10

PERSISTENZA DEI TORI INVARIANTI

In questo capitolo darò la dimostrazione del celebre teorema di Kolmogorov sulla persistenza di tori invarianti in sistemi hamiltoniani quasi integrabili.

Kolmogorov annunciò il suo teorema nel suo intervento al Congresso Internazionale dei matematici tenutosi ad Amsterdam nel 1954. Ne pubblicò poi l'enunciato con una breve traccia della dimostrazione nella nota [55]. In quella nota viene esposto in forma sintetica ma esauriente lo schema formale del calcolo, e vengono dati i suggerimenti tecnici essenziali per la dimostrazione della convergenza del procedimento iterativo. Successivamente, come risulta dalla testimonianza di diversi matematici russi, egli ne diede la dimostrazione completa in una serie di lezioni, il cui testo però non venne pubblicato.

La dimostrazione completa venne pubblicata circa dieci anni dopo in due corposi lavori di Arnold[2][3], preceduti da un lavoro di Moser[72] in cui si dava un risultato equivalente per una mappa piana che conserva l'area. Alcuni anni più tardi Moser riprese la dimostrazione per il caso hamiltoniano ed arrivò a stabilire per via indiretta la convergenza delle serie di Lindstedt.[73]. Da questi lavori prese le mosse quella che oggi viene chiamata teoria KAM, acronimo di Kolmogorov–Arnold–Moser.

La dimostrazione che riporto qui riproduce lo schema suggerito in origine da Kolmogorov nella sua breve nota, fatto salvo un cambiamento tecnico che consiste nel far uso della tecnica delle serie di Lie al fine di semplificare la dimostrazione. Inoltre, ho cercato di ridurre al minimo le difficoltà tecniche considerando il caso particolare di un'Hamiltoniana che sia quadratica nelle azioni.[1]

Il contesto che consideriamo è quello del problema generale della dinamica, come enunciato da Poincaré ed esposto all'inizio del capitolo 6, e che richiamo per comodità. Si considera un sistema canonico con Hamiltoniana

$$H(p,q) = h(p) + \varepsilon f(p,q) \ , \tag{10.1}$$

dove $p \in \mathcal{G} \subset \mathbb{R}^n$ sono variabili d'azione in un aperto $\mathcal{G}$ e $q \in \mathbb{T}^n$ sono variabili angolari,

[1] Per una dimostrazione nel caso generale anch'essa ispirata al lavoro originale di Kolmogorov si veda, ad esempio, [8].

sotto l'ipotesi che $H(p,q)$ sia analitica in tutte le variabili e che sia sviluppabile in serie convergente di un piccolo parametro ε.

Il sistema imperturbato, per $\varepsilon = 0$, è caratterizzato da moti quasi periodici su tori invarianti. La domanda che si pone è quale sia il destino di questi tori quando si tenga in conto la perturbazione, ossia si ponga $\varepsilon \neq 0$. La risposta di Kolmogorov è che se ε è sufficientemente piccolo allora molti di questi tori vengono deformati ma non distrutti dalla perturbazione: sono i tori caratterizzati da frequenze fortemente non risonanti.

10.1 La forma normale di Kolmogorov

L'idea sottostante, ammirevole per la sua semplicità, è che si può identificare una classe di Hamiltoniane per le quali l'esistenza di un toro invariante è del tutto evidente: si tratta di quelle che hanno la forma

$$H(q,p) = \langle \omega, p \rangle + R(q,p) \tag{10.2}$$

dove $R(q,p)$ è almeno quadratica nelle azioni p, ovvero $R(q,p) = O(p^2)$. In effetti, le equazioni canoniche per una tale Hamiltoniana si scrivono

$$\dot{q}_j = \omega_j + \frac{\partial R}{\partial p_j} \ , \quad \dot{p}_j = -\frac{\partial R}{\partial q_j} \ , \quad j = 1, \ldots, n \ .$$

Poiché $\frac{\partial R}{\partial p_j} = O(p)$ e $\frac{\partial R}{\partial q_j} = O(p^2)$ si conclude immediatamente che il toro $p = 0$ è invariante per il flusso canonico, e che tale flusso è lineare con frequenze ω.

Nel seguito dirò che l'Hamiltoniana (10.2) è in *forma normale di Kolmogorov*.

10.1.1 Lo schema formale generale

Consideriamo dunque un sistema Hamiltoniano della forma (10.1). Poniamo anzitutto $\varepsilon = 0$, e scegliamo un punto $p^* \in \mathcal{G}$ tale che la frequenza imperturbata corrispondente $\omega = \frac{\partial h}{\partial p}(p^*)$ soddisfi la condizione diofantea[2]

$$|\langle k, \omega \rangle| \geq \gamma |k|^{-\tau} \quad \forall\, k \in \mathbb{Z}^n \ , \ k \neq 0 \ , \tag{10.3}$$

con $\gamma > 0$ e $\tau \geq n - 1$. In altre parole, abbiamo selezionato un toro invariante con frequenze fortemente irrazionali.

Sviluppiamo ora l'Hamiltoniana in serie di Taylor intorno a p^*, cosa che possiamo fare sotto le ipotesi di analiticità che abbiamo enunciato all'inizio. Otteniamo così

$$H(q,p) = \langle \omega, p - p^* \rangle + A(q) + \langle B(q), p - p^* \rangle + \frac{1}{2} \langle C(q)(p - p^*), p - p^* \rangle + g(q, p - p^*) \ ,$$

[2] L'ipotesi $\tau \geq n - 1$ sull'esponente merita una breve osservazione. In conseguenza di un teorema di approssimazione di Dirichlet non esistono vettori reali ω che soddisfino la condizione diofantea per $\tau < n - 1$. Per $\tau > n - 1$ il complemento degli ω che la soddisfano ha misura proporzionale a γ, e dunque la misura relativa degli ω diofantei tende a 1 al tendere di γ a zero. Per $\tau = n - 1$ esistono dei vettori ω diofantei, ma formano un insieme di misura nulla.

dove $g(q, p - p^*) = O\big((p - p^*)^3\big)$. Le funzioni $A(q)$, $B_j(q)$, $C_j(q)$ si calcolano come

$$\begin{aligned} A(q) &= \varepsilon f(q, p^*) \ , \\ B_j(q) &= \varepsilon \frac{\partial f}{\partial p_j}(q, p^*) \ , \\ C_{j,k}(q) &= \frac{\partial^2 h}{\partial p_j \partial p_k}(p^*) + \frac{\partial^2 f}{\partial p_j \partial p_k}(q, p^*) \ . \end{aligned}$$

Con una traslazione, che è una trasformazione canonica, possiamo anche trasportare l'origine delle azioni in p^*, il che equivale a porre $p^* = 0$ nell'espressione precedente dell'Hamiltoniana, sicché essa prende la forma più leggibile

$$H(q,p) = \langle \omega, p \rangle + A(q) + \langle B(q), p \rangle + \frac{1}{2} \langle C(q)(p), p \rangle + g(q,p) \ , \tag{10.4}$$

Osserviamo che se avessimo $A(q) = B(q) = 0$ allora l'Hamiltoniana sarebbe già in forma normale di Kolmogorov. In generale non sarà così, ma possiamo osservare che $A(q)$ e $B_j(q)$, per come sono definite, sono quantità di ordine ε.

Il suggerimento di Kolmogorov è di tentare di rimuovere i termini sgraditi, ossia quelli che contengono $A(q)$ e $B_j(q)$, mediante un procedimento di forma normale. Precisamente, si applica una trasformazione canonica prossima all'identità con funzione generatrice

$$\chi(q,p) = X(q) + \langle \xi, q \rangle + \langle Y(q), p \rangle \ ,$$

dove $X(q)$ è una funzione, $Y(q)$ un vettore di funzioni, e $\xi \in \mathbb{R}^n$ è un vettore reale, e tutte queste quantità devono determinarsi in modo da rimuovere dall'Hamiltoniana il termine $A(q) + \langle B(q), p \rangle$. Il procedimento di trasformazione è di fatto la costruzione di una forma normale particolare: quella di Kolmogorov.

Come avviene in tutti i procedimenti di forma normale, l'Hamiltoniana trasformata ha ancora la forma (10.4), il che significa che i termini sgraditi $A(q) + \langle B(q), p \rangle$ non sono stati rimossi completamente. Però possiamo sperare che essi risultino più piccoli. Euristicamente: se $A(q) + \langle B(q), p \rangle = O(\varepsilon)$, allora possiamo sperare che anche per la funzione generatrice valga $\chi(q,p) = O(\varepsilon)$, e che il nuovo termine $A(q) + \langle B(q), p \rangle$ sia $O(\varepsilon^2)$. Accettando per il momento che ciò sia vero, potremo allora iterare il procedimento, riducendo via via i termini sgraditi in modo che diventino successivamente $O(\varepsilon^2)$, $O(\varepsilon^4)$, $O(\varepsilon^8)$, $O(\varepsilon^{16})$, &c. In altre parole, ci si attende che dopo s passi del procedimento di costruzione della forma normale si abbia $A(q) + \langle B(q), p \rangle = O(\varepsilon^{(2^s)})$, sicché al limite di infinite iterazioni l'Hamiltoniana dovrebbe assumere la forma normale di Kolmogorov.

10.1.2 Il problema della convergenza

Le considerazioni euristiche che abbiamo appena svolto ci inducono a pensare che la convergenza del procedimento sia molto rapida, simile a quella che si ha nel metodo di Newton per la ricerca degli zeri di una funzione. In effetti, Kolmogorov stesso parlava di metodo simile a quello di Newton. Vale la pena di aggiungere tre osservazioni.

Nel procedimento che ho illustrato non c'è traccia di uno sviluppo in ε: si fa uso del solo fatto che certe funzioni sono più piccole di altre. Questo viene formalizzato tecnicamente dicendo che la norma di tali funzioni è piccola "di ordine ε^s" con s intero. Il lettore che, avendo acquisito familiarità con gli sviluppi perturbativi classici in un parametro, si trovasse a disagio potrà riscrivere tutte le formule anteponendo un ε alle funzioni $A(q)$ e $B_j(q)$. Ciò faciliterà la scrittura del singolo passo dell'algoritmo, ma lo stesso lettore si renderà conto che non si tratta di uno sviluppo nel parametro ε: questo serve solo per tenere memoria di cosa sia piccolo o grande.

Nello schema di Kolmogorov la rinuncia allo sviluppo in un parametro è lo strumento tecnico che permette di superare il problema dei piccoli divisori: la rapidità della convergenza quadratica riesce a dominarne comunque la crescita, anche quando essi si accumulano in un modo che rende apparentemente impossibile la convergenza degli sviluppi classici. Questo fatto è stato spesso presentato come il secondo elemento tecnico introdotto da Kolmogorov, considerato come cruciale per completare la dimostrazione del teorema. In molti testi si parla anche di metodo quadratico, per sottolineare l'efficacia del procedimento.

In realtà, il ragionamento euristico svolto sopra presenta il grave difetto di non tener conto dell'effetto per così dire esplosivo dell'accumulo dei piccoli divisori: nel corso della dimostrazione si vede come sia necessario sacrificare almeno parte della rapidità di convergenza dovuta al metodo quadratico proprio per tenere sotto controllo i piccoli divisori.

10.1.3 L'enunciato del teorema

Darò qui l'enunciato nel caso speciale di un'Hamiltoniana quadratica nelle azioni. Qui ed in tutto il resto del capitolo assumeremo che lo spazio delle fasi sia $\mathbb{R}^n \times \mathbb{T}^n$, in modo da rimuovere tutte le difficoltà tecniche dovute alla forma del dominio nelle azioni.

Teorema 10.1: *Consideriamo l'Hamiltoniana*

$$H(q,p) = \frac{1}{2}\langle \mathsf{J}p, p\rangle + \varepsilon f(p,q,\varepsilon) \ , \tag{10.5}$$

dove J è una matrice simmetrica reale $n \times n$, e $f(p,q,\varepsilon)$ è un polinomio (non necessariamente omogeneo) di grado 2 nelle azioni p ed è una funzione analitica di $q \in \mathbb{T}^n$ e del parametro ε, che si suppone sufficientemente piccolo. Assumiamo che la matrice J sia non degenere. Supponiamo ancora che p^* sia un toro invariante per il flusso dell'Hamiltoniana imperturbata, con $\varepsilon = 0$, e che le frequenze $\omega = \mathsf{J}p^*$ su tale toro soddisfino la condizione diofantea (10.3) con costanti $\gamma > 0$ e $\tau > n-1$. Allora esiste $\varepsilon_* > 0$ tale che per ogni $|\varepsilon| < \varepsilon_*$ l'Hamiltoniana (10.5) ammette un toro invariante contenuto in un intorno di p^* di raggio ε^a, con $0 < a < 1$, ed il flusso su tale toro è quasi periodico con frequenze ω.

La dimostrazione del teorema occuperà i prossimi paragrafi.

10.2 Algoritmo formale

Seguiamo il procedimento illustrato in modo sintetico nel paragrafo 10.1.1. Scegliamo un toro imperturbato p^* caratterizzato da frequenze $\omega = \mathsf{J}p^*$ di tipo diofanteo. Trasliamo poi l'origine delle azioni in p^*, e scriviamo l'Hamiltoniana come

$$H(q,p) = \langle \omega, p\rangle + A(q) + \langle B(q), p\rangle + \frac{1}{2}\langle \mathsf{C}(q)p, p\rangle \ , \tag{10.6}$$

dove

$$\begin{aligned} A(q) &= \varepsilon f(q,p^*) \ , \\ B_j(q) &= \varepsilon \frac{\partial f}{\partial p_j}(q,p^*) \ , \\ \mathsf{C}_{j,k}(q) &= \mathsf{J}_{j,k} + \varepsilon \frac{\partial^2 f}{\partial p_j \partial p_k}(q,p^*) \ . \end{aligned}$$

La matrice $\mathsf{C}(q)$, dipendente dagli angoli q, è simmetrica, e per ε sufficientemente piccolo è non degenere in virtù della non degenerazione di J.

Farò uso delle seguenti notazioni. Data una funzione $g(q,p)$ denoterò con $\overline{g}(p)$ la sua media sugli angoli q, ossia

$$\overline{g}(p) = \frac{1}{(2\pi)^n}\int_0^{2\pi} dq_1 \ldots \int_0^{2\pi} dq_n \, g(q,p) \ .$$

La parentesi di Poisson con la parte lineare $\langle \omega, p\rangle$ dell'Hamiltoniana imperturbata è un operatore lineare che avrà un ruolo particolare nel seguito. Per questo motivo denoterò

$$\partial_\omega \cdot = \{\cdot, \langle \omega, p\rangle\} = \left\langle \omega, \frac{\partial}{\partial q}\cdot \right\rangle$$

da vedersi come un operatore lineare che agisce sulle funzioni definite sullo spazio delle fasi. Nel seguito dovremo risolvere un'equazione della forma

$$\partial_\omega \varphi = \psi \tag{10.7}$$

dove $\psi(p,q)$ è una funzione nota. Si tratta di un'equazione lineare, che si risolve formalmente facendo uso dello sviluppo di Fourier. La soluzione non è unica, in quanto si può sempre aggiungere un elemento arbitrario del nucleo dell'operatore ∂_ω. Nel caso di frequenze non risonanti tale nucleo è costituito dalle funzioni a media nulla, ovvero dalle funzioni $\varphi(p)$ delle sole azioni. La soluzione $\varphi(p,q)$ può rendersi unica assumendo che valga $\overline{\varphi}(p) = 0$. Sotto questa condizione scriverò $\varphi = \partial_\omega^{-1}\psi$, il che definisce in modo univoco l'inverso ∂_ω^{-1} dell'operatore ∂_ω. L'algoritmo esplicito della soluzione verrà riportato più avanti, nel corso della dimostrazione del lemma 10.4 (lemma sui piccoli divisori).

10.2.1 L'algoritmo iterativo

Veniamo alla scrittura dell'algoritmo di forma normale in una forma algebrica esplicita. Raccoglierò qui tutte le formule necessarie per portare a termine un singolo passo di riduzione alla forma normale. Il procedimento completo risulterà dall'iterazione di questo passo.

Si applicano all'Hamiltoniana (10.6) due trasformazioni canoniche con funzioni generatrici

$$\chi_1(q) = X(q) + \langle \xi, q \rangle \ , \quad \chi_2(q,p) = \langle Y(q), p \rangle \ . \tag{10.8}$$

La funzione $X(q)$, il vettore reale ξ ed il vettore di funzioni $Y(q)$ sono determinati mediante le equazioni

$$\partial_\omega X - A = 0 \ , \tag{10.9}$$

$$\overline{\mathsf{C}}\xi - \overline{\left(B - \mathsf{C}\frac{\partial X}{\partial q} \right)} = 0 \ , \tag{10.10}$$

$$\partial_\omega Y - B + \mathsf{C}\left(\frac{\partial X}{\partial q} + \xi \right) = 0 \ . \tag{10.11}$$

L'Hamiltoniana trasformata viene calcolata come

$$H'(q,p) = \exp\bigl(L_{\chi_2}\bigr) \circ \exp\bigl(L_{\chi_1}\bigr) H(q,p) \ , \tag{10.12}$$

e si scrive ancora nella forma (10.6), con una nuova funzione $A'(q)$, un vettore di funzioni $B'(q)$ ed una matrice $\mathsf{C}'(q)$ date da

$$A' = \exp(L_{\langle Y,p \rangle}) \hat{A} \ , \tag{10.13}$$

$$\langle B', p \rangle = \sum_{j \geq 1} \frac{j}{(j+1)!} L^j_{\langle Y,p \rangle} \langle \hat{B}, p \rangle \ , \tag{10.14}$$

$$\langle \mathsf{C}'p, p \rangle = \langle \mathsf{C}p, p \rangle + \sum_{j \geq 1} \frac{1}{j!} L^j_{\langle Y,p \rangle} \langle \mathsf{C}p, p \rangle \ , \tag{10.15}$$

$$\hat{A} = \frac{1}{2} \left\langle \mathsf{C}\left(\frac{\partial X}{\partial q} + \xi \right), \left(\frac{\partial X}{\partial q} + \xi \right) \right\rangle + \left\langle B, \left(\frac{\partial X}{\partial q} + \xi \right) \right\rangle \ , \tag{10.16}$$

$$\hat{B} = B - \mathsf{C}\left(\frac{\partial X}{\partial q} + \xi \right) \ . \tag{10.17}$$

La trasformazione canonica sulle variabili d'angolo–azione si scrive in forma esplicita come

$$\begin{aligned} q &= \exp\bigl(L_{\langle Y,p \rangle}\bigr) q' \\ p &= \exp\bigl(L_{\langle Y,p \rangle}\bigr) \circ \exp\bigl(L_{X + \langle \xi,q \rangle}\bigr) p' \end{aligned} \tag{10.18}$$

Si noti che $\exp\bigl(L_{X+\langle \xi,q \rangle}\bigr) q' = q'$, sicché la trasformazione sugli angoli dovuta alla generatrice $\chi_1(q)$ si riduce all'identità; per questo è stata omessa nella prima riga dell'ultima formula.[3]

[3] A rigore, nell'Hamiltoniana trasformata (10.12) si dovrebbero denotare le nuove variabili con q', p', e non con q, p, tenendo conto della sostituzione. Ma l'uso delle serie di Lie giustifica il mantenimento dello stesso nome per le variabili: l'operatore esponenziale opera direttamente sulle funzioni, ed il nome delle variabili è irrilevante.

10.2.2 Giustificazione dell'algoritmo iterativo

Veniamo alla giustificazione dell'algoritmo formale. Si tratta di giocare un po' con gli sviluppi delle serie di Lie.

Iniziamo con la scrittura esplicita della prima trasformazione, con la funzione generatrice $\chi_1(q) = X(q) + \langle \xi, q\rangle$. La funzione $X(q)$ ed il vettore reale ξ dovranno considerarsi come incognite. Denotando con $\hat{H}$ l'Hamiltoniana trasformata troviamo[4]

$$\begin{aligned}\hat{H} = \exp\bigl(L_{\chi_1}\bigr)H =& \langle\omega,p\rangle + \frac{1}{2}\langle \mathsf{C}p,p\rangle\\ &+ A + L_X\langle\omega,p\rangle + L_{\langle\xi,q\rangle}\langle\omega,p\rangle\\ &+ \langle B,p\rangle + \frac{1}{2}L_{\chi_1}\langle \mathsf{C}p,p\rangle\\ &+ \frac{1}{4}L_{\chi_1}^2\langle \mathsf{C}p,p\rangle + L_{\chi_1}\langle B,p\rangle\ .\end{aligned}$$

Per inciso, si vede che $\hat{H}$ è ancora quadratica in p. Ci proponiamo ora di rimuovere il termine indesiderato A determinando X mediante l'equazione $\partial_\omega X + A = 0$, ossia la (10.9). Possiamo assumere che $A(q)$ abbia media nulla, ossia che $\overline{A} = 0$, dal momento che la sua media sarebbe una costante del tutto irrilevante nell'Hamiltoniana. Di conseguenza l'equazione ammette una soluzione, almeno formale (ci occuperemo più avanti della convergenza). Lasciando ξ ancora indeterminato, scriviamo l'Hamiltoniana trasformata $\hat{H}$ come

$$\hat{H}(q,p) = \langle\omega,p\rangle + \hat{A}(q) + \langle\hat{B}(q),p\rangle + \frac{1}{2}\langle\mathsf{C}(q)p,p\rangle\ , \tag{10.19}$$

dove $\hat{A}$ e $\hat{B}$ sono date dalle (10.16) e (10.17). Per ricavare le espressioni di $\hat{A}$ e $\hat{B}$ basta raccogliere tutti i termini indipendenti da p, che poniamo in $\hat{A}$, e lineari in p, che poniamo in $\hat{B}$. Si noti che $L_{\langle\xi,q\rangle}\langle\omega,p\rangle = \langle\omega,\xi\rangle$ è una costante, che si può omettere.

Passiamo alla scrittura della seconda trasformazione, con funzione generatrice $\chi_2(q,p) = \langle Y(q),p\rangle$. Scriviamo l'Hamiltoniana trasformata H' come

$$\begin{aligned}H' &= \exp\bigl(L_{\langle Y,p\rangle}\bigr)\hat{H}\\ &= \langle\omega,p\rangle + \exp\bigl(L_{\langle Y,p\rangle}\bigr)\hat{A}\\ &\quad+ \langle\hat{B},p\rangle + L_{\langle Y,p\rangle}\langle\omega,p\rangle + \sum_{j\geq 2}\frac{1}{j!}L_{\langle Y,p\rangle}^j\langle\omega,p\rangle + \sum_{j\geq 1}\frac{1}{j!}L_{\langle Y,p\rangle}^j\langle\hat{B},p\rangle\\ &\quad+ \frac{1}{2}\langle\mathsf{C}p,p\rangle + \frac{1}{2}\sum_{j\geq 1}\frac{1}{j!}L_{\langle Y,p\rangle}^j\langle\mathsf{C}p,p\rangle\ .\end{aligned} \tag{10.20}$$

---

[4] Si noti che la trasformazione mediante serie di Lie qui produce solo un numero finito di termini. Ciò si deve al fatto che χ_1 è indipendente dalle azioni p. Infatti si nota subito che la parentesi di Poisson con χ_1 decrementa di uno il grado di un polinomio in p. Questo implica $L_{\chi_1}A = L_{\chi_1}^2\langle B(q),p\rangle = L_{\chi_1}^3\langle\mathsf{C}p,p\rangle = 0$.

Anche qui si vede che H' resta quadratica in p. Per rimuovere il termine indesiderato $\langle\hat{B},p\rangle$ dobbiamo imporre la condizione

$$\langle\hat{B},p\rangle + L_{\langle Y,p\rangle}\langle\omega,p\rangle = 0 \ . \tag{10.21}$$

Ricordando la (10.17), questa equazione si riscrive in modo più esplicito

$$\partial_\omega Y + B + \mathsf{C}\left(\frac{\partial X}{\partial q} + \xi\right) = 0 \ ,$$

che altro non è che la (10.11). Dobbiamo però osservare che Y può determinarsi per inversione dell'operatore ∂_ω solo se la media del termine noto è nulla. Qui interviene il vettore ξ, fin qui indeterminato. Infatti imponendo che la media si annulli troviamo l'equazione

$$\overline{\mathsf{C}}\xi + \overline{B + \mathsf{C}\frac{\partial X}{\partial q}} = 0 \ ,$$

ovvero la (10.10). Questa ammette una soluzione in virtù dell'ipotesi che $\overline{\mathsf{C}}$ sia non degenere. Avendo determinato ξ, possiamo risolvere anche l'equazione (10.21).

Scriviamo ora la forma esplicita dell'Hamiltoniana trasformata H'; di fatto dobbiamo ricavare le (10.13)–(10.15). A tal fine dobbiamo separare nella (10.20) i monomi di diverso grado in p. Per i termini che coinvolgono $\hat{A}$ e C si tratta solo di trascriverli, e si ricavano immediatamente la (10.13) e la (10.15). Per i termini che coinvolgono $\hat{B}$ occorre sostituire la (10.21) nella (10.20), ottenendo

$$\begin{aligned} B' &= \sum_{j\geq 2}\frac{1}{j!}L^j_{\langle Y,p\rangle}\langle\omega,p\rangle + \sum_{j\geq 1}\frac{1}{j!}L^j_{\langle Y,p\rangle}\langle\hat{B},p\rangle \\ &= \sum_{j\geq 1}\frac{1}{(j+1)!}L^j_{\langle Y,p\rangle}\Big(L_{\langle Y,p\rangle}\langle\omega,p\rangle + (j+1)\langle\hat{B},p\rangle\Big) \ ; \end{aligned}$$

Facendo ancora uso della (10.21) si ricava subito la (10.14). Questo conclude il calcolo dell'Hamiltoniana trasformata.

Avendo determinato le funzioni generatrici, risulta definita anche la trasformazione di coordinate, che ha la forma (10.18). L'algoritmo formale è dunque ben definito.

10.3 Stime quantitative

Veniamo ora alla parte quantitativa delle dimostrazione. L'obiettivo è dare una stima delle norme di tutte le funzioni che intervengono nel calcolo della forma normale, così da garantire la convergenza del procedimento.

10.3.1 Strumenti analitici

Iniziamo introducendo norme opportune per tutte le funzioni.

(i) Per i vettori $x \in \mathbb{R}^n$ farò uso della norma

$$|x| = \sum_{j=1}^{n} |x_j| \ .$$

(ii) Per una funzione analitica $f(q)$, dove $q \in \mathbb{T}^n$ sono le variabili angolari farò uso della *norma di Fourier con peso*, definita come

$$\|f\|_\sigma = \sum_{k\in\mathbb{Z}^n} |f_k| e^{|k|\sigma} \ ,$$

dove f_k sono i coefficienti dello sviluppo di Fourier di $f(q)$, in generale complessi, e σ è una costante positiva. L'ipotesi che $f(q)$ sia analitica garantisce che esista σ tale che la serie che definisce la norma sia convergente.

(iii) Per una funzione vettoriale $w(q) = (w_1(q), \ldots, w_n(q))$ farò uso della norma

$$\|w\|_\sigma = \sum_{j=1}^{n} \|w_j\|_\sigma \ ,$$

dove $\|w_j\|_\sigma$ è la norma di Fourier con peso definita al punto (ii).

Le norme qui introdotte hanno alcune proprietà che si riveleranno utili nel corso della dimostrazione, e che raccolgo proseguendo la numerazione romana delle definizioni appena date.

Lemma 10.2: *Sia dato il dominio $\Delta_\varrho(0) \times \mathbb{T}^n_\sigma$, dove $\varrho > 0$ è arbitrario e $\Delta_\varrho(p_0)$ denota un polidisco di raggio ϱ con centro in p_0. Siano $w(q)$ e $v(q)$ funzioni analitiche a valori vettoriali e $\mathsf{C}(q)$ una matrice $n \times n$ i cui elementi $c_{jk}(q)$ siano funzioni analitiche. Valgono le seguenti proprietà.*

(iv) Per la norma di Fourier con peso vale

$$\|\langle w, p\rangle\|_{(\varrho,\sigma)} \le \varrho \|w\|_\sigma \ .$$

(v) Se $\|\langle w, p\rangle\|_{(\varrho,\sigma)} \le D\varrho$, con D positivo, allora si ha $\|w\|_\sigma \le nD$.

(vi) Se $\|\langle \mathsf{C}(q)p, p\rangle\|_{(\varrho,\sigma)} \le D\varrho^2$, allora si ha $\|C_{j,k}\|_\sigma \le D$.

(vii) $\|\langle w, v\rangle\|_\sigma \le \|w\|_\sigma \|v\|_\sigma$.

(viii) $\|\overline{f}\|_\sigma \le \|f\|_\sigma$ e $\|f - \overline{f}\|_\sigma \le \|f\|_\sigma$.

La dimostrazione non presenta particolari difficoltà. Le sole proprietà che richiedono una minima attenzione sono la (v) e la (vi). Per verificarle occorre modificare di poco la dimostrazione della prima delle diseguaglianze (9.33), lemma 9.8; si fa uso delle diseguaglianze di Cauchy con l'accortezza di restringere a zero il dominio nelle p lasciando invariato σ. Il calcolo è facilitato dal fatto che la dipendenza dalle p è puramente polinomiale, sicché si tratta soltanto di identificare le derivate con i coefficienti.

10.3.2 Lemma iterativo

Il lemma che segue rende quantitative le ipotesi sull'Hamiltoniana e mostra che l'algoritmo iterativo formale esposto nel paragrafo 10.2.1 produce un'Hamiltoniana trasformata che soddisfa le ipotesi richieste per l'Hamiltoniana di partenza, sicché in linea di principio il passo iterativo si può ripetere almeno un numero finito di volte.

Lemma 10.3: *Sia $H(q,p)$ un'Hamiltoniana della forma (10.6), e assumiamo le ipotesi seguenti:*

(i) esistono delle costanti positive σ e ε tali che

$$\max\bigl(\|A\|_\sigma, \|B\|_\sigma\bigr) \le \varepsilon \ ;$$

(ii) esiste una costante positiva $m \le 1$ tale che

$$m|x| \le |\overline{\mathsf{C}}x| \quad \text{per ogni } x \in \mathbb{R}^n \ ;$$

(iii) per ogni funzione $w(q)$ con norma $\|w\|_\sigma$ limitata si ha

$$\|\mathsf{C}w\|_\sigma \le m^{-1}\|w\|_\sigma \ ;$$

(iv) le frequenze ω soddisfano la condizione diofantea

$$|\langle k,\omega\rangle| \ge \gamma|k|^{-\tau} \quad \text{per } 0 \ne k \in \mathbb{Z}^n \ .$$

Siano $d \le 1/6$ e σ_* delle costanti positive, soddisfacenti

$$(1-3d)\sigma \ge \sigma_* \ .$$

Allora esiste una costante positiva $\Lambda = \Lambda(n,\tau,\gamma,\sigma_*)$ per cui vale l'affermazione seguente: se

$$\frac{\Lambda\varepsilon}{m^5 d^{3\tau+4}} \le 1 \tag{10.22}$$

allora esiste una trasformazione canonica prossima all'identità $(q,p) = \mathcal{C}(q',p')$ che trasforma l'Hamiltoniana nella forma (10.6) con le stesse frequenze ω e con nuove funzioni A', B' e C' che soddisfano ancora le ipotesi (i)–(iii) con nuove costanti ε', σ' e m' date da

$$\begin{aligned} \varepsilon' &= \frac{\Lambda}{m^4 d^{3\tau+4}\sigma}\varepsilon^2 \ , \\ \sigma' &= (1-3d)\sigma \ , \\ m' &= (1-d^{\tau+1})m \ . \end{aligned} \tag{10.23}$$

Inoltre per la trasformazione canonica valgono le stime

$$\begin{aligned} |p-p'| &\le \frac{\Lambda\varrho\varepsilon}{m^2 d^{2\tau+2}\sigma} < d\varrho \ , \\ |q-q'| &\le \frac{\Lambda\varepsilon}{m^2 d^{2\tau+2}} < d\sigma \end{aligned} \tag{10.24}$$

per ogni $(q',p') \in \mathbb{T}^n_{(1-2d)\sigma} \times \Delta_{(1-2d)\varrho}$.

Le ipotesi (i–iv) pongono in una forma quantitativa quelle formulate nell'enunciato del teorema. In particolare la (ii) garantisce l'invertibilità della matrice $\overline{\mathsf{C}}$, che ha elementi costanti essendo una media sugli angoli; la (iii) limita la norma della matrice $\mathsf{C}(q)$ non mediata. Il ruolo di queste condizioni apparirà chiaro nel corso della dimostrazione, che sarà svolta in dettaglio nel paragrafo 10.3.4. Occorrono però alcuni risultati preliminari.

10.3.3 Lemma sui piccoli divisori

In questo paragrafo dimostrerò un lemma che ha un ruolo cruciale nel controllare i piccoli divisori. L'obiettivo è dare una stima quantitativa per la soluzione delle equazioni (10.9) e (10.11) per le funzioni generatrici.

Lemma 10.4: *Sia $\psi(q)$ una funzione a media nulla, ossia $\overline{\psi} = 0$, e supponiamo che la norma $\|\psi\|_{(1-\delta)\sigma}$ sia limitata per qualche $\delta < 1$. Supponiamo che ω soddisfi la condizione diofantea (10.3). Sia poi $\varphi(q)$ la soluzione a media nulla dell'equazione $\partial_\omega\varphi = \psi$. Allora per ogni $d < 1-\delta$ positivo si ha*

$$\|\varphi\|_{(1-\delta-d)\sigma} \le \frac{1}{\gamma}\left(\frac{\tau}{ed\sigma}\right)^{\tau}\|\psi\|_{(1-\delta)\sigma}\ , \tag{10.25}$$

$$\left\|\frac{\partial\varphi}{\partial q_j}\right\|_{(1-\delta-d)\sigma} \le \frac{1}{\gamma}\left(\frac{\tau+1}{ed\sigma}\right)^{\tau+1}\|\psi\|_{(1-\delta)\sigma} \tag{10.26}$$

Dimostrazione. Scrivendo lo sviluppo in serie di Fourier $\psi(q) = \sum_{0\neq k\in\mathbb{Z}^n}\psi_k e^{i\langle k,q\rangle}$ con coefficienti ψ_k noti, ed assumendo che $\varphi(q)$ abbia la stessa forma con coefficienti φ_k da determinarsi, si calcola

$$\partial_\omega\varphi = i\sum_{0\neq k\in\mathbb{Z}^n}\langle k,\omega\rangle\varphi_k e^{i\langle k,q\rangle}\ .$$

Per confronto dei coefficienti dello stesso modo di Fourier si ricava subito la soluzione dell'equazione come

$$\varphi = -i\sum_{0\neq k\in\mathbb{Z}^n}\frac{\psi_k}{\langle k,\omega\rangle}e^{i\langle k,q\rangle}\ ,$$

resa unica imponendo che la media sia nulla, ossia che non compaia nella somma il termine $k=0$. La condizione di non risonanza assicura che nessuno dei denominatori si annulla. Grazie alla definizione della norma possiamo stimare

$$\begin{aligned}\|\varphi\|_{(1-\delta-d)\sigma} &\le \frac{1}{\gamma}\sum_k |k|^{\tau}|\psi_k|_\varrho e^{|k|(1-\delta-d)\sigma}\\ &\le \frac{1}{\gamma}\left(\frac{\tau}{ed\sigma}\right)^{\tau}\sum_k|\psi_k|_\varrho e^{|k|(1-\delta)\sigma}\ ;\end{aligned}$$

qui si usa la diseguaglianza $|k|^\tau e^{-|k|d\sigma} \le \big(\tau/(ed\sigma)\big)^\tau$, che è una riscrittura della (9.35) sostituendo x con $|k|$, α con τ e β con $d\sigma$. La prima diseguaglianza è conseguenza diretta della definizione di norma. Per verificare la seconda diseguaglianza occorre calcolare

$$\frac{\partial \varphi}{\partial q_j} = \sum_k \frac{k_j \psi_k}{\langle k, \omega \rangle} e^{i\langle k,q\rangle} \ .$$

Poiché $|k| = |k_1| + \ldots + |k_n|$, possiamo stimare

$$\left\| \frac{\partial \varphi}{\partial q_j} \right\|_{(1-\delta-d)\sigma} \leq \frac{1}{\gamma} \sum_k |k|^{\tau+1} |\psi_k|_\varrho e^{|k|(1-\delta-d)\sigma} \ .$$

La (10.26) si ricava ripetendo l'argomento appena esposto, mutatis mutandis. *Q.E.D.*

10.3.4 Dimostrazione del lemma iterativo

La dimostrazione del lemma 10.3 si basa sull'applicazione sistematica delle stime del lemma 10.4 e di quelle per le parentesi di Poisson e le serie di Lie.

Iniziamo dalle stime sulla funzione generatrice. Per il lemma 10.4 la soluzione $X(q)$ dell'equazione (10.9), per l'ipotesi (i), soddisfa la diseguaglianza

$$\left\| \frac{\partial X}{\partial q} \right\|_{(1-d)\sigma} \leq \frac{K_1}{d^{\tau+1}} \varepsilon \ , \quad K_1 = \frac{1}{\gamma} \left(\frac{\tau+1}{e\sigma_*} \right)^{\tau+1} \ . \tag{10.27}$$

Per le ipotesi (i) e (iii), nella (10.10) abbiamo (ricordando che $m, d < 1$)

$$\left\| B + \mathsf{C} \frac{\partial X}{\partial q} \right\|_{(1-d)\sigma} \leq \frac{K_1 + 1}{m d^{\tau+1}} \varepsilon \ ; \tag{10.28}$$

di conseguenza, per l'ipotesi (ii) e per la (10.10) otteniamo

$$m|\xi| \leq \left| \overline{\mathsf{C}} \xi \right| \leq \frac{K_1 + 1}{m d^{\tau+1}} \varepsilon \ ,$$

e quindi anche

$$|\xi| \leq \frac{K_1 + 1}{m^2 d^{\tau+1}} \varepsilon \ . \tag{10.29}$$

Grazie a questa diseguaglianza ed alla (10.28) arriviamo alla stima di $\hat{B}$ data dalla (10.17). Ricordando poi che $m \leq 1$ e facendo uso ancora dell'ipotesi (iii) abbiamo

$$\|\hat{B}\|_{(1-d)\sigma} \leq \frac{2(K_1 + 1)}{m^3 d^{\tau+1}} \varepsilon \ . \tag{10.30}$$

Infine, risolvendo la (10.11) otteniamo

$$\|Y\|_{(1-2d)\sigma} \leq \frac{K_2}{m^3 d^{2\tau+1}} \varepsilon \ , \quad K_2 = \frac{2(K_1 + 1)}{\gamma} \left(\frac{\tau}{e\sigma_*} \right)^{\tau} \ . \tag{10.31}$$

Abbiamo così ricavato le stime per le funzioni generatrici.

Veniamo ora alle stime sull'Hamiltoniana trasformata. Dalla (10.16), tenendo conto dell'ipotesi (iii), e delle (10.27) e (10.29) abbiamo

$$\|\hat{A}\|_{(1-d)\sigma} \le \frac{(2K_1+1)^2}{m^5 d^{2\tau+2}}\varepsilon^2 \ ;$$

Ponendo questa stima nella (10.13), ed assumendo la condizione di convergenza

$$\frac{4eK_2}{m^5 d^{2\tau+3}\sigma}\varepsilon \le 1 \ , \tag{10.32}$$

possiamo stimare

$$\|A'\|_{(1-3d)\sigma} \le \frac{K_3}{m^5 d^{2\tau+2}}\varepsilon^2 \ , \quad K_3 = \frac{(e+1)(2K_1+1)^2}{e} \ . \tag{10.33}$$

Qui è opportuno aggiungere un commento. Per la proprietà (iv) enunciata all'inizio di questo paragrafo e per la (10.31) la norma della funzione generatrice $\langle Y,p\rangle$ viene stimata sul dominio $\mathbb{T}^n_{(1-2d)\sigma}\times\Delta_{(1-2d)\varrho}$, con ϱ arbitrario, da

$$\|\langle Y,p\rangle\|_{(1-2d)(\varrho,\sigma)} \le \frac{K_2\varrho}{m^3 d^{2\tau+1}}\varepsilon \ .$$

D'altra parte, sappiamo dalle stime sulle serie di Lie che per una funzione $\psi(q,p)$ analitica in $\mathbb{T}^n_{(1-2d)\sigma}\times\Delta_{(1-2d)\varrho}$ con norma $\|\psi\|_{(1-2d)(\varrho,\sigma)}$ limitata, sotto la condizione di convergenza (10.32), abbiamo

$$\begin{aligned}\big\|\exp L_{\langle Y,p\rangle}\psi-\psi\big\|_{(1-3d)(\varrho\sigma)} &\le \frac{2\|\langle Y,p\rangle\|_{(1-2d)(\varrho,\sigma)}}{ed^2\varrho\sigma}\|\psi\|_{(1-2d)(\varrho,\sigma)}\\ &\le \frac{2K_2\varepsilon}{em^3 d^{2\tau+3}\sigma}\|\psi\|_{(1-2d)(\varrho,\sigma)} \le \frac{1}{e}\|\psi\|_{(1-2d)(\varrho,\sigma)} \ .\end{aligned}$$

Facendo uso della stima per $\|\hat{A}\|_{(1-d)\sigma}$ che abbiamo già ricavato otteniamo la (10.33). La stima di $\langle B',p\rangle$ data dalla (10.14) si ricava con un argomento simile. Basta infatti osservare che vale

$$\begin{aligned}\|\langle B',p\rangle\|_{(1-3d)(\varrho,\sigma)} &\le \sum_{j\ge1}\frac{j}{(j+1)!}\big\|L^j_{\langle Y,p\rangle}\langle\hat{B},p\rangle\big\|_{(1-3d)(\varrho,\sigma)}\\ &\le \sum_{j\ge1}\frac{1}{j!}\big\|L^j_{\langle Y,p\rangle}\langle\hat{B},p\rangle\big\|_{(1-3d)(\varrho,\sigma)} \ ;\end{aligned}$$

in altre parole, si usa di nuovo lo stesso tipo di stima che abbiamo già visto per $\exp\big(L_{\langle Y,p\rangle}\big)\langle\hat{B},p\rangle-\langle\hat{B},p\rangle$. Stimando $\hat{B}$ mediante la (10.30) otteniamo

$$\|\langle B',p\rangle\|_{(1-3d)(\varrho,\sigma)} \le \frac{K_4\varrho}{2nm^6 d^{3\tau+4}}\varepsilon^2 \ , \quad K_4 = \frac{8n(K_1+1)}{e\sigma_*} \ . \tag{10.34}$$

Da qui, ricordando che $d\le 1/6$, e facendo uso della proprietà (v) del lemma 10.2, abbiamo

$$\|B'\|_{(1-3d)(\varrho,\sigma)} \le \frac{K_4\varrho}{m^6 d^{3\tau+4}}\varepsilon^2 \ . \tag{10.35}$$

Veniamo infine alla stima di C'. Per l'ipotesi (iii)

$$\|\langle \mathsf{C}p, p\rangle\|_{(\varrho,\sigma)} \le \frac{\varrho^2}{m} \ .$$

Stimando la serie nella (10.15) mediante l'argomento che abbiamo appena usato per le funzioni abbiamo

$$\bigl\|\langle (\mathsf{C}' - \mathsf{C})p, p\rangle\bigr\|_{(1-3d)(\varrho,\sigma)} \le \frac{K_5\varrho^2}{4nm^4 d^{2\tau+3}}\varepsilon \ , \quad K_5 = \frac{8nK_2}{e\sigma_*} \ . \tag{10.36}$$

Da qui, grazie alla proprietà (vi) del paragrafo 10.3.1, otteniamo

$$\|\mathsf{C}'_{j,k} - \mathsf{C}_{j,k}\|_{(1-3d)\sigma} \le \frac{K_5}{nm^4 d^{2\tau+3}}\varepsilon \ .$$

Assumendo che ε soddisfi la condizione

$$\tilde{m} := m - \frac{K_5}{m^4 d^{2\tau+3}}\varepsilon \ge m_* > 0 \ . \tag{10.37}$$

possiamo dunque stimare

$$\bigl|\overline{\mathsf{C}'}x\bigr| \ge \bigl|\overline{\mathsf{C}}x\bigr| - \bigl|\bigl(\overline{\mathsf{C}' - \mathsf{C}}\bigr)x\bigr| \ge \left(m - \frac{K_5}{m^4 d^{2\tau+3}}\varepsilon\right)|x| \ , \tag{10.38}$$

D'altra parte, tenuto conto della definizione (10.37) di $\tilde{m}$ e facendo uso della disegua-glianza elementare $a^{-1} + b < (a-b)^{-1}$ per $0 < b < a < 1$, per una funzione vettoriale $w(q)$ abbiamo

$$\begin{aligned}\|\mathsf{C}'w\|_{(1-3d)\sigma} &\le \|\mathsf{C}w\|_{(1-3d)\sigma} + \|(\mathsf{C}' - \mathsf{C})w\|_{(1-3d)\sigma}\\ &\le \left(\frac{1}{m} + \frac{K_5}{m^4 d^{2\tau+3}}\varepsilon\right)\|w\|_{(1-3d)\sigma}\\ &< \frac{1}{\tilde{m}}\|w\|_{(1-3d)\sigma} \ ,\end{aligned} \tag{10.39}$$

Abbiamo così completato le stime sulla funzione generatrice.
Resta da stimare la trasformazione canonica (10.18). Anche qui, procediamo in due passi eseguendo la trasformazione con $\exp\bigl(L_{X+\langle\xi,q\rangle}\bigr)$, seguita da quella con $\exp\bigl(L_{\langle Y,p\rangle}\bigr)$. La prima trasformazione si scrive in modo esplicito come

$$q = \hat{q} \ , \quad p = \hat{p} + \xi + \frac{\partial X}{\partial q}\Big|_{q=\hat{q},p=\hat{p}} \ . \tag{10.40}$$

La seconda trasformazione è

$$\hat{q} = \exp\bigl(L_{\langle Y,p\rangle}\bigr)q\Big|_{q=q'} \ , \quad \hat{p} = \exp\bigl(L_{\langle Y,p\rangle}\bigr)p\Big|_{q=q',p=p'} \ . \tag{10.41}$$

La stima fa uso del fatto che la norma di Fourier con peso fornisce anche un limite superiore sul massimo di una funzione. Per la (10.27) e la (10.29) la prima trasformazione viene stimata da

$$\bigl|p_j - \hat{p}_j\bigr| \le \frac{2K_1 + 1}{m^2 d^{\tau+1}}\varepsilon \tag{10.42}$$

Per la seconda trasformazione, dalla stima generale per le serie di Lie abbiamo

$$
\begin{aligned}
\bigl|\hat{q}_j - q_j\bigr| &\le \frac{K_2}{m^3 d^{2\tau+2}}\varepsilon \\
\bigl|\hat{p}_j - p_j\bigr| &\le \frac{K_2 \varrho}{m^3 d^{2\tau+2}\sigma}\varepsilon
\end{aligned}
\tag{10.43}
$$

Questo completa la stima sulla trasformazione canonica.
Per concludere la dimostrazione del lemma dobbiamo determinare Λ raccogliendo tutte le stime fatte fin qui. Definiamo

$$
\Lambda = \max\left(\frac{4eK_2}{\sigma_*}, K_3, K_4, K_5\right)
\tag{10.44}
$$

con K_2 , K_3 , K_4 e K_5 dati dalle (10.31), (10.33), (10.34) e (10.36). Da qui si vede che Λ dipende da n , τ , γ e σ_*, come affermato. Inoltre la definizione stessa di Λ mostra che la condizione di convergenza (10.32) è soddisfatta in virtù dell'ipotesi (10.22). Dalle (10.33) e (10.35) abbiamo $\max\bigl(\|A'\|_{\sigma'}, \|B'\|_{\sigma'}\bigr) \le \varepsilon'$, con ε' dato dalla prima delle (10.23). Per la stima su C', occorre usare la condizione di convergenza (10.22), e verificare che nella (10.37) si ha

$$
\frac{K_5 \varepsilon}{m^4 d^{2\tau+3}} \le \frac{\Lambda\varepsilon}{m^4 d^{2\tau+3}} \le m d^{\tau+1} \ ,
$$

Ne segue che la (10.38) e la (10.39) sono soddisfatte a patto di sostituire $\tilde{m}$ con m', così come dato dalla (10.23), grazie al fatto che $m' \le \tilde{m}$; questo dà la stima su C'. Le stime sulla trasformazione canonica si ricavano in modo simile facendo uso della condizione (10.22) e della definizione di Λ. Il lemma 10.3 è così dimostrato.

10.4 Iterazione delle stime e conclusione della dimostrazione.

Applicando successivamente il lemma iterativo possiamo costruire una successione infinita di trasformazioni canoniche, che denoteremo con $\{\hat{\mathcal{C}}^{(k)}\}_{k\ge 1}$, della forma

$$
\bigl(q^{(k-1)}, p^{(k-1)}\bigr) = \hat{\mathcal{C}}^{(k)}\bigl(q^{(k)}, p^{(k)}\bigr)
$$

(l'indice in alto identifica le coordinate al passo k–esimo nell'applicazione iterativa dell'algoritmo). Si costruisce dunque una successione di Hamiltoniane $\{H^{(k)}\}_{k\ge 0}$, in cui $H^{(0)} = H$ è quella di partenza, che soddisfano

$$
\begin{aligned}
&\max\bigl(\|A^{(k)}\|_{\sigma_k}, \|B^{(k)}\|_{\sigma_k}\bigr) \le \varepsilon_k \ , \\
&\bigl|\overline{\mathsf{C}}^{(k)} v\bigr| \ge m_k |v| \quad \text{per ogni } v \in \mathbb{R}^n \ , \\
&\bigl\|\mathsf{C}^{(k)} w\bigr\|_{\sigma_k} \le \frac{1}{m_k}\|w\|_{\sigma_k} \ ;
\end{aligned}
\tag{10.45}
$$

l'ultima stima vale per qualunque funzione vettoriale analitica $w(q)$. Le successioni $\{\varepsilon_k\}_{k\ge 0}$, $\{\sigma_k\}_{k\ge 0}$ e $\{m_k\}_{k\ge 0}$ sono definite ricorsivamente a partire da $\varepsilon_0 = \varepsilon$, $\sigma_0 = \sigma$,

$m_0 = m$ ponendo

$$\varepsilon_k = \frac{\Lambda}{m_{k-1}^5 d_k^{3\tau+4}} \varepsilon_{k-1}^2 \ , \tag{10.46}$$

$$\sigma_k = (1 - 3d_k)\sigma_{k-1} \ , \tag{10.47}$$

$$m_k = (1 - d_k^{\tau+1}) m_{k-1} \ . \tag{10.48}$$

Queste successioni dipendono ancora dalla successione $\{d_k\}_{k\geq 1}$, e questa a sua volta è soggetta alle sole condizioni, $d_k \leq 1/6$ e

$$\frac{\Lambda \varepsilon_{k-1}}{m_{k-1}^5 d_k^{3\tau+4}} \leq 1 \ , \tag{10.49}$$

$$(1 - 3d_k)\sigma_{k-1} \geq \sigma_* > 0 \ , \tag{10.50}$$

$$(1 - d_k^{\tau+1}) m_{k-1} \geq m_* > 0 \tag{10.51}$$

che devono essere soddisfatte per ogni $k \geq 1$ e per opportune costanti σ_* e m_*, e per il resto è arbitraria. Inoltre le trasformazioni canoniche soddisfano

$$\begin{aligned} |p^{(k)} - p^{(k-1)}| &\leq \frac{\Lambda \varrho}{m_{k-1}^3 d_k^{2\tau+2} \sigma_k} \varepsilon_{k-1} < d_k \varrho \ , \\ |q^{(k)} - q^{(k-1)}| &\leq \frac{\Lambda}{m_{k-1}^3 d_k^{2\tau+2}} \varepsilon_{k-1} < d_k \sigma_{k-1} \ . \end{aligned} \tag{10.52}$$

Dobbiamo anzitutto scegliere la successione $\{d_k\}_{k\geq 1}$ in modo che per ogni $k \geq 1$ siano soddisfatte le condizioni (10.49), (10.50) e (10.51) e che la successione $\{\varepsilon_k\}_{k\geq 0}$ converga a zero, e poi mostrare che la successione delle trasformazioni canoniche converge ad una trasformazione analitica che pone l'Hamiltoniana in forma normale di Kolmogorov.

Un modo conveniente consiste nel vedere la (10.46) come una relazione che ci permette di determinare d_k se ε_{k-1} e ε_k sono noti. Ricavando d_k dalla (10.46) abbiamo

$$d_k = \left(\frac{\Lambda \varepsilon_{k-1}^2}{m_{k-1}^5 \varepsilon_k} \right)^{1/(3\tau+4)} \ .$$

Possiamo così guardare alla successione $\{\varepsilon_k\}_{k\geq 0}$ come arbitraria, e calcolare tutto il resto in funzione di essa. A tal fine poniamo[5]

$$\varepsilon_k = \frac{\varepsilon_0}{(k+1)^{2(3\tau+4)}} \ . \tag{10.53}$$

Supponiamo per un momento che le condizioni $m_k \geq m_* > 0$ restino soddisfatte, cosa

[5] La scelta qui è soggetta alle sole condizioni $\varepsilon_k \to 0$ per $k \to +\infty$ e $\sum_k d_k \leq 1/6$, e per il resto è del tutto arbitraria. Si può anche scegliere, ad esempio, $\varepsilon_k \sim C^{-k}$ con una qualunque costante $C > 1$. Alcuni autori preferiscono porre, ad esempio, $\varepsilon_k = \varepsilon_{k-1}^{3/2}$, in modo da mantenere comunque la proprietà di superconvergenza del procedimento.

che verificheremo tra poco. Allora abbiamo

$$d_k \le \frac{4}{k^2}\left(\frac{\Lambda\varepsilon_0}{m_*^5}\right)^{1/(3\tau+4)} ,$$

e dunque anche[6]

$$\sum_{k\ge1} d_k < \frac{2\pi^2}{3}\left(\frac{\Lambda\varepsilon_0}{m_*^5}\right)^{1/(3\tau+4)} . \tag{10.54}$$

Qui imponiamo la condizione che $\varepsilon \equiv \varepsilon_0$ sia sufficientemente piccolo, e la rendiamo quantitativa richiedendo

$$4\pi^2\left(\frac{\Lambda\varepsilon_0}{m_*^4}\right)^{1/(3\tau+4)} \le 1 . \tag{10.55}$$

Da qui otteniamo

$$\sum_{k\ge1} d_k < \frac{1}{6} , \tag{10.56}$$

e questo implica che valga anche anche $d_k < 1/6$ per ogni $k \ge 1$. Dimostriamo ora che sono soddisfatte le (10.50) e (10.51). Iniziamo dalla (10.50). Scriviamo

$$\ln\prod_{k\ge1}(1-3d_k) = \sum_{k\ge1}\ln(1-3d_k) .$$

Facendo uso della diseguaglianza elementare

$$0 \ge \ln(1-x) \ge -2x\ln 2 \quad \text{per } 0 \le x \le 1/2 ,$$

possiamo valutare

$$0 \ge \sum_{k\ge1}\ln(1-3d_k) \ge -6\ln 2\sum_{k\ge1} d_k \ge -\ln 2 .$$

Ne segue che $\prod_{k\ge1}(1-3d_k) \ge 1/2$, sicché la (10.50) è soddisfatta scegliendo, ad esempio,

$$\sigma_* = \frac{\sigma}{2} . \tag{10.57}$$

Analogamente, tenendo conto che $d_k^{\tau+1} < d_k$, valutiamo $\prod_{k\ge1}(1-d_k) > 1/2^{1/3}$, sicché la (10.51) è soddisfatta scegliendo, ad esempio,

$$m_* = \frac{m_0}{2^{1/3}} . \tag{10.58}$$

La condizione (10.49) è chiaramente soddisfatta per come è stata scelta la successione ε_k nella (10.53). Infatti, per confronto con la (10.46), ci si riconduce alla condizione

---

[6] Si ricordi che $\sum_{k>0} k^{-2} = \pi^2/6$.

$\varepsilon_k \leq \varepsilon_{k-1}$, che è palesemente vera. Abbiamo dunque ridotto tutte le condizioni imposte nella dimostrazione alla sola (10.55) che deve essere soddisfatta da ε_0.

Resta da dimostrare che la trasformazione canonica è ben definita in un dominio opportuno. Fissato $\varrho_0 > 0$, consideriamo la successione di domini $\{\mathbb{T}^n_{\sigma_k} \times \Delta_{\varrho_k}(0)\}_{k\geq 0}$, con σ_k dato dalla (10.47) e $\varrho_k = (1-3d_k)\varrho_{k-1}$. Allora la trasformazione canonica

$$\begin{array}{rcl} \hat{\mathcal{C}}^{(k)} : \mathbb{T}^n_{\sigma_k} \times \Delta_{\varrho_k}(0) & \to & \mathbb{T}^n_{\sigma_{k-1}} \times \Delta_{\varrho_{n-1}} \\ (q^{(k)}, p^{(k)}) & \mapsto & (q^{(k-1)}, p^{(k-1)}) = \mathcal{C}^{(k)}(q^{(k)}, p^{(k)}) \end{array}$$

è analitica. Ne segue che anche la trasformazione

$$\hat{\mathcal{C}}^{(k)} \ : \ \mathbb{T}^n_{\sigma_k} \times \Delta_{\varrho_k}(0) \quad \to \quad \mathbb{T}^n_{\sigma_0} \times \Delta_{\varrho_0}$$

definita come

$$\mathcal{C}^{(k)} = \hat{\mathcal{C}}^{(k)} \circ \cdots \circ \hat{\mathcal{C}}^{(1)}$$

è canonica ed analitica, essendo composizione di trasformazioni canoniche ed analitiche. D'altra parte, per la (10.24), abbiamo

$$\left|q^{(k)} - q^{(0)}\right| < \left|q^{(k)} - q^{(k-1)}\right| + \ldots + \left|q^{(1)} - q^{(0)}\right| < \sigma \sum_{j=1}^{k} d_j \ ,$$

$$\left|p^{(k)} - p^{(0)}\right| < \left|p^{(k)} - p^{(k-1)}\right| + \ldots + \left|p^{(1)} - p^{(0)}\right| < \varrho \sum_{j=1}^{k} d \ .$$

Grazie alla convergenza della serie $\sum_{j\geq 1} d_j$ la successione delle trasformazioni $\mathcal{C}^k$ converge assolutamente ad una trasformazione

$$\mathcal{C}^{(\infty)} \ : \ \mathbb{T}^n_{\sigma_*} \times \Delta_{\varrho_*}(0) \quad \to \quad \mathbb{T}^n_{\sigma_0} \times \Delta_{\varrho_0}(0)$$

con, ad esempio, $\varrho_* = \varrho_0/2$; ciò in virtù della (10.57). La convergenza assoluta assicura anche la convergenza uniforme in ogni sottinsieme compatto di $\mathbb{T}^n_{\sigma_*} \times \Delta_{\varrho_*}(0)$. Possiamo dunque applicare il teorema di Weierstrass, sicché $\mathcal{C}^{(\infty)}$ è analitica. Infine, denotando con $q^{(\infty, p^{(\infty)})}$ le coordinate canoniche su $\mathbb{T}^n_{\sigma_*} \times \Delta_{\varrho_*}(0)$, e facendo ancora uso delle (10.52) e (10.53), abbiamo

$$\left|q^{(\infty)} - q^{(0)}\right| < \frac{2\pi^2}{3}\left(\frac{\Lambda\varepsilon_0}{m_*^4}\right)^{1/(3\tau+4)} \varrho \ ,$$

$$\left|p^{(\infty)} - p^{(0)}\right| < \frac{2\pi^2}{3}\left(\frac{\Lambda\varepsilon_0}{m_*^4}\right)^{1/(3\tau+4)} \sigma \ ,$$

sicché per $\varepsilon = 0$ la trasformazione è l'identità. Per le proprietà delle serie di Lie si ha anche $H^{(k)} = H^{(0)} \circ \mathcal{C}^{(k)}$, sicché la successione $H^{(k)}$ converge ad una funzione analitica $H^{(\infty)}$ che per costruzione è in forma normale di Kolmogorov. Il teorema è dunque dimostrato.