>57)</sup> al pari del precedente, è discretizzato e arrestato a  $T_{\max}$ .

- (c) Lo “spettro di potenza”  $W_j$  di  $I_j$ , la cui definizione, includendo già la discretizzazione e il troncamento a  $T_{\max} = sN_{\max}$ , è

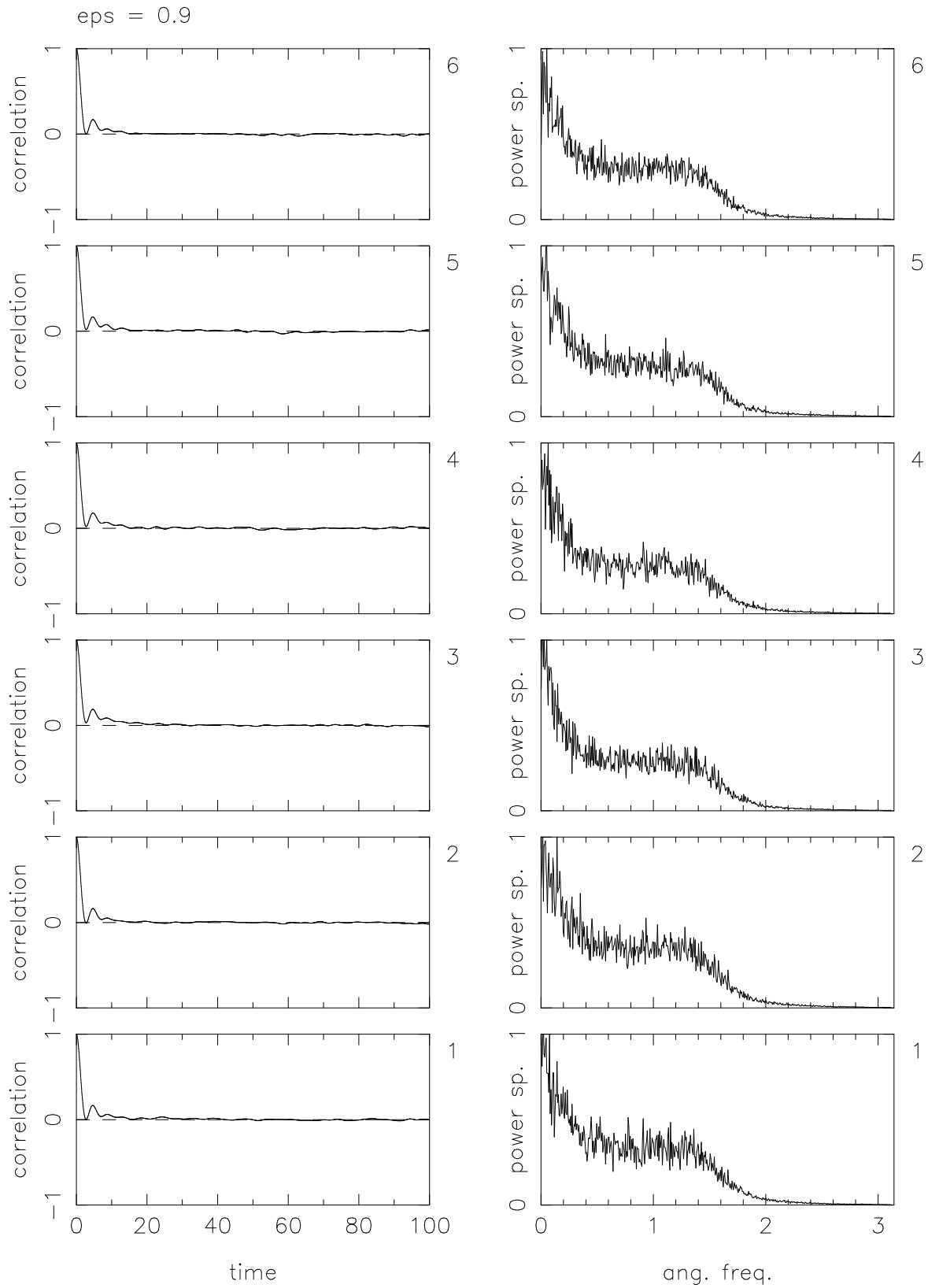
$$W_j\left(\frac{2\pi k}{N_{\max}}\right) = \left| \mathcal{I}_j\left(\frac{2\pi k}{N_{\max}}\right) \right|^2$$

$$\mathcal{I}_j\left(\frac{2\pi k}{N_{\max}}\right) = \frac{1}{sN_{\max}} \sum_{s=0}^{N_{\max}-1} I_j(ks) e^{-2\pi i ks / N_{\max}}.$$

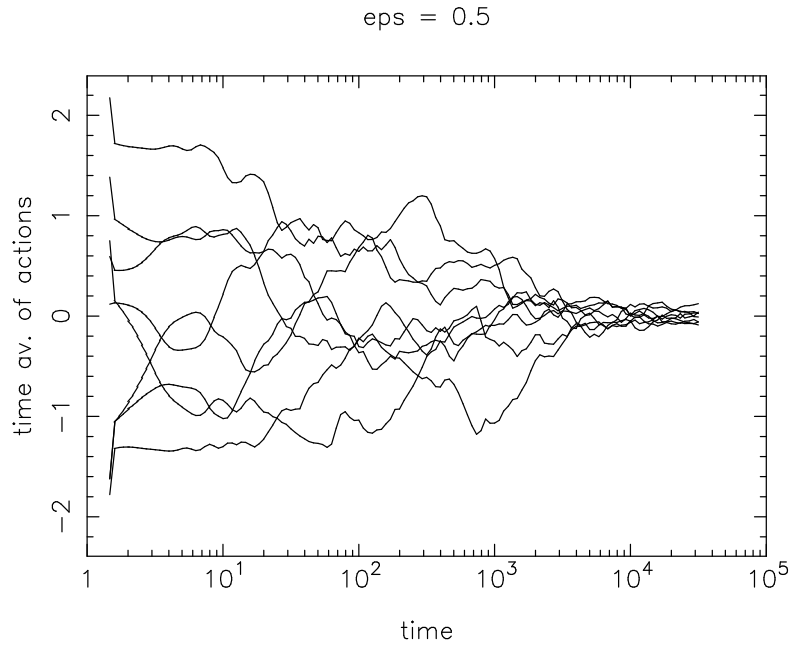
Lo spettro di potenza risulta essere la trasformata di Fourier della funzione di correlazione, pertanto esso non fornisce un’informazione indipendente, ma è ugualmente utile prenderlo in considerazione.

La figura 7.1 mostra il comportamento delle medie  $\overline{I_j}$  per grande accoppiamento,  $\varepsilon = 0.9$ . Il risultato è consistente con l’idea che il sistema, in queste condizioni, sia ergodico, cosicché

<sup>57)</sup> Il calcolo sfrutta la proprietà sotto enunciata che lo spettro di potenza è la trasformata di Fourier di  $g$ : ovvero prima si calcola lo spettro di Fourier e poi, antitrasformando, la funzione di correlazione.



**Figura 7.2.** Dal basso in alto:  $g_j$  (sinistra) e  $W_j$  (destra),  $j = 1, \dots, 6$ ;  $\varepsilon = 0.9$ .

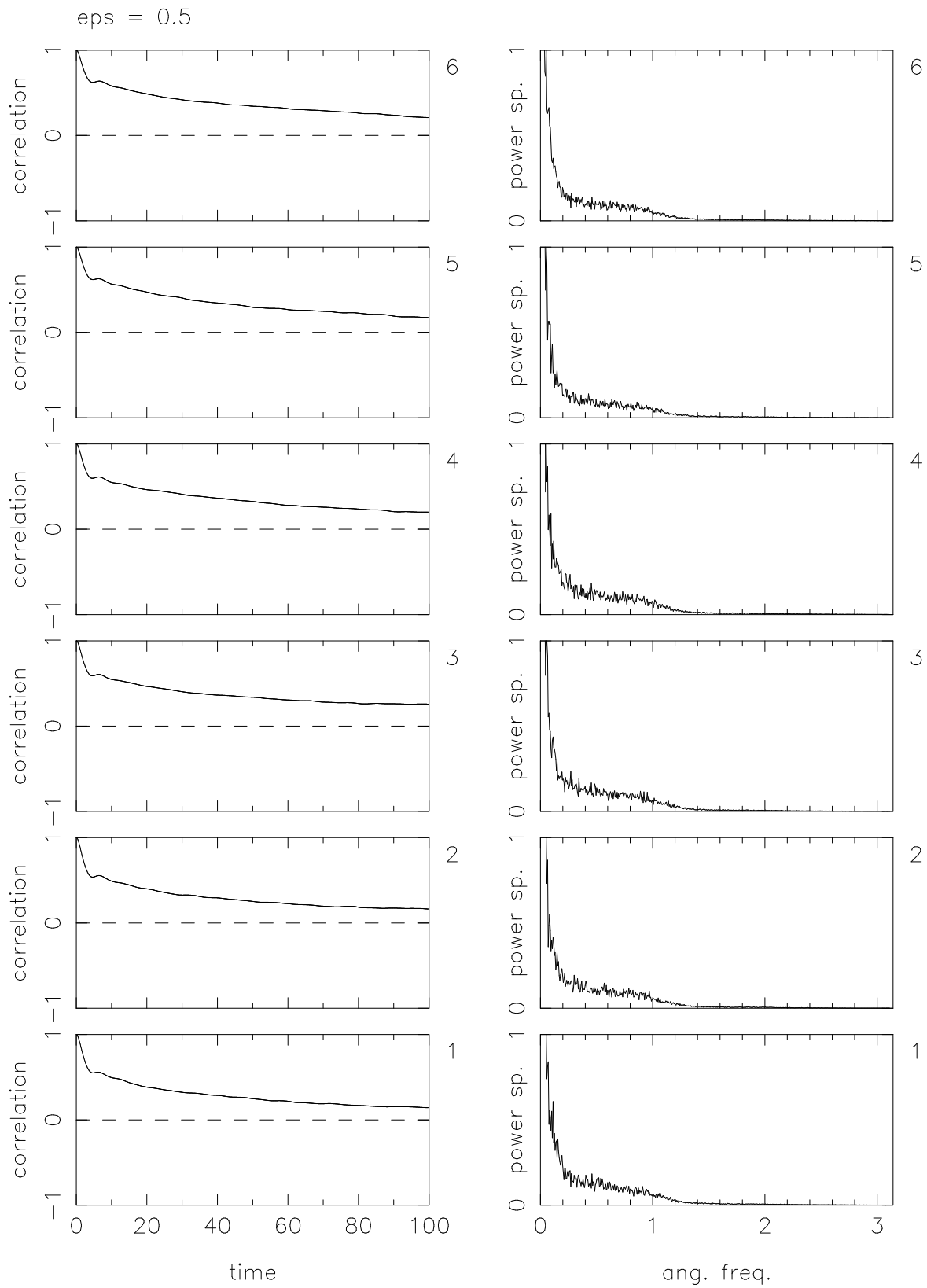


**Figura 7.3.** Simile alla figura 7.1, per  $\varepsilon = 0.5$ .

$\bar{I}_j(\infty) = 0$  per ogni  $j$ . La figura 7.2 mostra nelle stesse condizioni 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 una banda continua, che si estende da zero a un valore  $\omega_{\max}$  prossimo a  $\pi$ , 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.

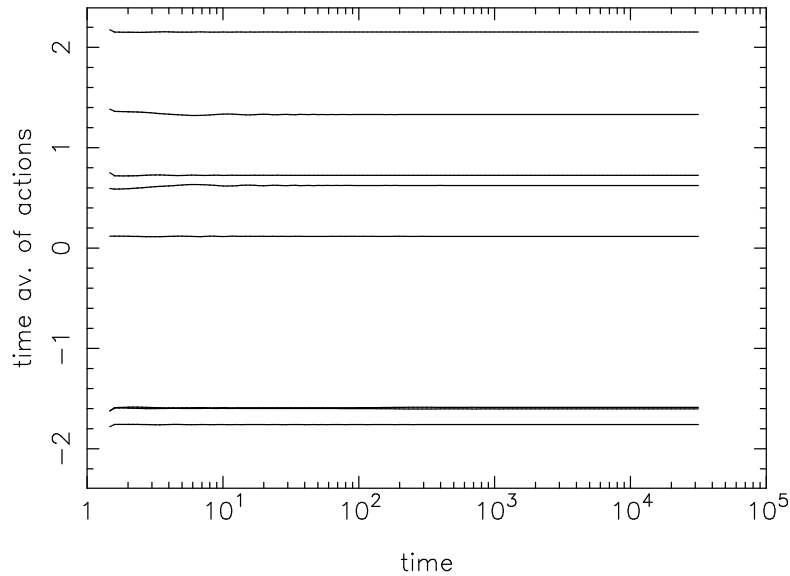
Le figure 7.3 e 7.4 riportano le medesime quantità, ma per accoppiamento più basso,  $\varepsilon = 0.5$ . L'aspetto qualitativo non è molto diverso, ma con evidenza la scala di tempo sulla quale si manifestano i diversi fenomeni si è parecchio dilatata; corrispondentemente la banda continua dello spettro è più stretta. Le figure successive 7.5 e 7.6 si riferiscono invece a un valore di  $\varepsilon$  sensibilmente più basso, precisamente  $\varepsilon = 0.03$ , e a 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 appare discreto; si potrebbe vedere che le linee nello spettro del  $j$ -esimo rotatore sono esattamente a  $\omega^+ = I_{j+1}$  e  $\omega^- = I_i - I_{j-1}$ .

Le successive figure 7.7 e 7.8 si riferiscono al caso risonante, precisamente alla doppia risonanza  $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\varepsilon}$ , con  $\sum_{j=1}^3 I_j = 0$ ; il valore di  $\varepsilon$  è 0.03. Con evidenza i rotatori  $1 - 2 - 3$  formano un sottosistema 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. 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 un aspetto essenziale



**Figura 7.4.** Simile alla figura 7.2, per  $\varepsilon = 0.5$ .

eps = 0.03 (nonresonant)



**Figura 7.5.** Simile alle figure 7.1 e 7.3, per  $\varepsilon = 0.03$  e dato iniziale non risonante.

del moto risonante — si può fare riducendo  $\varepsilon$ , ad esempio, di un fattore 4, e controllando se il tempo di decorrelazione e l'ampiezza della banda continua sono rispettivamente aumentati e diminuiti di un fattore 2. La verifica è riportata in figura 7.9. La parte superiore si riferisce a  $\varepsilon = 0.03$ , quella inferiore a  $\varepsilon = 0.0075$ . Per quest'ultimo valore di  $\varepsilon$  le scale dei tempi e delle frequenze sono rispettivamente raddoppiate e dimezzate rispetto al caso  $\varepsilon = 0.03$ , e come si vede, le figure si assomigliano in modi significativo (per rendere più leggibili le figure con lo spettro, la scala delle frequenze è stata dilatata, allo stesso modo per i due valori di  $\varepsilon$ , rispetto alle figure precedenti).

### 7.2. Il ruolo della convessità di $h$

Ci proponiamo qui di generalizzare ad hamiltoniane del tipo (7.1) i risultati ottenuti per il modello dei rotatori. Per quanto riguarda la perturbazione, l'unica proprietà rilevante è che essa contenga un numero finito di componenti di Fourier: diciamo

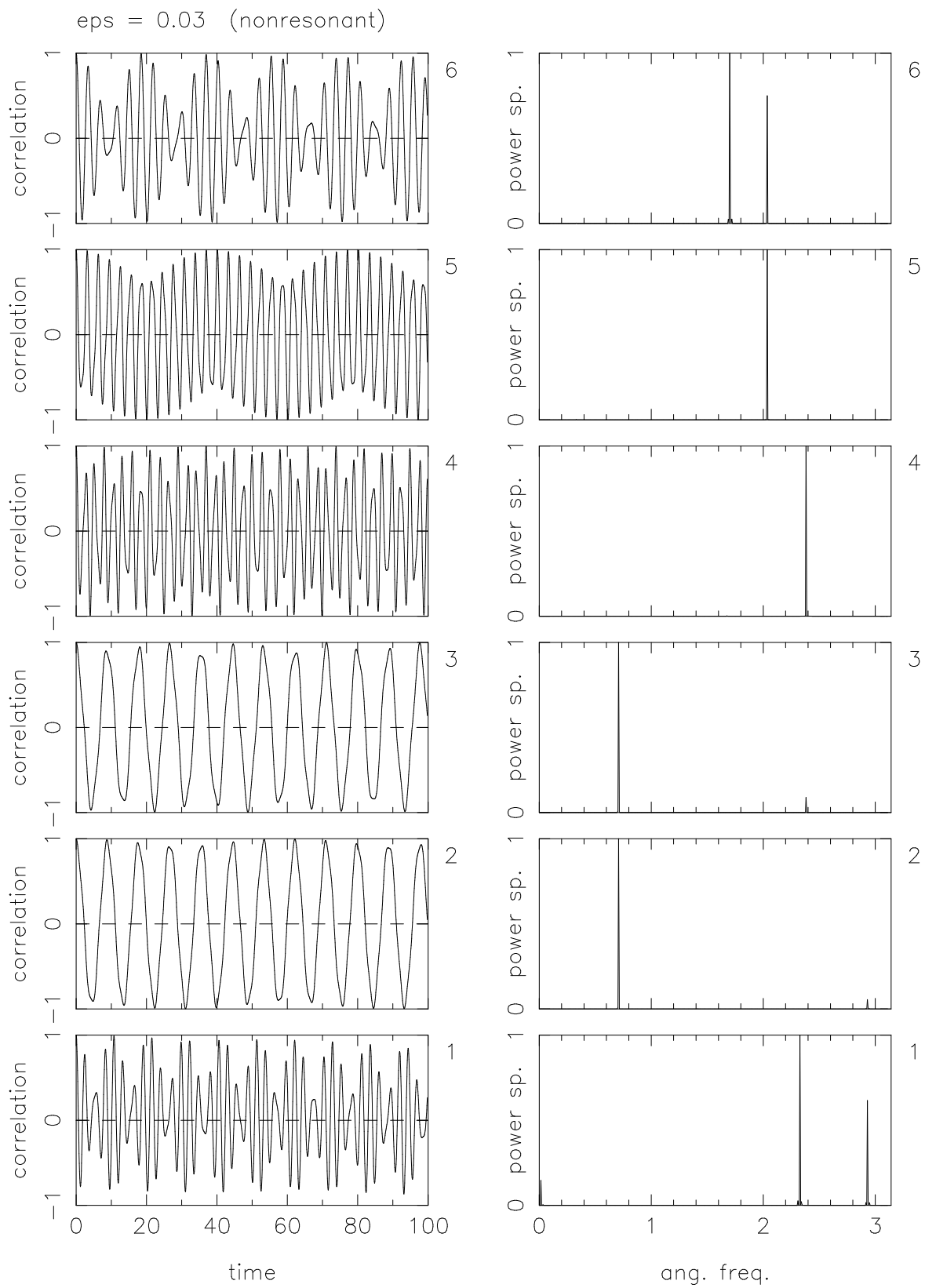
$$f(I, \varphi) = \sum_{|k| \leq K} \hat{f}_k(I) e^{ik \cdot \varphi} , \tag{7.11}$$

per qualche  $K > 0$ . Per quanto riguarda  $h$ , la forma particolarissima dell'hamiltoniana dei rotatori certo non è necessaria, tuttavia (si riveda il controesempio (3.3)) qualche cosa certamente è necessario supporre, se si vuole ottenere la stabilità delle azioni su tempi  $\mathcal{O}(\varepsilon^{-1})$ .

L'ipotesi adeguata su  $h$  risulta essere un'ipotesi di convessità, che nella formulazione più semplice si scrive

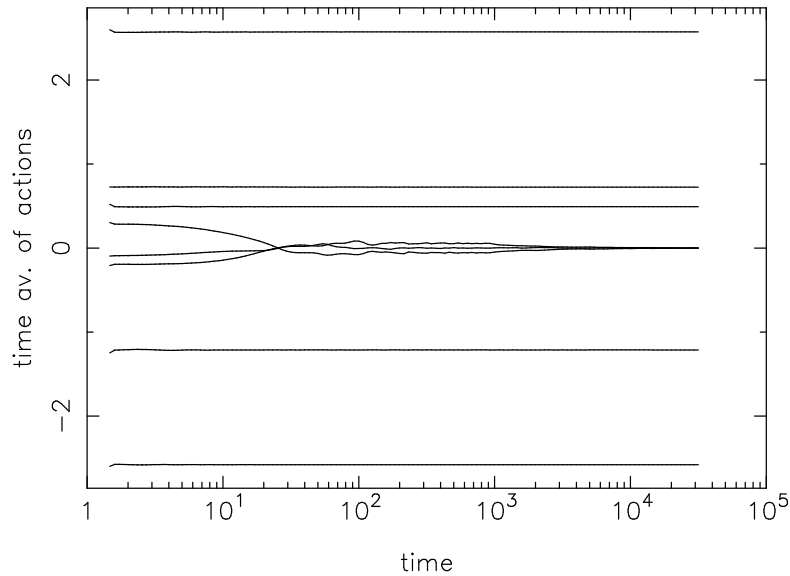
$$h''(I) v \cdot v \geq m v \cdot v \quad \forall v \in \mathbb{R}^n \setminus \{0\} , \quad \forall I \in \mathcal{B} ,$$

ove  $h''$  è la matrice hessiana  $(\frac{\partial^2 h}{\partial I_i \partial I_j})$ , e  $m$  è una costante positiva detta costante di convessità. Al di fuori delle risonanze in realtà la convessità non serve: l'equazione per  $\tilde{\chi}$  è la



**Figura 7.6.** Simile alle figure 7.2 e 7.4, per  $\varepsilon = 0.03$  e dato iniziale non risonante.

eps = 0.03 (resonant)



**Figura 7.7.** Simile alla figura 7.5, per  $\varepsilon = 0.03$  e dato iniziale nella risonanza  $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\varepsilon}$ .

solita

$$\omega \cdot \frac{\partial \tilde{\chi}}{\partial \varphi} = f - \langle f \rangle_\varphi, \quad \omega = \frac{\partial h}{\partial I},$$

ed è risolta da

$$\tilde{\chi}(I, \varphi) = \sum_{0 < |k| \leq K} \frac{\hat{f}_k(I)}{ik \cdot \omega} e^{ik \cdot \varphi};$$

la nuova hamiltoniana è

$$\tilde{H} = h + \varepsilon g + \varepsilon^2 \tilde{f},$$

con  $g = \langle f \rangle_\varphi$  indipendente dagli angoli e

$$\tilde{f} = \frac{1}{2} \{ \{ h, \tilde{\chi} \}, \tilde{\chi} \} + \{ f, \tilde{\chi} \} + \mathcal{O}(\varepsilon) = \frac{1}{2} \{ f, \tilde{\chi} \} + \mathcal{O}(\varepsilon).$$

L'espressione di  $\tilde{f}$  è complicata, ma si vede facilmente che compaiono, accanto a termini meno singolari, termini contenenti i quadrati dei piccoli denominatori, del tipo

$$\frac{[h''(I) k \cdot l] \hat{f}_k(I) \hat{f}_l(I)}{(k \cdot \omega(I))^2} e^{i(k+l) \cdot \varphi};$$

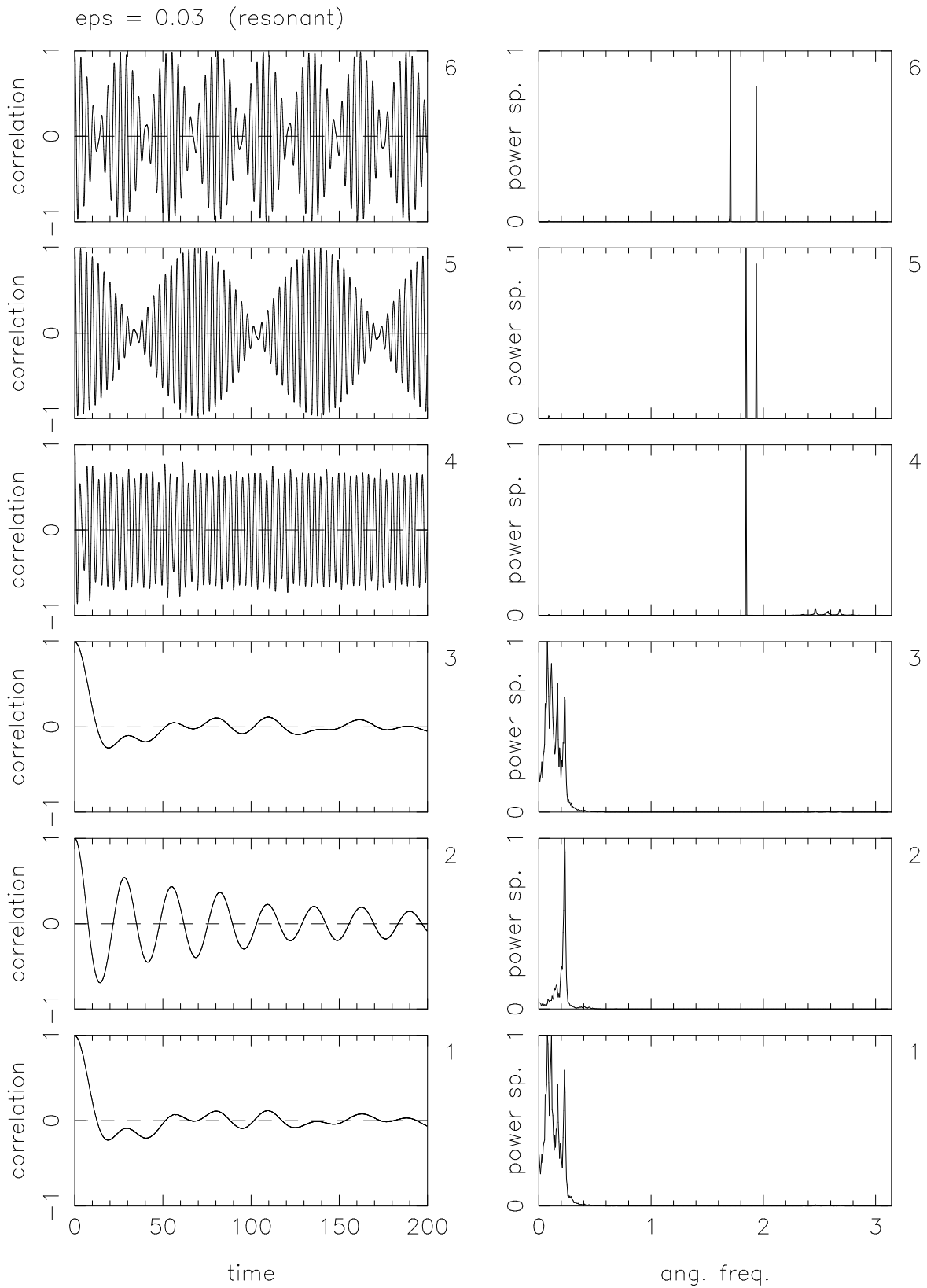
perciò si guadagna un ordine in  $\varepsilon$  se  $k \cdot \omega(I) = \mathcal{O}(1)$  per  $0 < |k| \leq K$ , mentre se  $k \cdot \omega(I) = \mathcal{O}(\sqrt{\varepsilon})$  la nuova perturbazione è ancora di ordine  $\varepsilon$ , e non si è guadagnato niente.

Consideriamo ora una generica risonanza di molteplicità  $r \geq 1$ : precisamente prendiamo un reticolo  $r$ -dimensionale  $\mathcal{R} \in \mathbb{Z}^n$ , generato da  $r$  vettori indipendenti  $k^{(1)}, \dots, k^{(r)}$  con  $|k^{(s)}| \leq K$ , e consideriamo la sottovarietà<sup>58)</sup>  $(n - r)$ -dimensionale di  $\mathcal{B}$

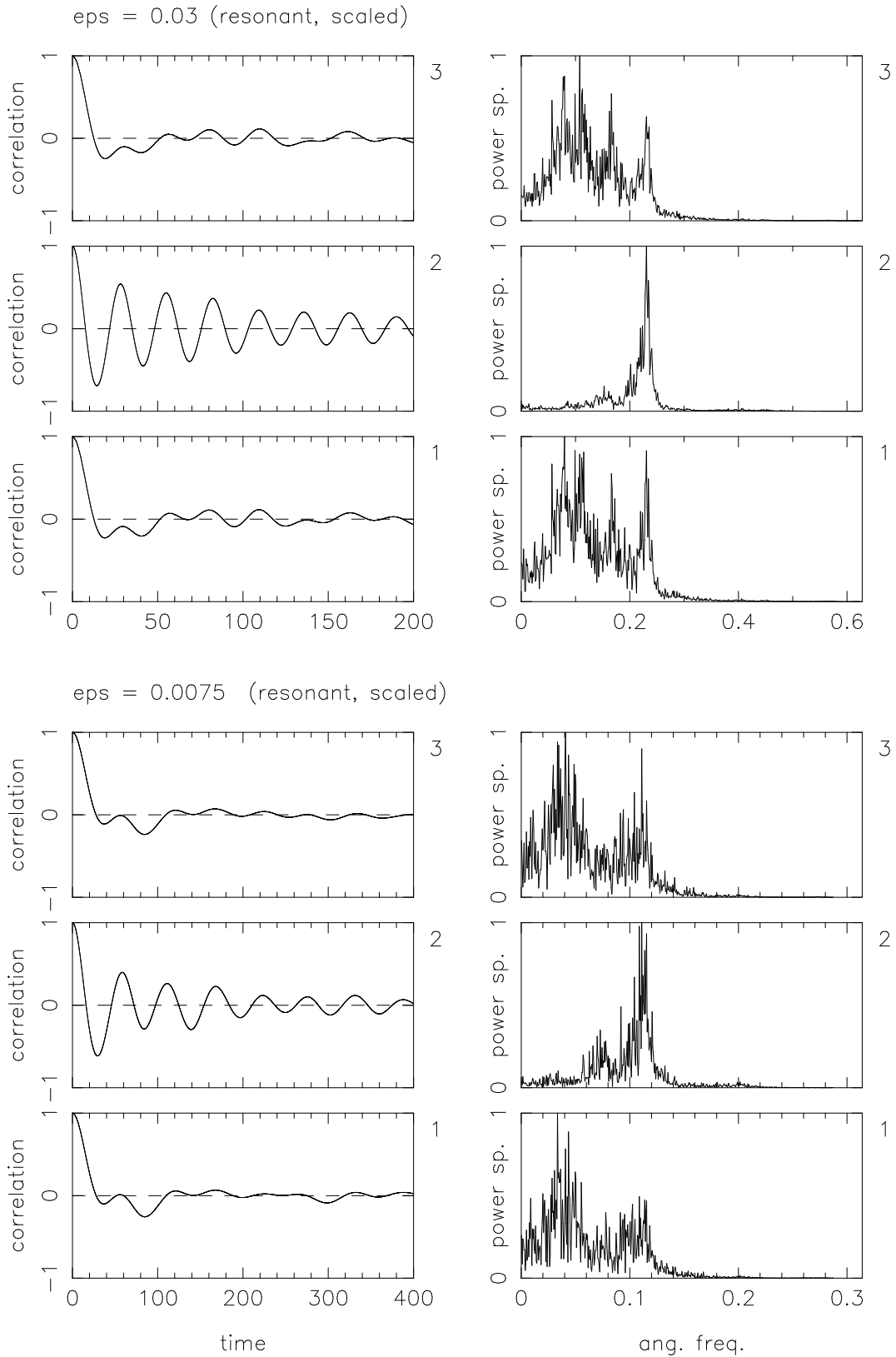
$$\Sigma_{\mathcal{R}} = \{ I \in \mathcal{B} : k^{(s)} \cdot \omega(I) = 0, \quad s = 1, \dots, r \},$$

<sup>58)</sup> E' una sottovarietà grazie al fatto che  $\det h'' \neq 0$ .





**Figura 7.8.** Simile alla figura 7.6, per  $\varepsilon = 0.03$  e dato iniziale nella risonanza  $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\varepsilon}$ .



**Figura 7.9.** Funzioni di correlazione e spettri di potenza di  $I_1, I_2, I_3$  nella risonanza  $I_1 \simeq I_2 \simeq I_3 \simeq \sqrt{\epsilon}$ , per  $\epsilon = 0.03$  (sopra) e  $\epsilon = 0.0075$  (sotto).

detta *varietà risonante*. In prossimità di  $\Sigma_{\mathcal{R}}$ , e lontano da altre risonanze, si può ancora eseguire un passo perturbativo, con

$$\chi(I, \varphi) = \sum_{k \notin \mathcal{R}} \frac{\hat{f}_k(I)}{ik \cdot \omega} e^{ik \cdot \varphi} ;$$

la nuova hamiltoniana è allora della forma

$$\tilde{H}(I, \varphi) = h(I) + \varepsilon g(I, \varphi) + \mathcal{O}(\varepsilon^2) , \quad g(I, \varphi) = \sum_{k \in \mathcal{R}} \hat{f}_k(I) e^{ik \cdot \varphi} . \quad (7.12)$$

Trascuriamo il termine  $\mathcal{O}(\varepsilon^2)$ , che non muove significativamente le azioni se non su tempi  $\mathcal{O}(\varepsilon^{-2})$ , e guardiamo all'effetto sulle azioni del termine  $g$ . Con evidenza,  $\dot{I}$  è parallelo a  $\mathcal{R}$ :

$$\dot{I} = -\varepsilon \sum_{k \in \mathcal{R}} (i \hat{f}_k e^{ik \cdot \varphi}) k ,$$

pertanto  $I(t)$  è confinato al piano  $\Pi_{\mathcal{R}}$  parallelo a  $\mathcal{R}$ , passante per  $I(0)$ . Lo spostamento delle azioni dovuto a  $g$  è detto *spostamento rapido* (inglese *fast drift*), contrapposto al più lento moto trasverso dovuto al resto,<sup>59)</sup> detto (impropriamente) *diffusione*.

Sfruttiamo ora l'ipotesi di convessità, e lo facciamo in due modi, osservando che:

- i.  $\Sigma_{\mathcal{R}}$  e  $\Pi_{\mathcal{R}}$  sono trasversi. La perdita di trasversalità infatti implicherebbe l'esistenza di  $v \in \Pi_{\mathcal{R}}$  che sia tangente a  $\Sigma_{\mathcal{R}}$ , e dunque ortogonale al piano  $\Lambda_{\mathcal{R}}$  ortogonale a  $\Sigma_{\mathcal{R}}$ . Questo piano è generato dai gradienti delle equazioni che definiscono  $\Sigma_{\mathcal{R}}$ , cioè dagli  $r$  vettori

$$\lambda^{(s)} = \nabla((k^{(s)} \cdot \omega(I)) = h''(I) k^{(s)} , \quad s = 1, \dots, r .$$

La perdita di trasversalità si ha perciò se esiste  $v \in \Pi_{\mathcal{R}}$  tale che  $\lambda^{(s)} \cdot v = 0$  per ogni  $s$ , ovvero tale che

$$h'' u \cdot v = 0 \quad \forall u \in \mathcal{R} ,$$

in particolare  $h'' v \cdot v = 0$ : il che però è impedito dalla convessità.  $\Sigma_{\mathcal{R}}$  e  $\Pi_{\mathcal{R}}$  hanno dimensione complementare (rispettivamente  $n-r$  e  $r$ ) nello spazio delle azioni, dunque si intersecano in un punto, che denotiamo  $I^*$ .

- ii. L'hamiltoniana imperturbata  $h$ , ristretta a  $\Pi_{\mathcal{R}}$ , ha un minimo in  $I^*$ . Questo è evidente se si sviluppa  $h$  attorno a  $I^*$ ,

$$h(I) = h(I^*) + \omega(I^*) \cdot (I - I^*) + \frac{1}{2} h''(I^*) (I - I^*) \cdot (I - I^*) + \dots ,$$

e si osserva che il termine lineare si annulla per  $I - I^*$  parallelo a  $\mathcal{R}$ , cioè per  $I \in \Pi_{\mathcal{R}}$ .

- **Osservazione.** Si vede facilmente che la vera proprietà che stiamo usando non è la convessità di  $h$ , ma una proprietà poco più debole detta *quasi convessità*: una funzione  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  si dice  $(m, l)$ -quasi convessa in  $\mathcal{B} \in \mathbb{R}^n$ , se  $h'' v \cdot v \geq m v \cdot v$  per ogni  $v$

---

<sup>59)</sup> Più lento soprattutto se invece di fare un solo passo perturbativo se ne fanno molti, e il resto diventa una potenza elevata di  $\varepsilon$ , o addirittura diviene esponenzialmente piccolo.

tale che  $|\omega \cdot v| \leq l$ . E' la convessità ristretta al piano ortogonale a  $\omega = \nabla h$ , cioè convessità ristretta alle superfici di livello di  $h$ . E' quasi convessa  $h(I_1, I_2) = \frac{1}{2}I_1^2 + I_2$ , per  $I_1 \neq 0$ ; oppure anche  $h(I_1, I_2) = \frac{1}{2}(I_1^2 - \eta I_2^2)$ , per  $0 < \eta < 1$ , nel dominio  $I_1 > 0$ ,  $|I_2| < I_1$  (qusst'ultimo è il caso del corpo rigido di Eulero, scritto in coordinate di azione–angolo).

Mettiamoci nel piano di moto rapido  $\Pi_{\mathcal{R}}$ , e poniamo per semplicità  $h(I^*) = 0$ ;  $h$  ha un minimo in  $I^*$ , perciò le superfici di livello di  $h$ , in vicinanza di  $I^*$ , sono sfere concentriche. Ma, grazie al fatto che  $h$  differisce da  $\tilde{H}$  per termini  $\mathcal{O}(\varepsilon)$ , si ha che  $h$  stessa ha oscillazioni, nel tempo, di ordine  $\varepsilon$ . Perciò il moto con, ad esempio,  $I(0)$  in un intorno  $\mathcal{O}(\sqrt{\varepsilon})$  dalla risonanza, ha inizialmente energia  $\mathcal{O}(\varepsilon)$ , e dunque per ogni  $t$  mantiene  $h = \mathcal{O}(\varepsilon)$ : pertanto per ogni  $t$  si ha  $\|I(t) - I^*\| = \mathcal{O}(\sqrt{\varepsilon})$ . Nella sostanza: l'ipotesi di convessità fa sì che il moto, ove si trascuri la diffusione dovuta ai termini  $\mathcal{O}(\varepsilon^2)$ , sia confinato a intorni di ampiezza  $\mathcal{O}(\sqrt{\varepsilon})$  attorno alle risonanze, ed è immediato convincersi che la scala di tempo di tali moti, diciamo il tempo di attraversamento di questi intorni per effetto del termine  $\varepsilon g$ , è  $\mathcal{O}(\sqrt{\varepsilon})$ . *Il quadro è sostanzialmente identico al caso dei rotatori.* In conclusione, per hamiltoniane del tipo (7.1), con  $h$  convessa e  $f$  polinomio di Fourier, *continua a valere il principio della media.*

- **Esercizio.** Si introducano variabili adattate alla risonanza, cioè angoli lenti e veloci (e azioni corrispondenti); si osservi che effettuando riscaldamenti nelle azioni e nei tempi si ottiene, come nel caso dei rotatori, un sistema sempre fortemente accoppiato per ogni  $\varepsilon$ .
- **Esercizio.** Si ripercorra il procedimento euristico sopra descritto, e si cerchi di rigorizzarlo tenendo conto della perturbazione  $\varepsilon^2 \tilde{f}$ , per tempi  $|t| = \mathcal{O}(\varepsilon^{-1})$ .

### 7.3. Perturbazioni con infinite componenti di Fourier.

Consideriamo infine il caso di una hamiltoniana della forma (7.1), con  $h$  convessa, ma senza più ipotesi su  $f$ , se non che sia funzione regolare di  $\varphi$ ; benché per fare un singolo passo perturbativo basti la differenziabilità finita, qui supporremo che  $f$ , come funzione di  $\varphi$ , sia analitica in una striscia  $\mathcal{S}_\rho$ . Entrambe le ipotesi i. e ii. del teorema di Poincaré 4.1 sono allora soddisfatte. Come si vedrà, questo non preclude che si possa ugualmente lavorare su  $H$  in modo perturbativo, mostrando in particolare la validità del principio della media (pur senza contraddire, ovviamente, il teorema).

L'idea, molto semplice, è di introdurre per  $f$  la decomposizione

$$f = f^{\leq K} + f^{> K}, \quad f^{\leq K} := \sum_{|k| \leq K} \hat{f}_k(I) e^{ik \cdot \varphi},$$

e di sfruttare il fatto (Lemma 5.2) che per una funzione analitica degli angoli le armoniche di Fourier decrescono esponenzialmente con  $|k|$ ;  $f^{> K}$  è detta *parte ultravioletta* di  $f$ , mentre  $K$  è detto *taglio (cut-off) ultravioletto*.

**Lemma 7.1.** *Se  $f$ , funzione degli angoli, è analitica nella striscia  $\mathcal{S}_\rho$ , allora esiste una costante  $C$  (dipendente da  $n$  e da  $\rho$ ) tale che*

$$\|f^{> K}\|_{\frac{1}{2}\rho} \leq C \|f\|_\rho e^{-\frac{1}{4}\rho K}.$$

**Dimostrazione.** Sappiamo che, per l'analiticità di  $f$  in  $\mathcal{S}_\rho$ ,

$$|\hat{f}_k| \leq \|f\|_\rho e^{-\rho|k|} ;$$

segue immediatamente

$$\|f^{>K}\|_{\frac{1}{2}\rho} \leq \|f\|_\rho \sum_{|k|>K} e^{-\frac{1}{2}\rho|k|} = \|f\|_\rho \sum_{s=K+1}^{\infty} N(n, s) e^{-\frac{1}{2}\rho s} ,$$

ove  $N_s < s^{n-1}$  (appendice D) denota il numero di  $n$ -ple intere con  $|k| = s$ . Segue allora

$$\|f^{>K}\|_{\frac{1}{2}\rho} \leq \|f\|_\rho \sum_{s=K+1}^{\infty} (s^{n-1} e^{-\frac{1}{4}\rho s}) e^{-\frac{1}{4}\rho s} \leq \|f\|_\rho C' \sum_{s=K+1}^{\infty} e^{-\frac{1}{4}\rho s} ,$$

ove  $C'$  è il massimo su  $s$  di  $s^{n-1} e^{-\frac{1}{4}\rho s}$ , e d'altra parte

$$\sum_{s=K+1}^{\infty} e^{-\frac{1}{4}\rho s} \leq \frac{4}{\rho} e^{-\frac{1}{4}\rho K} .$$

■

Questo lemma si può sfruttare in vario modo. Ad esempio, se poniamo  $K = K(\varepsilon) = \frac{4l}{\rho} \log \varepsilon^{-1}$ , otteniamo

$$\|f^{>K}\|_{\frac{1}{2}\rho} \leq C \|f\|_\rho \varepsilon^l ;$$

se invece prendiamo  $K = \varepsilon^{-b}$ ,  $b > 0$ , la parte ultravioletta di  $f$  diviene esponenzialmente piccola:

$$\|f^{>K}\|_{\frac{1}{2}\rho} \leq C \|f\|_\rho e^{-\frac{\rho}{4}\varepsilon^{-b}} .$$

Si vede allora che la parte ultravioletta della perturbazione, benchè non contenga a fattore una potenza elevata di  $\varepsilon$ , tuttavia muove le azioni solo su tempi  $\mathcal{O}(\varepsilon^l)$ , o addirittura  $\mathcal{O}(e^{\frac{\rho}{4}\varepsilon^{-b}})$ , e può essere ignorata per tempi inferiori.

La prescrizione per eseguire, secondo il nostro programma, un passo perturbativo, è allora la seguente:

- i. Si fissa il taglio ultravioletto  $K(\varepsilon)$ ; per eseguire un singolo passo basta  $K = \frac{4}{\rho} \log \varepsilon^{-1}$ . Si prendono in considerazione solo le armoniche con  $|k| \leq K$ , che sono in numero finito.
- ii. Si considerano le varietà risonanti  $\Sigma_{\mathcal{R}}$  per ogni reticolo risonante  $\mathcal{R}$ ,  $\dim \mathcal{R} = r \leq n-1$ , generato da vettori  $k^{(s)}$  con  $|k^{(s)}| \leq K$ . Queste vanno infittendosi al tendere di  $\varepsilon$  a zero e di  $K$  all'infinito: ma per ogni  $\varepsilon$  sono in numero finito.
- iii. Si lavora perturbativamente fuori dalle risonanze, e anche attorno a ogni risonanza, come illustrato nel paragrafo precedente. C'è una difficoltà in più, legata al fatto che le varietà risonanti si infittiscono: bisogna fare attenzione che partendo da una regione

ove si è costruita e si sta usando l'hamiltoniana (7.12), nel corso del moto e per la scala di tempo che interessa, non intervengano altre risonanze, che renderebbero invalido l'uso di  $\tilde{H}$ . E' un punto delicato, sul quale in questa trattazione euristica sorvoliamo.

Il moto acquista proprietà diverse in zone diverse dello spazio delle azioni, e può accadere ad esempio che al calare di  $\varepsilon$  una regione non risonante improvvisamente diventi risonante (e compaiano moti caotici). Ma *il principio della media resta valido*: su scala di tempo  $\mathcal{O}(\varepsilon^{-1})$  le variazioni delle azioni sono piccole con  $\varepsilon$ .

- **Osservazione.** La compatibilità col teorema di Poincarè 4.1 è piena, perché resta violato il terzo punto: in nessuna regione il resto ha a fattore  $\varepsilon^2$ ; i termini ultravioletti mantengono a fattore  $\varepsilon$ , e sono piccoli per altra ragione. Inoltre l'infittirsi delle risonanze per  $\varepsilon \rightarrow 0$  fa sì che in nessun aperto si possano eliminare gli angoli, per  $\varepsilon$  in un intorno dell'origine.

## 8. Il teorema di Nekhoroshev e il teorema KAM

## 9. Un sistema degenere: il corpo rigido in rapida rotazione

## 10. Applicazioni alla Meccanica Celeste

## 11. Applicazioni alla meccanica Statistica

## APPENDICI

## A. Prova della proposizione 1.4

Osserviamo innanzitutto che non è restrittivo supporre  $s$  e  $t$  piccoli, tanto da poter lavorare localmente in una carta: infatti per il primo punto la cosa è ovvia, mentre per il secondo basta osservare che si può sempre scrivere  $s = k\sigma$  e  $t = k\tau$ , con  $k$  grande, e come è immediato verificare, la commutazione di  $\Phi_X^\sigma$  e  $\Phi_Y^\tau$  implica quella di  $\Phi_X^s$  e  $\Phi_Y^t$ .

La dimostrazione del primo punto è un semplice calcolo: posto

$$\Delta_f(s, t) = f \circ \Phi_Y^t \circ \Phi_X^s - f \circ \Phi_X^s \circ \Phi_Y^t ,$$

si ha  $\Delta_f(0, t) = \Delta_f(s, 0) = 0$ , e dunque

$$\Delta_f(s, t) = st \frac{\partial^2 \Delta_f}{\partial s \partial t} \Big|_{s=t=0} + \mathcal{O}_3(s, t) .$$

Si calcola poi facilmente

$$\frac{\partial f}{\partial t}(\Phi_Y^t(\Phi_X^s(x))) \Big|_{t=0} = (L_Y f)(\Phi_X^s(x)) , \quad \frac{\partial^2 f}{\partial s \partial t}(\Phi_Y^t(\Phi_X^s(x))) \Big|_{s=t=0} = (L_X L_Y f)(x) ,$$

e analogamente  $\frac{\partial^2 f}{\partial s \partial t}(\Phi_X^s(\Phi_Y^t(x))) \Big|_{s=t=0} = (L_Y L_X f)(x)$ . Segue, come richiesto,

$$\frac{\partial^2 \Delta_f}{\partial s \partial t} \Big|_{s=t=0} = [L_X, L_Y] f = L_{[X, Y]} f .$$

Per quanto riguarda il secondo punto, in un verso la prova è banale (se  $\Delta_f = 0$  per ogni  $f$  certamente  $[X, Y] = 0$ ), 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

$$\Phi_Y^t \circ \Phi_X^s = \Phi_Y^\tau \circ \dots \circ \Phi_Y^\tau \circ \Phi_X^\sigma \circ \dots \circ \Phi_X^\sigma .$$

Per quanto già dimostrato, si ha

$$\Phi_Y^\tau \circ \Phi_X^\sigma = \Phi_X^\sigma \circ \Phi_Y^t + \mathcal{O}(k^{-3}) ,$$

e se  $\lambda$  denota la costante di Lipshitz di  $Y$  segue

$$\Phi_Y^{(k-1)\tau} \circ \Phi_Y^\tau \circ \Phi_X^\sigma \circ \Phi_X^{(k-1)\sigma} = \Phi_Y^{(k-1)\tau} \circ \Phi_X^\sigma \circ \Phi_Y^\tau \circ \Phi_X^{(k-1)\sigma} + \mathcal{O}(e^{\lambda(k-1)\sigma} k^{-3}) .$$

Iterando, con un numero di scambi pari a  $k$  si porta  $\Phi_X^\sigma$  in testa, e con  $\mathcal{O}(k^2)$  scambi si portano a uno a uno in testa tutti i  $\Phi_X^\sigma$ . L'errore ogni volta è di ordine  $k^{-3}$  per un coefficiente che non supera

$$c := e^{\lambda(l\sigma + (k-l-1)\tau)} \leq e^{\lambda \max(s, t)} ,$$

ove si è supposto che  $\lambda$  sia costante di Lipshitz anche per  $X$  ( $l$  è il numero di fattori  $\Phi_Y^\tau$  rimasti a sinistra,  $k-l-1$  il numero di fattori  $\Phi_X^\sigma$  già portati in testa). Segue

$$\Phi_Y^t \circ \Phi_X^s - \Phi_X^s \circ \Phi_Y^t = \mathcal{O}(k^2 k^{-3}) = \mathcal{O}(k^{-1}) ,$$

e per l'arbitrarietà di  $k$  si conclude che i flussi commutano.

## B. Lemmi per il teorema di Liouville–Arnol’d

### B.1. Dimostrazione del lemma 2.3.

Sia  $k \leq n$  la dimensione di  $\mathcal{T}$ , ovvero il massimo numero di vettori linearmente indipendenti di  $\mathcal{T}$ ; allora ogni elemento  $\tau$  di  $\mathcal{T}$  si può scrivere come combinazione lineare

$$\tau = \sum_{j=1}^k \mu_j e^j, \tag{B.1}$$

con  $e^1, \dots, e^k$  vettori indipendenti di  $\mathcal{T}$ . Dobbiamo mostrare che si possono sempre scegliere  $e^1, \dots, e^k$  in modo che tutti gli elementi di  $\mathcal{T}$  siano combinazioni intere di tali vettori (l’affermazione inversa è banale: poiché  $\mathcal{T}$  è un gruppo, allora tutte le combinazioni intere dei suoi elementi, in particolare di  $e^1, \dots, e^k$ , sicuramente appartengono a  $\mathcal{T}$ ). Per questo denotiamo con  $E$  il sottospazio di  $\mathbb{R}^n$  generato da  $e^1, \dots, e^k$ ,  $\mathcal{T} \subset E$ , e con  $\mathcal{E}(e^1, \dots, e^k) \subset E$  il singolo parallelogramma, o cella, che ha un vertice nell’origine ed è generato da  $e^1, \dots, e^k$ . Il lemma si riduce a dimostrare che *per ogni sottogruppo discreto  $k$ -dimensionale  $\mathcal{T}$  di  $\mathbb{R}^n$  esiste sempre una cella “elementare”  $\mathcal{E}(e^1, \dots, e^k)$ , nella quale non cadono altri elementi di  $\mathcal{T}$  al di fuori dei vertici* (figura B.1). Ammessa infatti questa proprietà la conclusione è immediata: se  $e^1, \dots, e^k$  generano una cella elementare  $\mathcal{E}$ , e  $\tau$  della forma (B.1) appartiene a  $\mathcal{T}$ , allora i coefficienti  $\mu_1, \dots, \mu_k$  sono necessariamente interi, altrimenti

$$\tilde{\tau} = \sum_{j=1}^k (\mu_j - [\mu_j]) e^j, \quad [.] = \text{parte intera}$$

(che certamente appartiene a  $\mathcal{T}$ ) starebbe in  $\mathcal{E}$  senza essere uno dei suoi vertici.

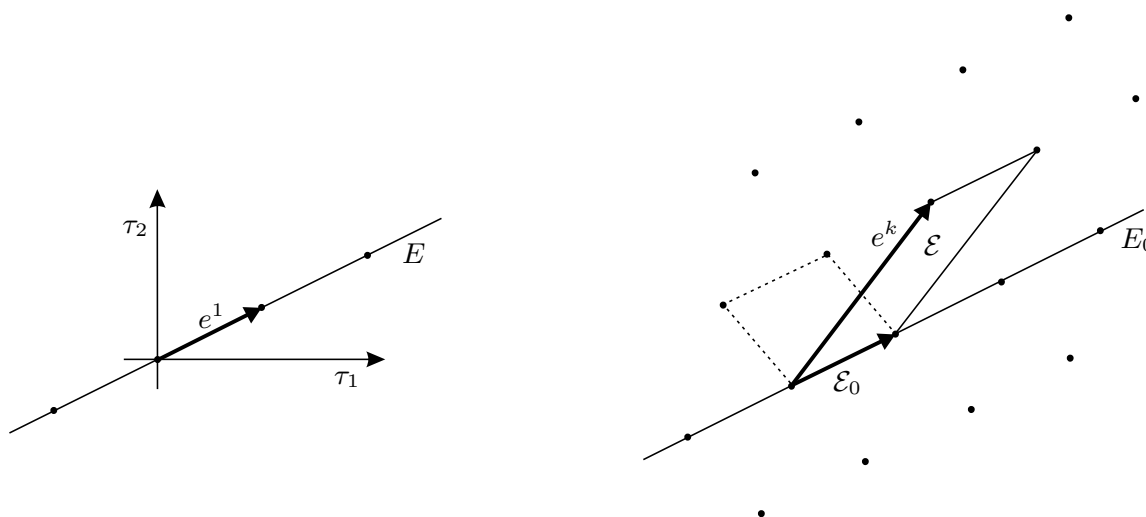


Figura B.1. Celle elementari in una e due dimensioni.



L'esistenza della cella elementare  $\mathcal{E}$  è evidente per  $k = 1$ : in tal caso  $E$  è una retta, e basta prendere per  $e^1$  uno dei due elementi di  $\mathcal{T}$  che hanno distanza minima dall'origine (esistono sempre, perché il gruppo è discreto). Procedendo per induzione, assumiamo vera la proprietà per  $k - 1 < n$ , e verifichiamola per  $k$ . Siano allora  $u^1, \dots, u^k$  vettori indipendenti qualsiasi di  $\mathcal{T}$ ; denotiamo con  $E_0$  il sottospazio di  $E$  generato dai primi  $k - 1$  vettori, e poniamo  $\mathcal{T}_0 = \mathcal{T} \cap E_0$ . Con evidenza  $\mathcal{T}_0$  è un sottogruppo discreto di  $R^n$  di dimensione  $k - 1$ , dunque per l'ipotesi induttiva esiste una sua cella elementare  $\mathcal{E}_0$ , generata da opportuni  $e^1, \dots, e^{k-1} \in \mathcal{T}_0$ . Mostriamo che si può sempre prendere  $e^k \in \mathcal{T} \setminus \mathcal{T}_0$ , tale che la cella  $\mathcal{E}(e^1, \dots, e^k)$  sia elementare.

Per questo osserviamo innanzitutto che al variare di  $\tau$  in  $\mathcal{T} \setminus \mathcal{T}_0$  il volume (euclideo  $k$ -dimensionale) della cella  $\mathcal{E}(e^1, \dots, e^{k-1}, \tau)$  è limitato inferiormente, o equivalentemente, la distanza di  $\tau$  da  $E_0$  è limitata inferiormente. Infatti, per assurdo, sia  $\tau^1, \tau^2, \dots$  una sequenza di elementi la cui distanza da  $E_0$  decresce monotonamente a zero, e sia

$$\sigma^i = \sum_{j=1}^{k-1} \mu_j e^j$$

il piede della perpendicolare da  $\tau^i$  a  $E_0$ ; allora tutti i punti

$$\tilde{\sigma}^i = \sigma^i - \sum_{j=1}^{k-1} [\mu_j] e^j$$

cadono in  $\mathcal{E}_0$ , e corrispondentemente tutti i punti

$$\tilde{\tau}^i = \tau^i - \sum_{j=1}^{k-1} [\mu_j] e^j,$$

che sicuramente appartengono a  $\mathcal{T}$ , sono diversi tra loro (hanno infatti diversa distanza da  $E_0$ ) e hanno distanza limitata dall'origine: di conseguenza hanno un'accumulazione, contro l'ipotesi che il gruppo  $\mathcal{T}$  sia discreto. Si vede allora facilmente che il volume delle celle è una variabile discreta: se infatti ci fossero  $\tau$  e  $\tau'$ , tali che le celle  $\mathcal{E}(e^1, \dots, e^{k-1}, \tau)$  e  $\mathcal{E}(e^1, \dots, e^{k-1}, \tau')$  hanno volume arbitrariamente vicino, allora la cella  $\mathcal{E}(e^1, \dots, e^{k-1}, \tau' - \tau)$  avrebbe volume arbitrariamente piccolo, e questo è già stato escluso. Dunque esiste (almeno) un elemento  $e^k \in \mathcal{T} \setminus \mathcal{T}_0$ , tale che la cella  $\mathcal{E}(e^1, \dots, e^{k-1}, e^k)$  ha volume minimo. In tale cella non possono cadere punti di  $\mathcal{T}$  diversi dai vertici: non nella base che sta in  $E_0$ , e di conseguenza non nella base opposta; non altrove, altrimenti esisterebbe una cella di volume inferiore. Questo basta per la conclusione del lemma. ■

## B.2. Una dimostrazione del lemma 2.4.

In questa dimostrazione useremo le coordinate adattate definite nel successivo lemma 2.5, che è indipendente dal lemma 2.4. La cosa da verificare è che per ogni  $i, j$  si ha

$$\frac{\partial P_i}{\partial q_j} - \frac{\partial P_j}{\partial q_i} = 0, \quad (B.2)$$

o equivalentemente che per ogni  $l, m$  si ha

$$\alpha_{l,m} := \sum_{i,j=1}^n \frac{\partial F_l}{\partial p_i} \frac{\partial F_m}{\partial p_j} \left( \frac{\partial P_i}{\partial q_j} - \frac{\partial P_j}{\partial q_i} \right) = 0 . \quad (B.3)$$

A questo scopo deriviamo la (2.13) rispetto alle  $q$ ; otteniamo

$$\sum_{i=1}^n \frac{\partial F_l}{\partial p_i} \frac{\partial P_i}{\partial q_j} + \frac{\partial F_l}{\partial q_j} = 0 , \quad \sum_{j=1}^n \frac{\partial F_m}{\partial p_j} \frac{\partial P_j}{\partial q_i} + \frac{\partial F_m}{\partial q_i} = 0 ,$$

che inserite nella (B.3) danno, come richiesto

$$\alpha_{l,m} = \sum_{i=1}^n \left( \frac{\partial F_m}{\partial q_i} \frac{\partial F_l}{\partial p_i} - \frac{\partial F_m}{\partial p_i} \frac{\partial F_l}{\partial q_i} \right) = 0 .$$

■

Questa dimostrazione ha il difetto di fare riferimento alle variabili adattate, ovvero alla condizione (2.12), garantita dal lemma 2.5 ma di per sé spuria.

### B.3. Dimostrazione del lemma 2.5

Supponiamo che, nelle variabili originali  $(p, q)$ , risulti

$$\text{rango} \left( \frac{\partial F_i}{\partial p_j} \right) = k < n . \quad (B.4)$$

Allora con opportuna trasformazione lineare nelle  $p$ , estendibile canonicamente alle  $q$ , si ottiene (nelle nuove variabili)

$$\left( \frac{\partial F_i}{\partial(p_j, q_j)} \right) = \left( \frac{\partial F_i}{\partial(p_1, \dots, p_k)} \mid 0 \mid \frac{\partial F_i}{\partial(q_1, \dots, q_k)} \mid \frac{\partial F_i}{\partial(q_{k+1}, \dots, q_n)} \right) ;$$

Mostriamo allora che la matrice quadrata  $Z$  di ordine  $n$  formata dalle prime  $k$  e dalle ultime  $n - k$  colonne (una colonna dunque per ogni grado di libertà  $j$ , relativa o alla coordinata  $p_j$  oppure alla coordinata coniugata  $q_j$ ) ha determinante diverso da zero; è chiaro che questo basta per la dimostrazione del lemma, perchè con trasformazione canonica banale si possono poi scambiare le  $p$  con le  $q$  limitatamente agli ultimi  $n - k$  gradi di libertà. Nella verifica possiamo evidentemente sostituire a piacere  $F_1, \dots, F_n$  con loro opportune combinazioni lineari indipendenti (ciò corrisponde a sostituire alle righe di  $Z$  corrispondenti combinazioni lineari): il determinante non cambia, inoltre le nuove  $F$  restano in involuzione. Approfittiamo di questa possibilità per far sì che le ultime  $n - k$  righe della matrice  $k \times n$  formata dalle prime  $k$  colonne siano nulle, ottenendo così la matrice a blocchi

$$\left( \frac{\partial F_i}{\partial(p_j, q_j)} \right) = \left( \begin{array}{c|c|c|c} A & 0 & B & C \\ \hline 0 & 0 & D & E \end{array} \right) ;$$

con le stesse notazioni si ha

$$Z = \left( \begin{array}{c|c} A & C \\ \hline 0 & E \end{array} \right) .$$

Prendiamo ora  $1 \leq i \leq k$ ,  $k+1 \leq j \leq n$ . E' immediato verificare che risulta

$$0 = \{F_i, F_j\} = - \sum_{l=1}^k \frac{\partial F_i}{\partial p_l} \frac{\partial F_j}{\partial q_l} = - \sum_{l=1}^k A_{il} D_{jl} ,$$

ovvero ciascuna riga di  $D$  è ortogonale a tutte le righe di  $A$ . Ma per la (B.4) si ha  $\det A \neq 0$ : se ne deduce che la matrice  $D$  è nulla. D'altra parte, si è supposto  $\text{rango} \left( \frac{\partial F_i}{\partial (p_j, q_j)} \right) = n$ , perciò deve essere necessariamente  $\det Z \neq 0$ . Questo basta alla conclusione del lemma. ■

### C. Le variabili di azione-angolo per il moto centrale

L'hamiltoniana (per unità di massa) del moto centrale piano si scrive, in coordinate polari  $(r, \theta)$ ,

$$H(p_r, p_\theta, r, \theta) = \frac{p_r^2}{2} + \frac{p_\theta^2}{2r^2} + V(r) .$$

Un'attenzione particolare andrà riservata al caso kepleriano,

$$V(r) = -\frac{k}{r} , \quad k > 0 .$$

In questo sistema sono costanti del moto  $H$  stessa e il momento angolare, cioè la funzione

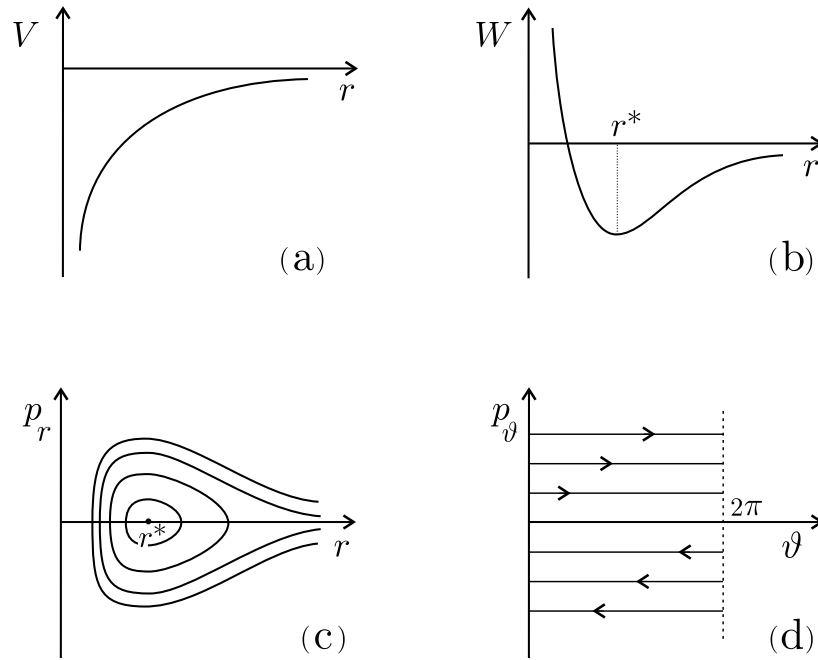
$$L(p_r, p_\theta, r, \theta) = p_\theta ,$$

ed evidentemente si ha  $\{H, L\} = 0$ . Per ogni valore  $l$  di  $L$  la coppia di variabili canoniche  $(p_r, r)$  segue l'hamiltoniana ridotta

$$H_l(p_r, r) = \frac{p_r^2}{2m} + W_l(r) , \quad W_l(r) = \frac{l^2}{2mr^2} + V(r) .$$

Supporremo che  $V$  sia attrattivo, diciamo della forma tracciata in figura C.1a, e che corrispondentemente,<sup>60)</sup> per  $l \neq 0$ , il "potenziale efficace"  $W_l(r)$  abbia la forma qualitativa riportata in figura C.1b. Denoteremo con  $r^*$  il punto di minimo di  $W_l$  (il raggio dell'orbita circolare di momento angolare  $l$ ) e con  $W_l^*$  il corrispondente valore di  $W_l$ . Le

<sup>60)</sup> Per questo bisogna che (come nel caso kepleriano)  $V(r)$  diverga più lentamente di  $r^{-2}$  per  $r \rightarrow 0$  e si annulli più lentamente di  $r^{-2}$  per  $r \rightarrow \infty$ .



**Figura C.1.** (a) Un potenziale attrattivo di tipo kepleriano; (b) il corrispondente potenziale efficace; (c) il ritratto in fase nel piano  $r, p_r$ ; (d) il ritratto in fase nel piano  $\theta, p_\theta$ .

figure C.1c,d riportano il ritratto in fase del sistema nel piano  $(r, p_r)$  e nel piano  $(\theta, p_\theta)$ ; si osservi che  $p_\theta$  resta costante mentre  $\theta$  avanza monotonamente (con velocità  $\dot{\theta} = l/(2r^2)$ ).

L'applicazione del teorema di Liouville–Arnol'd richiede innanzitutto di esaminare la matrice

$$\frac{\partial(H, L)}{\partial(p_r, p_\theta, r, \theta)} = \begin{pmatrix} \frac{\partial H}{\partial p_r} & \frac{\partial H}{\partial p_\theta} & \frac{\partial H}{\partial r} & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{C.1}$$

Come si è visto, la condizione per applicare il teorema nell'intorno di una superficie di livello  $H = h, L = l$  è che ovunque su di essa il rango della matrice sia due. Ma dalla (C.1) si vede che si ha una caduta di rango quando si annullano simultaneamente  $p_r$  e  $\frac{\partial H}{\partial r}$ , ovvero in corrispondenza agli equilibri del sistema ridotto  $(p_r, r) = (0, r^*(l))$  e  $(p_r, r) = (0, \infty)$ . Gli insiemi di livello “eccezionali” sono dunque costituiti dai moti circolari e dai moti che avvengono con velocità di fuga; i valori eccezionali di  $h$  e  $l$  sono dati, rispettivamente, dalle equazioni  $h = W_l^*$  e  $h = 0$ .

Al di fuori di queste eccezioni il teorema si applica. In effetti è evidente dalle figure C.1c,d che nel caso legato, cioè per

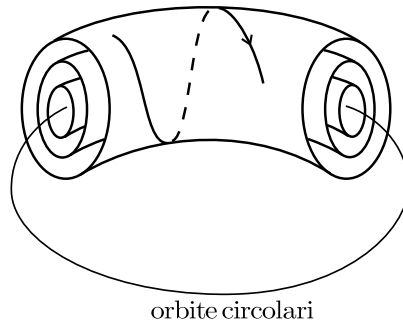
$$W^*(l) < h < 0 ,$$

la sottovarietà definita da

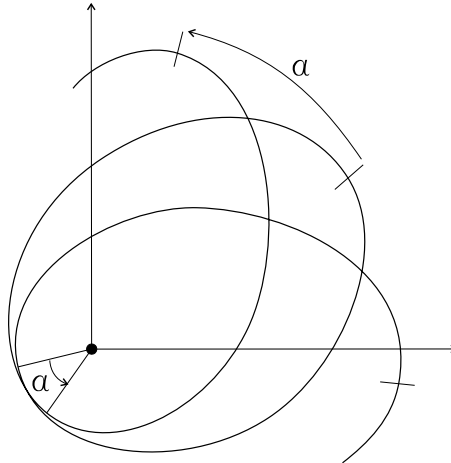
$$H(p_r, p_\theta, r, \theta) = h , \quad L(p_r, p_\theta, r, \theta) = l$$

è il prodotto di due cerchi, ovvero è un toro  $\mathbb{T}^2$ , mentre invece è un cilindro  $\mathbb{R} \times S^1$  per  $h > 0$ . Si osservi (figura C.1c) che per ogni  $l$ , al tendere di  $h$  al valore minimo  $W_l^*$ , una famiglia

di curve nel piano  $r, p_r$  si stringe in un punto; corrispondentemente nello spazio delle fasi del sistema una famiglia di tori  $\mathbb{T}^2$  si stringe su un circolo  $S^1$ , corrispondente all'orbita circolare di momento angolare  $l$  (figura C.2).<sup>61)</sup> Alla caduta di rango della matrice (C.1) si accompagna, come si vede, una singolarità della foliazione in tori. Ci restringiamo di qui in poi al caso legato.



**Figura C.2.** Una famiglia di tori  $\mathbb{T}^2$  si stringe su un circolo  $S^1$ , corrispondente all'orbita circolare di momento angolare  $l$ .



**Figura C.3.** Il moto a rosetta; a ogni passaggio, perielio e afelio avanzano di un angolo  $\alpha$ .

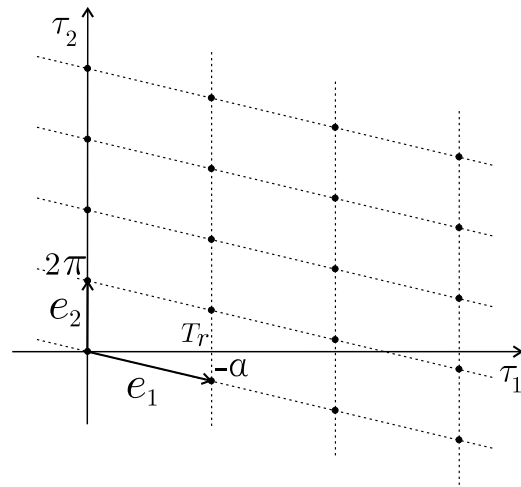
Per potenziali generici il moto, si sa, è “a rosetta” (figura C.3): nel periodo  $T_r(h, l)$  della coppia  $(p_r, r)$  l'angolo  $\theta$  avanza di

$$\Delta\theta(h, l) = \int_0^{T_r} \frac{l}{r^2(t)} \frac{d}{dt} ;$$

l'avanzamento del perielio (e dell'afelio) ad ogni passaggio è allora

$$\alpha(h, l) = \Delta\theta(h, l) - 2\pi .$$

<sup>61)</sup> Se facciamo variare anche  $l$ , abbiamo una famiglia a due parametri di tori  $\mathbb{T}^2$  che si stringe su una famiglia a un parametro di circoli, ciascuno dei quali è identificato da  $l$  o indifferentemente dal valore  $h$  dell'energia.



**Figura C.4.** Il gruppo dei periodi per il moto centrale piano.

Nel caso kepleriano, come è noto, si ha  $\alpha(h, l) = 0$ .

Per la costruzione delle variabili di azione–angolo servono due cicli indipendenti sul toro, per mezzo dei quali si potranno identificare il gruppo e la matrice dei periodi. Un primo ciclo è del tutto ovvio: si lascia agire  $L$ , il cui campo vettoriale hamiltoniano è  $X_L = (0, 0, 0, 1)$ , per il tempo  $\tau_2 = 2\pi$ , mentre non si fa agire  $H$ , ovvero si prende  $\tau_1 = 0$ . Un periodo di  $\Phi^\tau = \Phi_L^{\tau_2} \circ \Phi_H^{\tau_1}$  è dunque  $\tau = (0, 2\pi)$ . Un secondo ciclo indipendente si ottiene facendo agire  $H$  per il tempo  $\tau_1 = T_r(h, l)$ , in modo che  $(p_r, r)$  ritornino al valore iniziale, e poi  $L$  per il tempo  $-\alpha$ , in modo da portare  $\theta$  a  $2\pi$  lasciando fermi  $p_r$  e  $r$ . Un periodo indipendente dal precedente, questa volta dipendente da  $h$  e  $l$ , è allora

$$\tau(h, l) = (T_r(h, l), -\alpha(h, l)) .$$

Assoceremo ai due cicli, nell'ordine, gli indici 2 e 1. Il gruppo dei periodi è formato precisamente dalle combinazioni lineari intere<sup>62)</sup> di

$$e^{(1)} = (T_r, -\alpha) , \quad e^{(2)} = (0, 2\pi)$$

(si veda la figura C.4), mentre la matrice dei periodi  $T$  e la sua inversa sono

$$T = \begin{pmatrix} T_r & 0 \\ -\alpha & 2\pi \end{pmatrix} , \quad T^{-1} = \frac{1}{2\pi} T_r \begin{pmatrix} 2\pi & 0 \\ \alpha & T_r \end{pmatrix} .$$

Le due azioni sono

$$I_1(h, l) = \frac{1}{2\pi} \left[ \oint_{\text{ciclo}_1} p_r dr + p_\theta d\theta \right] = \frac{1}{2\pi} \oint_{\text{ciclo}_1} p_r dr + l = \frac{\mathcal{A}(h, l)}{2\pi} + l , \quad (C.2)$$

ove  $\mathcal{A}(h, l)$  denota l'area sottesa alla curva  $H_l(p_r, r) = h$  nel piano  $r, p_r$ , e

$$I_2(h, l) = \frac{1}{2\pi} \oint_{\text{ciclo}_2} p_\theta d\theta = l . \quad (C.3)$$

<sup>62)</sup> Non è difficile convincersi che non vi sono altri periodi che le combinazioni intere di  $e_1$  e  $e_2$ .

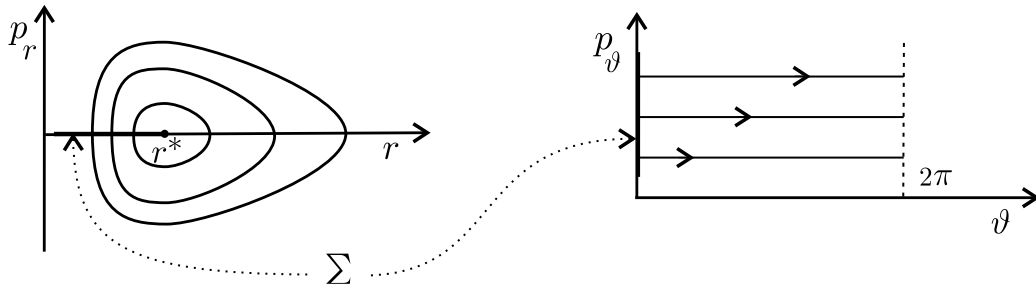
Per definire in maniera univoca gli angoli gli angoli va scelta una superficie  $\Sigma$  trasversa ai tori, che stabilisca un'origine su ciascuno di essi. Una scelta conveniente si ottiene, per questo sistema, fissando non le coordinate  $r$  e  $\theta$  (come si è fatto nella dimostrazione del teorema di Liouville–Arnol'd), ma il momento  $p_r$  e la coordinata  $\theta$ . E' una scelta altrettanto legittima, anzi è proprio la stessa se si pensa di effettuare preliminarmente la trasformazione canonica che scambia tra loro  $r$  e  $p_r$ . La sezione prescelta (figura C.5) è precisamente  $p_r = 0$ , nel tratto  $0 < r < r^*$ , e  $\theta = 0$ . I punti di  $\Sigma$  corrispondono così a stati nei quali il corpo orbitante è al perielio. Con questa scelta gli angoli sono dati da

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = 2\pi T^{-1} \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix},$$

ovvero

$$\varphi_1 = 2\pi \frac{\tau_1}{T_r}, \quad \varphi_2 = \alpha \frac{\tau_1}{T_r} + \tau_2, \quad (C.4)$$

con origine dei tempi su  $\Sigma$ . Si vede così che l'angolo  $\varphi_1$  ha origine nell'ultimo perielio, e avanza uniformemente fino a  $2\pi$  al momento di passaggio al perielio successivo;<sup>63)</sup> invece l'angolo  $\varphi_2$  ha origine sull'asse  $x$  (arbitrario) e avanza uniformemente di  $\alpha$  nel tempo intercorrente tra due passaggi successivi al perielio ( $\tau_2$  resta evidentemente costante nel moto vero, cioè per effetto di  $\Phi_H^t$ ). I due angoli sono chiamati rispettivamente *anomalia media* e *anomalia del perielio*.



**Figura C.5.** Una buona scelta della sezione  $\Sigma$  per il moto centrale piano.

L'espressione (C.2) di  $I_1$  permette di calcolare  $I_1$  stessa come funzione di  $h$  e  $l$ , e poi per inversione, ricordando  $I_2 = l$ , si ottiene  $h(I_1, I_2)$ , ovvero l'hamiltoniana espressa in funzione delle variabili di azione; l'espressione da calcolare è

$$I_1(h, l) = \frac{1}{\pi} \int_{r_-}^{r_+} p_r dr + l = \frac{1}{\pi} \int_{r_-}^{r_+} \sqrt{2 \left( h - \frac{l^2}{2r^2} + \frac{k}{r} \right)} dr + l, \quad (C.5)$$

con  $r_{\pm}$  zeri del radicando,

$$r_{\pm} = -\frac{k}{2h} \pm \frac{1}{2} \sqrt{\frac{k^2}{h^2} + \frac{2l^2}{h}}.$$

<sup>63)</sup> Per la legge delle aree,  $\varphi_1$  è dunque proporzionale all'area spazzata dal corpo orbitante a partire dal perielio.

Il risultato dipende evidentemente dal problema in questione. In ogni caso tuttavia le (C.4) mostrano che per la frequenza  $\omega = \frac{\partial h}{\partial I} = \dot{\varphi}$  si ha l'espressione

$$\omega = (\omega_1, \omega_2) = \frac{1}{T_r}(2\pi, \alpha) .$$

Nel caso kepleriano i calcoli si potrebbero eseguire in dettaglio, ma li semplifichiamo, accettando per buono dalla meccanica Newtoniana il fatto cruciale che in questo caso  $\alpha = 0$  (l'orbita si chiude, il perielio non avanza). Ma allora  $\omega_2 = \frac{\partial h}{\partial I_2} = 0$ , ovvero  $h(I_1, I_2)$  non dipende da  $I_2$ , e pertanto l'espressione (C.5) di  $I_1$  si può calcolare per un valore arbitrario di  $l$ , anche al limite  $l = 0$ . In questo limite si ha  $r_-(h) = 0$ ,  $r_+(h) = -k/h$ , e

$$I_1(h) = \frac{1}{\pi} \int_0^{-k/h} \sqrt{2(h + k/r)} \, dr .$$

A conti fatti viene

$$I_1(h) = \frac{1}{\sqrt{2}} \frac{k}{\sqrt{-h}} ,$$

cosicché

$$h(I_1) = -\frac{k^2}{2I_1^2} ;$$

corrispondentemente la frequenza  $\omega_1$  è

$$\omega_1 = \frac{k}{I_1^3} .$$

D'altra parte, l'ellisse kepleriana ha semiasse maggiore

$$a = \frac{1}{2}(r_+ + r_-) = -\frac{k}{h} = \frac{I^2}{k} ,$$

e dunque  $\omega_1$  è proporzionale a  $a^{-3/2}$ , in conformità alla terza legge di Keplero.

## D. Prova della proposizione 5.2

Poniamo

$$\bar{\Omega}_{\gamma, \tau, k} = \left\{ \omega \in \Omega : |k \cdot \omega| \leq \frac{\gamma}{\|k\|} \right\} ,$$

cosicché

$$\Omega \setminus \Omega_{\gamma, \tau} = \bigcup_{k \in \mathbb{Z}^n \setminus \{0\}} \bar{\Omega}_{\gamma, \tau, k} , \quad \text{mes}(\Omega \setminus \Omega_{\gamma, \tau}) \leq \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \text{mes}(\bar{\Omega}_{\gamma, \tau, k}) .$$



Si trova facilmente la stima

$$\text{mes}(\overline{\Omega}_{\gamma,\tau,k}) < \gamma \frac{\text{Vol}(\Omega)}{L} \frac{n\sqrt{n}}{|k|^{\tau+1}}. \quad (D.1)$$

Infatti, posto

$$\omega_{\parallel} := \frac{k \cdot \omega}{\|k\|},$$

si vede subito che la condizione diofantea è violata in corrispondenza a  $k$  se  $|\omega_{\parallel}| < \gamma(\|k\| |k|^{\tau})^{-1}$ ; corrispondentemente si ha

$$\text{mes}(\overline{\Omega}_{\gamma,\tau,k}) < \text{Vol}(B^{n-1}(L)) \frac{2\gamma}{\|k\| |k|^{\tau}},$$

ove  $B^{n-1}(L)$  denota la sfera  $n - 1$ -dimensionale di raggio  $L$ . Una stima facile è<sup>64)</sup>

$$\text{Vol}(B^{n-1}(L)) < \frac{n}{2L} \text{Vol}(B^n(L)) = \frac{n}{2L} \text{Vol}(\Omega),$$

da cui segue la (D.1). Si ottiene allora

$$\text{mes}(\overline{\Omega}_{\gamma,\tau}) < \gamma n \sqrt{n} \frac{\text{Vol}(\Omega)}{L} \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{1}{|k|^{\tau+1}} = \gamma n \sqrt{n} \frac{\text{Vol}(\Omega)}{L} \sum_{s \geq 1} N_s \frac{1}{s^{\tau+1}},$$

ove  $N_s$  denota il numero di  $n$ -ple intere  $k$  con  $|k| = s$ . Ora è facile verificare per induzione che risulta

$$N_s = \binom{s+n-1}{s} < s^{n-1}$$

(una stima immediata più rozza è  $N_s < (2s+1)^{n-1} < (3s)^{n-1}$ ); per  $\tau > n - 1$  la serie  $\sum_{s \geq 1} s^{n-\tau-2}$  converge, e segue la (5.13). Per  $\tau = n$  si stima facilmente  $\sum_{s \geq 1} s^{-2} < 2$ , ad esempio sovrastimando la sommatoria con  $1 + \int_1^{\infty} s^{-2} ds$ .

---

<sup>64)</sup> Si sottostima  $B^n(L)$  con una coppia di coni di base  $B^{n-1}(L)$  e altezza  $L$ . La stima esatta è

$$B^n(L) = C_n L \frac{(n-1)!!}{n!!} \text{Vol}(B^{n-1}(L)),$$

con  $C_n = 2$  per  $n$  dispari,  $C_n = \pi$  per  $n$  pari.