DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **12**, Number **4**, April **2005**

Website: http://aimSciences.org

pp. **569–594**

EXPONENTIAL STABILITY FOR THE RESONANT D'ALEMBERT MODEL OF CELESTIAL MECHANICS

Luca Biasco and Luigi Chierchia

Dipartimento di Matematica
Università "Roma Tre"
Largo S. L. Murialdo 1, 00146 Roma (Italy)

(Communicated by Carlangelo Liverani)

Abstract. We consider the classical D'Alembert Hamiltonian model for a rotationally symmetric planet revolving on Keplerian ellipse around a fixed star in an almost exact "day/year" resonance and prove that, notwithstanding proper degeneracies, the system is stable for exponentially long times, provided the oblateness and the eccentricity are suitably small.

1. **Introduction.** Perturbative techniques are a basic tool for a deep understanding of the long time behavior of conservative Dynamical Systems. Such techniques, which include the so–called KAM and Nekhoroshev theories (compare [1] for generalities), have, by now, reached a high degree of sophistication and have been applied to a great corpus of different situations (including infinite dimensional systems and PDE's). KAM and Nekhoroshev techniques work under suitable non–degeneracy assumptions (e.g, invertibility of the frequency map for KAM or steepness for Nekhoroshev: see [1]); however such assumptions are strongly violated exactly in those typical examples of Celestial Mechanics, which – as well known – were the main motivation for the Dynamical System investigations of Poincaré, Birkhoff, Siegel, Kolmogorov, Arnold, Moser,...

In this paper we shall consider the day/year (or spin/orbit) resonant planetary D'Alembert model (see [10]) and will address the problem of the long time stability for such a model.

The D'Alembert planetary model is a Hamiltonian model for a rotationally symmetric planet (or satellite) with polar radius slightly smaller than the equatorial radius; the center of mass of the planet revolves periodically on a given Keplerian ellipse of small eccentricity around a fixed star (or "major body") occupying one of the foci of the ellipse; the planet is subject only to the gravitational attraction of the major body. This system is modelled by a Hamiltonian system of two and a half degrees of freedom depending on two action–variables J_1 and J_2 corresponding, respectively, to the absolute value of the angular momentum of the planet and

2000 *Mathematics Subject Classification.* 37C75, 70H08, 70K43, 70K45, 34D10, 34C29.
Key words and phrases. KAM and Nekhoroshev theory, stability, nonlinear dynamics, Hamiltonian systems,D'Alembert model, Celestial Mechanics, proper degeneracies, fast averaging, action-angle variables.
Supported by M.I.U.R. *Variational Methods and Nonlinear Differential Equations*.

its projection onto the unit normal to the ecliptic plane[1], on the two conjugated angle–variables and on time; the dependence upon time is T–periodic, T being the "year" of the planet (i.e., the period of the Keplerian motion). The D'Alembert model is integrable if the oblateness[2] ε of the planet vanishes, while if the eccentricity μ of the Keplerian ellipse is zero the system becomes time–independent. The integrable approximation ($\varepsilon = 0$) is *properly degenerate*: in suitable variables, the D'Alembert Hamiltonian depends only upon the absolute value of the angular momentum J_1 of the planet. Clearly, J_1 is stable (i.e., stay close to its initial value) for the full Hamiltonian and the (action) stability problem consists then in studying the stability of J_2, the projection of the angular momentum onto the unit normal to the ecliptic plane (which measures the inclination of the spin axis on the ecliptic plane). The "resonant model" deals with phase space regions around exact day/year resonances, i.e., regions where the ratio between the periods of rotation and revolution is a rational number p/q. More precisely, we shall fix $0 < \ell < 1/2$ and consider an ε^ℓ–neighborhood of the exact resonance $\bar{J}_1 = (2\pi/T)\,(p/q)$. Then, we shall prove the following

Theorem 1.1. *Fix $c > 0$ and consider the D'Alembert model with $0 < \mu \le \varepsilon^c$ in an $O(\varepsilon^\ell)$ neighborhood of an exact day/year resonance $\bar{J}_1$ as above. Assume that $\bar{J}_1 \ne \sqrt{3}L$, where L denotes the projection of the angular momentum of the planet onto the polar axis. Then, the evolution of the angular momentum of the planet stay close to its initial position for times exponentially long in $1/\varepsilon$, provided ε is small enough.*

A more precise formulation of this theorem is given in the next section, where the Hamiltonian description of the resonant D'Alembert model is recalled. In § 3 the proof of Theorem 1.1 is given; such proof is significantly easier in the case the resonance (p, q) is different from $(1, 1)$ and $(2, 1)$. This difference is related to the fact that the "secular" part, in the case $(p, q) \ne (1, 1), (2, 1)$, depends only on the action variables, while in the other cases it depends explicitly also on one angle. To overcome this difficulty, we will make use of (detailed, analytic information on) action–angle variables for generalized pendula; even though this subject is classical, we could not find in the literature any suitable reference and we decided, therefore, to include it in appendix. The results presented here were announced in the note [3].

Let us make a few remarks:

- In view of the proper degeneracy, Nekhoroshev theorem [13] does not apply to the D'Alembert model and an *ad hoc* proof is needed. Indeed, the "Nekhoroshev exponent" that we find is better than the one predicted for general systems with two-and-a-half degrees of freedom; compare Theorem 2.1 below with [13], [11] or [14]; this fact is related to the appearance of a "fast" time scale[3]. For Nekhoroshev estimates on a related model, see [2].

[1]The "ecliptic plane" is the (fixed) plane containing the Keplerian ellipse described by the periodic motion of the center of mass of the planet.

[2]The oblateness of the planet is essentially the ratio between the polar and the equatorial radius.

[3]The appearance of different time scales in Celestial Mechanics and its exploitation in perturbation theory is a well known fact going back at least to Lagrange; for more modern implications of the appearance of different time scales in connection with the problems considered here, we refer the reader to, e.g., [12].

- As mentioned above, the main difficulty in the proof of Theorem 1.1 is the appearance (in the cases $(p, q) = (1, 1)$, $(2, 1)$) of separatrices in the (integrable) secular Hamiltonian. To overcome such problem, we use energy conservation arguments in the region close to the separatrices and averaging theory in the region far away from it (after having introduced action–angle variables for the secular Hamiltonian). Clearly, the key technical point consists in proving the overlap of these two regions so as to obtain Nekhoroshev stability in *the whole* phase space.
- We mention that in [7], it was claimed that the planetary D'Alembert model, near the resonance $(p, q) = (2, 1)$, has an instability region where the variable J_2 undergoes a variation of order one (i.e., independent of the perturbative parameters) in finite time, provided ε and $\mu = \varepsilon^c$ (for a suitable $c > 1$) are positive and small enough. The proof of this claim proposed in [7] contained an algebraic error (see the Erratum in [7]) and, even though such error has been corrected ([8]) and several technical progresses, in such direction, have been obtained (see, e.g., [9], [15]), a complete proof of the above claim is still missing.

2. **Exponential stability theorem for the D'Alembert model.** Let us proceed to formulate, in a more precise way, our main result. It is a classical fact (essentially due to Andoyer) that the planetary D'Alembert model near an exact $(p{:}q)$ resonance may be described (in suitable physical units) by a a real–analytic Hamiltonian of the form[4]

$$H_{\varepsilon,\mu} := \frac{I_1^2}{2} + \omega(pI_1 - qI_2 + qI_3) + \varepsilon F_0(I_1, I_2, \varphi_1, \varphi_2) + \varepsilon\mu F_1(I_1, I_2, \varphi_1, \varphi_2, \varphi_3; \mu)\ , \quad (1)$$

where:

- $(I, \varphi) \in A \times \mathbf{T}^3$ are standard symplectic coordinates; the domain $A \subset \mathbf{R}^3$ is given by
$$A := \Big\{ |I_1| < r\varepsilon^\ell\ ,\quad |I_2 - \bar{J}_2| < r\ ,\quad I_3 \in \mathbf{R} \Big\}\ , \quad (2)$$
with $0 < \ell < 1/2$, $r > 0$. In the terminology of the preceding item 1.2, $I_1 := J_1 - \bar{J}_1$, $I_2 := J_2$; $\bar{J}$ being a fixed "reference datum" corresponding to the exact $(p{:}q)$ resonance; p and q are two positive co–prime integers, which identify the spin–orbit resonance (the planet, in the unperturbed regime, revolves q times around the major body and p times around its spin axis): $\bar{J}_1 = p\omega$ and the period of the Keplerian orbit is $2\pi/(q\omega)$, $(\omega > 0)$. The action I_1 measures the displacement from the exact resonance, while I_3 is an artificially introduced variable canonically conjugated to φ_3 (the "mean anomaly"), which is proportional to time.
- $0 \le \varepsilon, \mu < 1$ are – as in the preceding item – two small parameters (measuring, respectively, the oblateness of the planet and the eccentricity of the Keplerian ellipse).
- The functions F_i are real–analytic functions in all their arguments, and may be computed (via Legendre expansions in the eccentricity μ) from the Lagrangian expression of the gravitational (Newtonian) potential; for explicit computations, see, e.g., [7]. While the explicit form of F_1 is not important in the sequel, and, in fact, *our result holds for any function F_1 real–analytic*

[4] Compare, e.g., [7], [5]

and bounded on A, the explicit form of F_0 plays a major rôle in the following analysis. The function F_0 is a trigonometric polynomial given by

$$F_0(I_1, I_2, \varphi_1, \varphi_2) = \sum_{j \in \mathbf{Z},\ |j| \le 2} c_j \cos(j\varphi_1) + d_j \cos(j\varphi_1 + 2\varphi_2) \ , \tag{3}$$

where c_j and d_j are suitable functions of $J = (\bar{J}_1 + I_1, I_2)$ which may be described as follows. Let

$$\kappa_1 := \kappa_1(I_1) := \frac{L}{\bar{J}_1 + I_1} \ , \qquad \kappa_2 := \kappa_2(I_1, I_2) := \frac{I_2}{\bar{J}_1 + I_1} \ ,$$
$$\nu_1 := \nu_1(I_1) := \sqrt{1 - \kappa_1^2} \ , \qquad \nu_2 := \nu_2(I_1, I_2) := \sqrt{1 - \kappa_2^2} \ ;$$

where L is a real parameter (L corresponds to the projection of the angular momentum of the planet onto the polar axis of the planet and, since the planet is rotational, it turns out to be a constant of the motion). *The parameters* $\bar{J}_i$, L *and the constant* r *are assumed to satisfy*

$$L + 3r\varepsilon^\ell < \bar{J}_1 \ , \qquad |\bar{J}_2| + 3r(\varepsilon^\ell + 1) < \bar{J}_1 \ , \tag{4}$$

so that $0 < \kappa_i < 1$ *and the* ν_i*'s are well defined on the domain* A. Then, the functions c_j and d_j are defined by

$$c_0(I_1, I_2) := \frac{1}{4}\Big(2\kappa_1^2\nu_2^2 + \nu_1^2(1 + \kappa_2^2)\Big) \ , \quad d_0(I_1, I_2) := -\frac{\nu_2^2}{4}(2\kappa_1^2 - \nu_1^2) \ ,$$
$$c_{\pm 1}(I_1, I_2) := \frac{\kappa_1\kappa_2\nu_1\nu_2}{2} \ , \quad d_{\pm 1}(I_1, I_2) := \mp\frac{(1 \pm \kappa_2)\kappa_1\nu_1\nu_2}{2} \ ,$$
$$c_{\pm 2}(I_1, I_2) := -\frac{\nu_1^2\nu_2^2}{8} \ , \quad d_{\pm 2}(I_1, I_2) := -\frac{\nu_1^2(1 \pm \kappa_2)^2}{8} \ . \tag{5}$$

With the above positions, for the motions governed by the $(p{:}q)$–resonant D'Alembert Hamiltonian $H_{\varepsilon,\mu}$, there holds the following

Theorem 2.1. *Let* $c > 0$, $0 < \ell < 1/2$ *and* $0 < C_0 < \min\{c, \ell\}$. *Assume that* (4) *holds and that*

$$\nu_1(0) \ne \frac{2}{3} \tag{6}$$

(which is equivalent to $\bar{J}_1 \ne \sqrt{3}L$*). Then, there exist* ε_0, $C_i > 0$ *such that, if* $0 \le \varepsilon \le \varepsilon_0$ *and* $0 \le \mu \le \varepsilon^c$, *then*

$$|I(t) - I(0)| < C_3 r\ \varepsilon^{C_1} \ , \qquad \forall \ \ |t| < T(\varepsilon) := \frac{C_5}{\omega\ \varepsilon^{C_4}}\ \exp\Big(\frac{C_2}{\varepsilon^{C_0}}\Big) \ , \tag{7}$$

where $(I(t), \varphi(t))$ *denotes the* $H_{\varepsilon,\mu}$*–evolution of an initial datum* $(I(0), \varphi(0)) \in A \times \mathbf{T}^3$.

3. **Proof of the theorem.**

3.1. Preliminaries

We shall use the following notations: if $\emptyset \ne D \subset \mathbf{R}^d$ and $\rho := (\rho_1, \rho_2, \ldots, \rho_d)$ with $0 < \rho_j \le \infty$ for $1 \le j \le d$, we denote

$$D_\rho := \{I = (I_1, \ldots, I_d) \in \mathbf{C}^d \ : \ |I_j - \bar{I}_j| < \rho_j, j = 1, \ldots, d, \ \text{ for some } \bar{I} \in D\} \ ;$$

$\mathbf{T}^d_\sigma$ denotes the complex set $\{z \in \mathbf{C}^d : \ |\operatorname{Im} z_j| < \sigma, \ j = 1, \dots, d\}$ (thought of as a complex neighborhood of $\mathbf{T}^d$).

We shall work in the Banach space $\mathcal{H}_{\mathbf{R}}(D_\rho \times \mathbf{T}^d_\sigma)$ of functions f real–analytic on $D_\rho \times \mathbf{T}^d_\sigma$ having finite norm

$$\|f\|_{\rho,\sigma} := \sum_{k\in\mathbf{Z}^d} \sup_{I\in D_\rho} |\hat{f}_k(I)| e^{|k|\sigma} \ ,$$

$\hat{f}_k(I)$ being the Fourier coefficients of the periodic function $\varphi \to f(I,\varphi)$. Notice that, by Liouville Theorem, if $f \in \mathcal{H}_{\mathbf{R}}(D_\rho \times \mathbf{T}^d_\sigma)$ and $\rho_j = \infty$ for some $1 \le j \le d$, then f does not depend on I_j.
We shall use the following standard result from normal form theory; see [14] for the proof with $\rho_1 = \cdots = \rho_d$; (for the simple modifications in the case of different analyticity radii, see [4] or [5]).

Lemma 3.1 (Normal Form Lemma)**.** *Let $H := H(I,\varphi) := h(I) + f(I,\varphi)$ be a real–analytic Hamiltonian on $D \times \mathbf{T}^d$ belonging to $\mathcal{H}_{\mathbf{R}}(D_\rho \times \mathbf{T}^d_\sigma)$ for certain $\rho := (\rho_1, \rho_2, \dots, \rho_d)$ with $0 < \rho_j \le \infty$, $1 \le j \le d$, and $\sigma > 0$. Let $\rho_0 := \min_{1\le j\le d} \rho_j$. Let Λ a sub–lattice of $\mathbf{Z}^d$, $\alpha > 0$, $K \in \mathbf{N}$ with $K\sigma \ge 6$. Suppose that $\forall\, I \in D_\rho$ and $\forall\, k \in \mathbf{Z}^d \setminus \Lambda$, $|k| \le K$, we have $|h'(I) \cdot k| \ge \alpha$ and that the following condition is satisfied:*

$$\|f\|_{\rho,\sigma} =: \eta \le \frac{\alpha\rho_0}{2^{10} K} \ . \tag{8}$$

Then, there exist a real–analytic symplectic transformation

$$\phi : (J,\psi) \in D_{\rho/2} \times \mathbf{T}^d_{\sigma/6} \mapsto (I,\varphi) = \phi(J,\psi) \in D_\rho \times \mathbf{T}^d_\sigma$$

and real–analytic functions $f_* := f_*(J,\psi)$, $g := g(J,\psi) := \sum_{k\in\Lambda} g_k(J) e^{ik\cdot\psi}$ belonging to the space $\mathcal{H}_{\mathbf{R}}(D_{\rho/2} \times \mathbf{T}^d_{\sigma/6})$ such that the following properties hold:

(i) $H \circ \phi(J,\psi) = h(J) + \sum_{k\in\Lambda, |k|\le K} f_k(J) e^{ik\cdot\psi} + g(J,\psi) + f_*(J,\psi)$;

(ii) $\|g\|_{\rho/2,\sigma/6} \le \dfrac{2^{11}}{\alpha\rho_0\sigma}\eta^2 \le \dfrac{2}{K\sigma}\eta \le \dfrac{1}{3}\eta$;

(iii) $\|f_*\|_{\rho/2,\sigma/6} \le \eta e^{-K\sigma/6}$;

(iv) $|I - J| \le \dfrac{2^5}{\alpha\sigma}\eta \le \dfrac{\rho_0}{2^7}$, $\quad |\varphi - \psi| \le \dfrac{2^6}{\alpha\rho_0}\eta \le \dfrac{\sigma}{2^5}$, $\quad \forall\ (J,\psi) \in D_{\rho/2} \times \mathbf{T}^d_{\sigma/6}$.

Remark 3.1. *If $h := h(I) := \hat{h}(I_1, \dots, I_{d-1}) + \omega I_d$ and $f := f(I_1, \dots, I_{d-1}, \varphi)$, then the symplectic transformation ϕ preserves the form of the Hamiltonian and has the form:*

$$\begin{cases} I_j = \tilde{I}_j(J_1, \dots, J_{d-1}, \psi) \ , \quad \varphi_j = \tilde{\varphi}_j(J_1, \dots, J_{d-1}, \psi) \ , & (1 \le j \le d-1) \ , \\ I_d = J_d + \tilde{I}_d(J_1, \dots, J_{d-1}, \psi) \ , \quad \varphi_d = \psi_d \ , & \end{cases}$$

and, also, f_* and g do not depend on J_d.

In what follows, we shall assume that, for any $0 \le \varepsilon \le \bar{\varepsilon}$, the Hamiltonian $H_{\varepsilon,\mu}$ in (1) belongs to $\mathcal{H}_{\mathbf{R}}(A_R \times \mathbf{T}^3_s)$, where

$$R := (r\varepsilon^\ell, r, \infty) \ , \tag{9}$$

$s > 0$ (and $0 < \bar{\varepsilon} < 1$). We shall also denote M_0 and M_1 (ε–independent) upper bounds on, respectively, $\|F_0\|_{R,s}$ and $\|F_1\|_{R,s}$.

3.2. Step 1: linear change of variables

Let ϕ_0 be the following linear symplectic map:

$$\phi_0(I',\varphi') := \Big((I'_1, I'_2, -\frac{p}{q}I'_1 + I'_2 + \frac{1}{q}I'_3), (\varphi'_1 + p\varphi'_3, \varphi'_2 - q\varphi'_3, q\varphi'_3)\Big) \ . \tag{10}$$

Then, ϕ_0 casts the Hamiltonian $H_{\varepsilon,\mu}$ into the form

$$\begin{aligned} H^{(0)}(I',\varphi';\varepsilon,\mu) &:= H_{\varepsilon,\mu} \circ \phi_0(I',\varphi') \\ &:= \tfrac{I'^2_1}{2} + \omega I'_3 + \varepsilon G_0(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3) + \varepsilon\mu G_1(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3; \mu) \ , \end{aligned} \tag{11}$$

which belongs to $\mathcal{H}_{\mathbf{R}}(A_R \times \mathbf{T}_S)$ with

$$S := c's \ , \qquad c' := \min\{1/(1+p), 1/(1+q)\} \ , \tag{12}$$

and

$$\|G_0\|_{R,S} \le M_0 \ , \qquad \|G_1\|_{R,S} \le M_1 \ .$$

Moreover:

$$G_0(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3) := H_{01}(I'_1, I'_2, \varphi'_1) + \widetilde{G}_0(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3)$$

with

$$\int_0^{2\pi} \widetilde{G}_0(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3) d\varphi'_3 = 0 \ ,$$

and

$$H_{01}(I'_1, I'_2, \varphi'_1) := \begin{cases} c_0(I'_1, I'_2) \ , & \text{if } (p,q) \ne (1,1), (2,1) \ , \\ c_0(I'_1, I'_2) + d_{j_p}(I'_1, I'_2)\cos(j_p\varphi'_1) \ , & \text{if } (p,q) = (1,1), (2,1) \ , \end{cases}$$

with $j_1 := 2$ and $j_2 := 1$.

Obviously, since ϕ_0 depends upon p and q, also the functions G_i and H_{01} depend upon p and q, but *we shall not indicate such dependence in the notation*.
We remark that, in general, ϕ_0 is *not* a diffeomorphism of $\mathbf{R}^3 \times \mathbf{T}^3$ (since the induced map on $\mathbf{T}^3$ has determinant equal to q); this fact, however, does not affect the following analysis.

If

$$a := 1 + \min\{c, \ell\} \ , \tag{13}$$

using the fact that $|I'_1| < 2r\varepsilon^\ell$ and $\mu \le \varepsilon^c$, one see that $H^{(0)}$ has the following form:

$$\frac{I'^2_1}{2} + \omega I'_3 + \varepsilon \overline{H}_{01}(I'_2, \varphi'_1) + \varepsilon \overline{G}_0(I'_2, \varphi'_1, \varphi'_2, \varphi'_3) + \varepsilon^a H_2^{(0)}(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3; \varepsilon)$$

where $\overline{G}_0(I'_2, \varphi) := \widetilde{G}_0(0, I'_2, \varphi')$,

$$\int_0^{2\pi} \overline{G}_0(I'_2, \varphi'_1, \varphi'_2, \varphi'_3) d\varphi'_3 = 0 \ , \tag{14}$$

and (recall (5)), $\overline{H}_{01}(I'_2,\varphi'_1) := H_{01}(0,I'_2,\varphi'_1)$ is equal to

$$\overline{H}_{01}(I'_2,\varphi'_1) := \begin{cases} \overline{c}_0(I'_2)\ , & \text{if } (p,q)\neq(1,1),(2,1)\ ,\\ \overline{c}_0(I'_2)+\overline{d}_{j_p}(I'_2)\cos(j_p\varphi'_1)\ , & \text{if } (p,q)=(1,1),(2,1)\ , \end{cases} \tag{15}$$

where

$$\overline{c}_0(I'_2) := c_{00}+c_{02}\frac{{I'}_2^2}{2},\quad c_{00} := \frac{1}{4}\Big(2-\overline{\nu}_1^2\Big),\quad c_{02} := \frac{1}{\overline{J}_1^2}\Big(\frac{3}{2}\overline{\nu}_1^2-1\Big) \tag{16}$$

and

$$\begin{aligned} \overline{d}_1(I'_2) &:= \overline{d}_1(0,I'_2) := -\frac{1}{2}\overline{\kappa}_1\overline{\nu}_1\sqrt{1-\frac{{I'}_2^2}{\overline{J}_1^2}}\Big(1+\frac{{I'}_2}{\overline{J}_1}\Big)\ ,\\ \overline{d}_2(I'_2) &:= \overline{d}_2(0,I'_2) := -\frac{1}{8}\overline{\nu}_1^2\Big(1+\frac{{I'}_2}{\overline{J}_1}\Big)^2\ , \end{aligned} \tag{17}$$

where

$$\overline{\kappa}_1 := \kappa_1(0) := \frac{L}{\overline{J}_1},\quad \overline{\nu}_1 := \nu_1(0) := \sqrt{1-\overline{\kappa}_1^2}\ . \tag{18}$$

The function $H_2^{(0)}(I'_1,I'_2,\varphi'_1,\varphi'_2,\varphi'_3;\varepsilon)$ belongs to $\mathcal{H}_{\mathbf{R}}(A_R\times\mathbf{T}^3_S)$ and

$$\|\varepsilon\overline{H}_{01}+\varepsilon\overline{G}_0+\varepsilon^a H_2^{(0)}\|_{R,S} = \|\varepsilon G_0+\varepsilon^{1+c}G_1\|_{R,S}\le\varepsilon(M_0+\varepsilon^c M_1)\ .$$

3.3. Step 2: time averaging

Here, we shall remove, up to exponentially small terms, the (fast) dependence upon φ'_3. To do this, we shall apply the Normal Form Lemma with $d := 3$, $(I,\varphi) := (I',\varphi')$, $H := H^{(0)}$, $h := {I'}_1^2/2+\omega I'_3$, $f := \varepsilon\overline{H}_{01}+\varepsilon\overline{G}_0+\varepsilon^a H_2^{(0)} = \varepsilon G_0+\varepsilon^{1+c}G_1$, $D := A$, $\rho := R$, $\rho_0 := r\varepsilon^\ell$, $\sigma := S$, $\Lambda := \{(k_1,k_2,0)$ s.t. $k_1,k_2\in\mathbf{Z}\}$, $\alpha := \omega/2$, $K := \omega/(4r\varepsilon^\ell)$. The condition $K\sigma\ge 6$ is implied by

$$\varepsilon\le(\omega S/24r)^{1/\ell}\ . \tag{19}$$

Condition (8) becomes

$$\varepsilon(M_0+\varepsilon^c M_1)\le 2^{-9}r^2\varepsilon^{2\ell}$$

which is verified, for example, if

$$\varepsilon\le\Big(\frac{r^2}{2^9(M_0+M_1)}\Big)^{1/(1-2\ell)}. \tag{20}$$

Hence, for ε small enough, we can apply the Normal Form Lemma, finding a real–analytic symplectic transformation

$$\phi_1 : (\hat I,\hat\varphi)\in A_{R/2}\times\mathbf{T}^3_{S/6}\mapsto(I',\varphi')\in A_R\times\mathbf{T}^3_S\ ,$$

such that

$$\begin{aligned} H^{(1)}(\hat I,\hat\varphi;\varepsilon,\mu) &:= H^{(0)}\circ\phi_1(\hat I,\hat\varphi;\varepsilon) \\ &:= \frac{\hat I_1^2}{2}+\omega\hat I_3+\varepsilon\overline{H}_{01}(\hat I_2,\hat\varphi_1)+\varepsilon^a H_1^{(1)}(\hat I_1,\hat I_2,\hat\varphi_1,\hat\varphi_2;\varepsilon)+H_*^{(1)}(\hat I_1,\hat I_2,\hat\varphi;\varepsilon)\ , \end{aligned} \tag{21}$$

and such that the following bounds hold. For any $(\hat I,\hat\varphi)\in A_{R/2}\times\mathbf{T}^3_{S/6}$,

$$|I'-\hat I|\le\frac{2^6}{\omega S}\varepsilon(M_0+\varepsilon^c M_1)\le\Big(\frac{r}{8S\omega}\Big)r\varepsilon^{2\ell}\le\frac{1}{2^7}r\varepsilon^\ell\ , \tag{22}$$

and

$$\|H_1^{(1)}\|_{R/2,S/6} \le M_1^{(1)},$$
$$\|H_*^{(1)}\|_{R/2,S/6} \le M_*^{(1)} := \varepsilon(M_0 + \varepsilon^c M_1)\exp(-\,\mathfrak{c}_1\varepsilon^{-\ell}) \ , \tag{23}$$

where

$$M_1^{(1)} := \Big(M_1 + \frac{8r}{\omega S}(M_0 + M_1)\Big), \qquad \mathfrak{c}_1 := \frac{\omega S}{24r} \ . \tag{24}$$

3.4. Step 3: averaging over $\hat\varphi_1$ (case $(p,q) \ne (1,1)$, $(2,1)$)

Let us assume, first, that (p,q) is different from $(1,1)$ and from $(2,1)$. Then the Hamiltonian $\overline{H}_{01}(\hat I_2, \hat\varphi_1)$ is independent of the angles, allowing to treat the angle $\hat\varphi_1$ as a "fast" angle in a suitable domain $\hat A$. Consider, therefore, the Hamiltonian

$$\hat H^{(1)}(\hat I_1, \hat I_2, \hat\varphi_1, \hat\varphi_2; \varepsilon) := \frac{\hat I_1^2}{2} + \varepsilon \overline{c}_0(\hat I_2) + \varepsilon^a H_1^{(1)}(\hat I_1, \hat I_2, \hat\varphi_1, \hat\varphi_2; \varepsilon) \ , \tag{25}$$

and let

$$\hat A := \{\hat I_1 \in (-\frac54 r\varepsilon^\ell, -\frac12 r\varepsilon^b) \cup (\frac12 r\varepsilon^b, \frac54 r\varepsilon^\ell), \ \ |\hat I_2 - \bar J_2| < \frac54 r\} \ . \tag{26}$$

In order to apply the Normal Form Lemma, we let C_0 be as in Theorem 2.1, a as in (13) and fix a number b so that

$$\frac12 < b \le \frac{a - C_0}{2} \ . \tag{27}$$

We, also, let: $d := 2$, $(I,\varphi) := (\hat I_1, \hat I_2, \hat\varphi_1, \hat\varphi_2)$, $h := \hat I_1^2/2 + \varepsilon\overline{c}_0(\hat I_2)$, $f := \varepsilon^a H_1^{(1)}$, $D := \hat A$, $\rho := (r\varepsilon^b/4, r/4)$, $\rho_0 := r\varepsilon^b/4$, $\sigma := S/6$, $\Lambda := \{(0,k_2) \text{ s.t. } k_2 \in \mathbf{Z}\}$, $\alpha := r\varepsilon^b/8$, $K := r^2(2^{15}M_1^{(1)}\varepsilon^{C_0})^{-1}$. With such positions, we see that we can apply Lemma 3.1, provided

$$\varepsilon \le \min\Big\{\Big[\frac{2^{12}M_1^{(1)}}{|c_{02}|r(\bar J_2 + 2r)}\Big]^{1/(1-b-C_0)}, \quad \mathfrak{c}_2^{1/C_0}\Big\} \ , \qquad \mathfrak{c}_2 := \frac{r^2 S}{9\cdot 2^{17}M_1^{(1)}} \ . \tag{28}$$

Under such condition, we can find a real–analytic symplectic transformation

$$\hat\phi_2 : (\tilde I_1, \tilde I_2, \tilde\varphi_1, \tilde\varphi_2) \in \hat A_{(r\varepsilon^b/8, r/8)} \times \mathbf{T}^2_{S/36} \mapsto (\hat I_1, \hat I_2, \hat\varphi_1, \hat\varphi_2) \in \hat A_{(r\varepsilon^b/4, r/4)} \times \mathbf{T}^2_{S/6}$$

such that

$$\begin{aligned}\hat H^{(2)}(\tilde I_1, \tilde I_2, \tilde\varphi_1, \tilde\varphi_2; \varepsilon) &:= H^{(1)} \circ \hat\phi_2(\tilde I_1, \tilde I_2, \tilde\varphi_1, \tilde\varphi_2; \varepsilon) \\ &:= \frac{\tilde I_1^2}{2} + \varepsilon\overline{c}_0(\tilde I_2) + \varepsilon^a H_1^{(2)}(\tilde I_1, \tilde I_2, \tilde\varphi_2; \varepsilon) + \hat H_*^{(2)}(\tilde I_1, \tilde I_2, \tilde\varphi_1, \tilde\varphi_2; \varepsilon)\end{aligned} \tag{29}$$

with

$$\begin{aligned}\|H_1^{(2)}\|_{(r\varepsilon^b/8, r/8), S/36} &\le M_1^{(2)} := \frac43 M_1^{(1)}, \\ \|\hat H_*^{(2)}\|_{(r\varepsilon^b/8, r/8), S/36} &\le \hat M_*^{(2)} := \varepsilon^a M_1^{(1)}\exp(-\,\mathfrak{c}_2\varepsilon^{-C_0}),\end{aligned}$$

and, for any $(\tilde I_1, \tilde I_2, \tilde\varphi_1, \tilde\varphi_2) \in \hat A_{(r\varepsilon^b/8, r/8)} \times \mathbf{T}^2_{S/36}$,

$$|\hat I_1 - \tilde I_1|, |\hat I_2 - \tilde I_2| \le \frac{3\cdot 2^9 M_1^{(1)}}{Sr}\varepsilon^{a-b} \le \frac{r\varepsilon^b}{2^9} \ . \tag{30}$$

Extend such symplectic transformation on $\hat{A}_{(r\varepsilon^b/8,r/8)}\times \mathbf{C}\times \mathbf{T}^3_{S/36}$ by setting

$$\phi_2(\tilde I_1,\tilde I_2,\tilde I_3,\tilde\varphi_1,\tilde\varphi_2,\tilde\varphi_3;\varepsilon):=(\hat\phi_2(\tilde I_1,\tilde I_2,\tilde\varphi_1,\tilde\varphi_2;\varepsilon),\tilde I_3,\tilde\varphi_3)\ .$$

In this way, denoting $(\tilde I,\tilde\varphi)=(\tilde I_1,\tilde I_2,\tilde I_3,\tilde\varphi_1,\tilde\varphi_2,\tilde\varphi_3)$, we see that

$$\begin{aligned} H^{(2)}(\tilde I,\tilde\varphi;\varepsilon) &:= H^{(1)}\circ\phi_2(\tilde I,\tilde\varphi;\varepsilon) \\ &:= \frac{\tilde I_1^2}{2}+\omega\tilde I_3+\varepsilon\overline{c}_0(\tilde I_2)+\varepsilon^a H_1^{(2)}(\tilde I_1,\tilde I_2,\tilde\varphi_2;\varepsilon)+H_*^{(2)}(\tilde I_1,\tilde I_2,\tilde\varphi;\varepsilon)\end{aligned} \tag{31}$$

with

$$\|H_*^{(2)}\|_{(r\varepsilon^b/8,r/8,\infty),S/36}\le M_*^{(2)}:=\hat M_*^{(2)}+M_*^{(1)}.$$

In order to simplify the calculus of the constants we assume that

$$\varepsilon\le\min\Big\{\Big(\frac{\mathfrak{c}_1}{2\,\mathfrak{c}_2}\Big)^{1/(\ell-C_0)},(\ell\,\mathfrak{c}_1)^{2/\ell}\Big\}. \tag{32}$$

Using (32) it is simple to prove that

$$\varepsilon^{-1}\exp(\,\mathfrak{c}_1\varepsilon^{-\ell})\ge\varepsilon^{-a}\exp(\,\mathfrak{c}_2\varepsilon^{-C_0}). \tag{33}$$

In fact, using (32) and the fact that $a<3/2$, it is sufficient to prove that $\exp(\,\mathfrak{c}_1\varepsilon^{-\ell})$ $\ge\varepsilon^{-1}$ which is guaranteed[5] again by (32).
By (33) we, also, obtain

$$M_*^{(2)}\le\overline{M}_2^{(2)}\varepsilon^a\exp(-\,\mathfrak{c}_2\varepsilon^{-C_0}),\qquad \overline{M}_2^{(2)}:=\Big(M_1+\Big(1+\frac{8r}{\omega S}\Big)(M_0+M_1)\Big). \tag{34}$$

3.5. Step 4: conclusion of proof (case $(p,q)\neq(1,1),(2,1)$)

We are, now, in the position of concluding the proof of Theorem 2.1 for the case $(p,q)\neq(1,1)$, $(2,1)$. The arguments we shall use, here, are based on energy conservation. However, such arguments, are not completely straightforward because we have to keep track of domains (recall that the variables $(\tilde I_1,\tilde I_2)$ are not defined in a neighborhood of the origin) and also because we shall freely use different sets of variables.

- *Energy conservation for the Hamiltonians $H^{(1)}$ and $H^{(2)}$*

Denote by $\hat z(t):=(\hat I(t),\hat\varphi(t))$ and $\tilde z(t):=(\tilde I(t),\tilde\varphi(t))$ the solutions of the Hamilton equations associated, respectively, to the Hamiltonians $H^{(1)}$ in (21) and $H^{(2)}$ in (31), with respective initial data $\hat z(0):=(\hat I(0),\hat\varphi(0))$ and $\tilde z(0):=(\tilde I(0),\tilde\varphi(0))$. Furthermore, if $F=F(\tilde I,\tilde\varphi)$, denote $\hat\Delta_t F:=F(\hat z(t))-F(\hat z(0))$ and $\tilde\Delta_t F:=F(\tilde z(t))-F(\tilde z(0))$. Then, conservation of energy for the Hamiltonians in (21) and (31) yields[6]:

$$\begin{aligned}\varepsilon c_{02}\Big[\hat I_2(0)\hat\Delta_t\hat I_2+\frac12(\hat\Delta_t\hat I_2)^2\Big]+\Big[\hat I_1(0)\hat\Delta_t\hat I_1+\frac12(\hat\Delta_t\hat I_1)^2\Big]&\\ +\omega\hat\Delta_t\hat I_3+\varepsilon^a\hat\Delta_t H_1^{(1)}+\hat\Delta_t H_*^{(1)}=0\ ,&\end{aligned} \tag{35}$$

[5] Setting $x:=\varepsilon^{-\ell}$ and $y:=1/\ell\,\mathfrak{c}_1$ we have to prove that $e^x\ge x^y$. This is obvious if $y\le1$; if $y>1$ it is true if, for example, $x\ge y^2$.

[6] Recall (15) and observe that for any numbers x,y, one has $\frac{x^2}{2}-\frac{y^2}{2}=\frac12(x-y)^2+y(x-y)$.

and

$$\varepsilon c_{02}\Big[\tilde I_2(0)\tilde\Delta_t\tilde I_2+\frac12(\tilde\Delta_t\tilde I_2)^2\Big]+\Big[\tilde I_1(0)\tilde\Delta_t\tilde I_1+\frac12(\tilde\Delta_t\tilde I_1)^2\Big]$$
$$+\omega\tilde\Delta_t\tilde I_3+\varepsilon^a\tilde\Delta_t H_1^{(2)}+\tilde\Delta_t H_*^{(2)}=0. \tag{36}$$

- *A–priori exponential estimates for the drift of $\tilde I_1$, $\hat I_3$ and $\tilde I_3$*

For[7] $0\le t\le T(\varepsilon)$ we have directly by Hamilton equations, (33) and Cauchy estimates[8]

$$|\tilde\Delta_t\tilde I_3| \;=|\hat\Delta_t\hat I_3|\le\sup_{0\le\tau\le t}|\partial_{\hat\varphi_3}H_2^{(1)}(\hat z(\tau);\varepsilon)|t\le\frac{6}{eS}M_2^{(1)}t\le\frac{M_0+M_1}{96\overline{M}_2^{(2)}}r\varepsilon^{2b}, \tag{37}$$

$$|\tilde\Delta_t\tilde I_1| \;\le\sup_{0\le\tau\le t}|\partial_{\tilde\varphi_1}H_2^{(2)}(\tilde z(\tau);\varepsilon)|t\le\frac{36}{eS}M_2^{(2)}t\le\frac{1}{16}r\varepsilon^{2b}. \tag{38}$$

Consider, now, (real) initial positions

$$(\hat I_1(0),\hat I_2(0))\in\{|\hat I_1|\le(1+2^{-7})r\varepsilon^\ell\}\times\{|\hat I_2-\bar J_2|\le(1+2^{-7})r\}\ ,$$

and let us consider, separately, two cases:

(i) $|\hat I_1(t)|<\frac12 r\varepsilon^b, \qquad \forall\, 0\le t\le T(\varepsilon)$;

(ii) $\exists\, 0\le t^*<T(\varepsilon)$ s.t. $|\hat I_1(t)|<\frac12 r\varepsilon^b \quad \forall\, 0\le t<t^*$ and $|\hat I_1(t^*)|\ge\frac12 r\varepsilon^b$.

- *Case (i) and stability of $\hat I_2$*

Consider case (i): by (37) and (35), we see that, until

$$|\hat I_2(t)-\bar J_2|\le 3r/2, \tag{39}$$

we have

$$\Big|2\hat I_2(0)\hat\Delta_t\hat I_2+(\hat\Delta_t\hat I_2)^2\Big|\le\ \mathfrak{c}_3 r^2\varepsilon^{2b-1} \tag{40}$$

where we can take[9]

$$\mathfrak{c}_3:=\frac{1}{|c_{02}|}\left(\frac14+\frac{M_0+M_1}{48\overline{M}_2^{(2)}}\frac{\omega}{r}+\frac{16}{3}\frac{M_1^{(1)}}{r^2}+4\frac{\overline{M}_2^{(2)}}{r^2}\right)\ . \tag{41}$$

We need, at this point, an elementary estimate (whose trivial proof is left to the reader):

Lemma 3.2. *Let $y,y_0\in\mathbf{R}$ and $C>0$ and suppose that*

$$|2y_0y+y^2|\le C^2\ .$$

Then:

(1) *if* $|y_0|\le C$ *then* $|y|\le|y_0|+\sqrt{y_0^2+C^2}\le(1+\sqrt2)C$,

---

[7] We shall consider only positive times since negative times are treated in a completely analogous way.

[8] "Cauchy estimates" allow to bound derivatives of analytic functions in terms of their sup–norm on larger domains; with our choice of norms, Cauchy estimates take the following form. Consider a 2π–periodic function $f(\varphi):=\sum_{k\in\mathbf{Z}}f_ke^{ik\cdot\varphi}$, analytic on $\mathbf{T}_s$ with $\|f\|_s:=\sum_k|f_k|e^{|k|s}$, then $\max_{\varphi\in\mathbf{T}}|\partial_\varphi f(\varphi)|\le\frac{1}{es}\|f\|_s$. In fact for all $0<\sigma<s$ we have $\max_{\varphi\in\mathbf{T}}|\partial_\varphi f(\varphi)|\le\|\partial_\varphi f\|_{s-\sigma}=\sum_k\ |k|e^{-|k|\sigma}\ |f_k|e^{|k|s}\le\frac{1}{e\sigma}\sum_k|f_k|e^{|k|s}$ and taking the sup over $\sigma<s$ of the right hand side, we have the thesis.

[9] Recall (6), which implies $c_{02}\neq0$.

(2) *if* $|y_0| > C$ *then*[10] $|y| \le C^2 |y_0|^{-1} \le C$.

Let us now assume that

$$\varepsilon \le \Big(\frac{1}{20\sqrt{\mathfrak{c}_3}}\Big)^{2/(2b-1)} \tag{42}$$

and let us apply the estimates of Lemma 3.2 to (40) with $C := \sqrt{\mathfrak{c}_3} r\varepsilon^{C_1}$, $y_0 := \hat{I}_2(0)$ and $y := \hat{\Delta}\hat{I}_2$. Then:

$$|\hat{I}_2(0)| \le \sqrt{\mathfrak{c}_3} r\varepsilon^{C_1} \implies |\hat{I}_2(t) - \hat{I}_2(0)| \le (1+\sqrt{2})\sqrt{\mathfrak{c}_3} r\varepsilon^{C_1} \le \Big(\frac{1}{8} - \frac{1}{2^8}\Big) r\,, \tag{43}$$

$$|\hat{I}_2(0)| > \sqrt{\mathfrak{c}_3} r\varepsilon^{C_1} \implies |\hat{I}_2(t) - \hat{I}_2(0)| \le \frac{\mathfrak{c}_3 r^2}{|\hat{I}_2(0)|}\varepsilon^{2b-1} \le \Big(\frac{1}{8} - \frac{1}{2^8}\Big) r\,, \tag{44}$$

which in particular imply (39).

- *Case (ii) and stability of* $\hat{I}_1$

If (ii) occurs, then, by (44), we have that

$$(1+2^{-7}) \ge |\hat{I}_1(t^*)| \ge r\varepsilon^b/2\,, \qquad |\hat{I}_2(t^*) - \bar{J}_2| \le 5r/4\,.$$

Then, by (30), we can find

$$(\tilde{I}_1^*, \tilde{I}_2^*, \tilde{\varphi}_1^*, \tilde{\varphi}_2^*) \in \hat{A}_{(\frac{r\varepsilon^b}{2^9}, \frac{r}{2^9})} \times \mathbf{T}^2$$

such that

$$\hat{\phi}(\tilde{I}_1^*, \tilde{I}_2^*, \tilde{\varphi}_1^*, \tilde{\varphi}_2^*) = (\hat{I}_1(t^*), \hat{I}_2(t^*), \hat{\varphi}_1(t^*), \hat{\varphi}_2(t^*))\,.$$

Now, as in (38), we have

$$|\tilde{I}_1(t) - \tilde{I}_1(t^*)| \le r\varepsilon^{2b}/16 \tag{45}$$

hence, using (30),

$$\begin{aligned} |\hat{I}_1(t) - \hat{I}_1(0)| &\le |\hat{I}_1(t) - \tilde{I}_1(t)| + |\tilde{I}_1(t) - \tilde{I}_1(t^*)| \\ &\quad + |\tilde{I}_1(t^*) - \tilde{I}_1(0)| + |\tilde{I}_1(0) - \hat{I}_1(0)| \\ &\le (1 + \frac{1}{16} + \frac{1}{2^8}) r\varepsilon^b \\ |\hat{I}_1(t)| &\le (1 + \frac{1}{16} + \frac{1}{2^6}) r\varepsilon^\ell. \end{aligned}$$

Finally, using (22), we obtain

$$|I_1(t) - I_1(0)| \le \frac{5}{4} r\varepsilon^b \qquad \text{and} \qquad |I_1(t)| \le \frac{5}{4} r\varepsilon^\ell\,, \tag{46}$$

provided

$$\varepsilon \le \Big(\frac{r}{2^{10}\omega S(M_0 + M_1)}\Big)^{1/1-b}\,. \tag{47}$$

- *Stability of* $\tilde{I}_2$

In order to prove stability for the I_2-variable, we can apply (36) until

$$|\tilde{I}_1(t) - \tilde{I}_1^*| \le \Big(\frac{1}{8} - \frac{1}{2^9}\Big) r\varepsilon^b \qquad \text{and} \qquad |\tilde{I}_2(t) - \tilde{I}_2^*| \le \Big(\frac{1}{8} - \frac{1}{2^9}\Big) r \tag{48}$$

[10] We set $x := C^2 y_0^{-2}$ and we have used that $\sqrt{1+x} - 1 \le x/2$ and $1 - \sqrt{1-x} \le x$ for $0 \le x \le 1$.

obtaining again (as for (40))

$$\left|2\tilde{I}_2^*(\tilde{I}_2(t)-\tilde{I}_2^*)+(\tilde{I}_2(t)-\tilde{I}_2^*)^2\right| \le \mathfrak{c}_3 r^2\varepsilon^{2b-1}. \tag{49}$$

We prove the first inequality in (48) using (30) and (45). As in case (i) we use Lemma 3.2 with $C := \sqrt{\mathfrak{c}_3}r\varepsilon^{C_1}$, $y_0 := \tilde{I}_2^*$ and $y := \tilde{I}_2(t) - \tilde{I}_2^*$. Using again (42) we have that

$$|\tilde{I}_2^*| \le \sqrt{\mathfrak{c}_3}r\varepsilon^{C_1} \implies |\tilde{I}_2(t)-\tilde{I}_2^*| \le (1+\sqrt{2})\sqrt{\mathfrak{c}_3}r\varepsilon^{C_1} \le \Big(\frac{1}{8}-\frac{1}{2^8}\Big)r\ , \tag{50}$$

$$|\tilde{I}_2^*| > \sqrt{\mathfrak{c}_3}r\varepsilon^{C_1} \implies |\tilde{I}_2(t)-\tilde{I}_2^*| \le \frac{\mathfrak{c}_3 r^2}{|\tilde{I}_2^*|}\varepsilon^{2b-1} \le \Big(\frac{1}{8}-\frac{1}{2^8}\Big)r\ , \tag{51}$$

which, in particular, imply the second condition in (48).

- *Conclusion*

Finally, if we define $C_1 := (2b-1)/2$, $C_6 := \sqrt{\mathfrak{c}_3}+2^{-6}$, (with $\mathfrak{c}_3$ is defined in (41)), then, by (43), (44), (50), (51), (22) and (30) we obtain

$$|I_2(0)| \le C_6 r\varepsilon^{C_1} \implies |I_2(t)-I_2(0)| \le (1+\sqrt{2})C_6 r\varepsilon^{C_1} \le \frac{r}{8} \tag{52}$$

$$|I_2(0)| > C_6 r\varepsilon^{C_1} \implies |I_2(t)-I_2(0)| \le \frac{C_6^2 r^2}{|\hat{I}_2(0)|}\varepsilon^{2b-1} \le \frac{r}{8}. \tag{53}$$

The proof of Theorem 2.1 is concluded in the case $(p,q) \ne (1,1),(2,1)$.

3.6. The case $(p,q) = (1,1)$ or $(p,q) = (2,1)$

We, now, turn to the case $(p,q) = (1,1)$ or $(2,1)$. In such a case, the Hamiltonian (21) has the form

$$H^{(1)} = \frac{\hat{I}_1^2}{2} + \omega\hat{I}_3 - \varepsilon k_p(\hat{I}_2)(1+\cos j_p\hat{\varphi}_1) + \varepsilon h_p(\hat{I}_2) + \varepsilon^a H_1^{(1)} + H_*^{(1)}, \tag{54}$$

where

$$k_p(\hat{I}_2) := -\overline{d}_{j_p}(\hat{I}_2), \quad h_p(\hat{I}_2) := \overline{c}_0(\hat{I}_2) + k_p(\hat{I}_2), \quad j_1 = 2, \quad j_2 = 1. \tag{55}$$

In this subsection ξ_i will denote positive (ε–independent) constants and we will take ε as small as we need.
Choose $1 < \lambda \le a - C_0$ (here λ corresponds to $2b$). From (23) we deduce that

$$|\hat{I}_3(t) - \hat{I}_3(t_0)| \le \xi_1\varepsilon^\lambda\ , \qquad \forall\, 0 \le t_0 \le t \le T_1(\varepsilon) := \xi_3\exp(-\xi_2/\varepsilon^\ell). \tag{56}$$

In order to prove stability in the other actions we state the following elementary Lemma concerning the conservation of energy.

Lemma 3.3. *Let $H := H(I,t;\mu) := h(I) + \mu f(t) \in \mathbf{R}$, $I,t,\mu \in \mathbf{R}$ and assume that h is analytic and not identically constant and that $|f(t)| \le 1$ for all $t \in \mathbf{R}$. Fix $r_0 > 0$. Then, there exist $0 < \mu_0, v \le 1$ and $c > 0$ such that, if for some continuous function $I(t) := I(t;\mu)$ with $|I_0| := |I(0)| \le r_0$ $H(I(t),t;\mu) \equiv 0$, then for all $0 \le \mu \le \mu_0$ we have $|I(t) - I_0| \le c\mu^v$.*

Proof. Being h analytic we have that, if $N := \{|I| \le 2r_0 \text{ s.t. } h'(I) = 0\}$, then $\#N < \infty$. Hence, there exists $p_* \in \mathbf{N}$ such that $\forall I_0 \in N$ there exist $1 < p_0 \le p_*$ for which[11] $h^{(p_0)}(I_0) \ne 0$ and $h^{(p)}(I_0) = 0\ \forall\, 1 \le p \le p_0$.

[11] We denote with $h^{(p)}$ the p-th derivative of h with respect to I.

There exist $b_0 > 0$ and $0 \le r_0' \le r_0$ such that $\forall I_0 \in N$, $|I_0| \le r_0$ we have that $|h^{(p_0)}(\tilde I)| \ge b_0$, $\forall |\tilde I - I_0| \le r_0'$. We claim that the Lemma holds with $v \le 1/p_*$, $c \le (2p_0!/b_0)^{1/p_0}$ and $\mu_0 \le (r_0'/c)^{p_*}$. In fact, by Taylor's formula, $\forall |I - I_0| \le r_0'$, $\exists |I_* - I_0| \le r_0'$ such that $h(I) - h(I_0) = h^{(p_0)}(I_*)(I - I_0)^{p_0}/p_0!$ and hence $|I(t) - I_0|^{p_0} \le 2p_0!\mu/b_0$.
On the other hand, if $|I_0| \le r_0$ with $I_0 \in \{|I| \le 2r_0,\ \text{s.t. } |I - I_1| \ge r_0'/2, \forall I_1 \in N\} =: M$, then, defining $m := \min_M |h'| > 0$, the Lemma holds with $v \le 1$, $c \le 2/m$, $\mu_0 \le r_0' m/4$, since $2\mu \ge |h(I) - h(I_0)| \ge m|I - I_0|$. □

We consider first the case $(p, q) = (2, 1)$; the analogous case $(p, q) = (1, 1)$ will be considered later. For brevity we will omit the dependence on $p = 2$ in the formulas. In the Hamiltonian (54) we analyze first the following part, which represents a pendulum with a small gravity depending on a parameter:

$$E := E(\hat I_1, \hat I_2, \hat\varphi_1) := E(\hat I_1, \hat I_2, \hat\varphi_1; \varepsilon) := \frac{\hat I_1^2}{2} - \varepsilon k(\hat I_2)(1 + \cos\hat\varphi_1)\ , \qquad (57)$$

where $k(\hat I_2) = k_2(\hat I_2)$. We denote $E(t) := E(\hat I_1(t), \hat I_2(t), \hat\varphi_1(t))$.
We claim that if $0 \le t_0 \le t \le T_1(\varepsilon)$ then

$$|E(t) - E(t_0)| \le 4\varepsilon^\lambda \implies |\hat I_1(t) - \hat I_1(t_0)| \le \xi_4\sqrt{\varepsilon}\ , \quad |\hat I_2(t) - \hat I_2(t_0)| \le \varepsilon^{\xi_5}\ . \quad (58)$$

In fact, if $|E(t) - E(t_0)| \le 4\varepsilon^\lambda$ then the variable $\hat I_1$ may vary, at most, by order $\sqrt{\varepsilon}$ and using (56), we can apply the energy conservation to the Hamiltonian (54) obtaining that

$$|h(\hat I_2(t)) - h(\hat I_2(t_0))| \le \xi_7 \varepsilon^{\lambda - 1}.$$

Hence, by the fact that $h(\cdot)$ is a non constant analytic function (as it is immediate to verify), using Lemma 3.3 we get (58).

Since the Hamiltonian $\overline{H}_{01}(\hat I_2, \hat\varphi_1)$ depends explicitly on $\hat\varphi_1$, in order to carry out the analogous of step 3 above (§ 3.4), we have, first, to introduce action–angle variables for the two–dimensional integrable system $\frac{\hat I_1^2}{2} + \varepsilon\overline{H}_{01}(\hat I_2, \hat\varphi_1)$, which may be viewed as a "suspended pendulum" (with potential $\cos\hat\varphi_1$ or $\cos 2\hat\varphi_1$) having a small gravity varying with a second action–variable. The results we need are contained in the following proposition, the proof of which is deferred to the appendix[12].

Proposition 3.1. *Let $k(I)$ real–analytic on $\Delta^0 := [\alpha, \beta]$ for $\alpha < \beta$ with analytic extension on $\Delta^0_{r_2}$ for $r_2 > 0$. Let*

$$\bar k := \max_{\bar I_1 \in \Delta^0} k(\bar I_1), \quad \hat k := \min_{\bar I_1 \in \Delta^0} k(\bar I_1), \quad \bar k' := \max_{I \in \Delta^0_{r_2}} |k'(I)|,$$

and suppose that $\hat k > 0$. Let $\varepsilon > 0$, $\eta > 0$, $R_0 \ge 2r_1 > 0$, $0 < s_1 \le 1$, $s_2 > 0$, $D^0 := [-R_0, R_0]$,

$$E(p, I, q) := p^2/2 - \varepsilon k(I)(1 + \cos q)$$

and[13]

$$\begin{aligned}\mathcal{M}^+ &:= \Big\{(p, I, q, \varphi) \in [0, R_0] \times \Delta^0 \times \mathbf{T}^2 :\ \eta \le E(p, I, q) \le R_0^2/2\Big\} \\ \mathcal{M}^- &:= \Big\{(p, I, q, \varphi) \in D^0 \times \Delta^0 \times \mathbf{T}^2 :\ -2\varepsilon k(I) + \eta \le E(p, I, q) \le -\eta\Big\}\ .\end{aligned}$$

---

[12] For the simpler case $k \equiv 1$, see Lemma 2.1 and Appendix B of [4].
[13] An analogous statement holds for $p \in [-R_0, 0]$. Of course, we are assuming that the various parameters are chosen so that $\mathcal{M}^\pm \ne \emptyset$.

Then, there exist positive constants c_1, c_2 (sufficiently large and depending only on $\bar{k}$, $\hat{k}$, $\bar{k}'$) and c_3, c_4, c_5, c_6, ε_0 (sufficiently small) such that if $\varepsilon \le \varepsilon_0$, $R_0 \ge c_1\sqrt{\varepsilon}$, $r_1 \ge c_2\sqrt{\varepsilon}$, $\eta \le c_3\varepsilon$ and

$$\rho_1 := c_4 \frac{\eta}{\sqrt{\varepsilon}}, \qquad \rho_2 := \min\{c_5 \frac{\eta}{\varepsilon \ln(\varepsilon/\eta)}, r_2\}, \qquad \sigma_1 := c_6 \frac{s_1}{\ln(\varepsilon/\eta)}, \qquad \sigma_2 := \frac{s_2}{2},$$

then the following holds. There exist two real–analytic symplectic transformations $\phi^\pm : \Omega^\pm_{(\rho_1,\rho_2)} \times \mathbf{T}_{\sigma_1} \times \mathbf{T}_{\sigma_2} \to D^0_{r_1} \times \Delta^0_{r_2} \times \mathbf{T}_{s_1} \times \mathbf{T}_{s_2}$ and two real–analytic functions $E^\pm : \Omega^\pm_{(\rho_1,\rho_2)} \to \mathbf{C}^2$ so that

$$\begin{aligned}
&\phi^\pm(P, J, Q, \psi) = (p, I, q, \varphi) \ , \\
&p = p^\pm(P, J, Q), \ I = J \ , \ q = q^\pm(P, J, Q) \ , \ \varphi = \varphi^\pm(P, J, Q, \psi) \ , \\
&E^\pm(P, J) = E(p^\pm(P, J, Q), J, q^\pm(P, J, Q)) \ , \\
&\phi^\pm(\Omega^\pm \times \mathbf{T}^2) = \mathcal{M}^\pm \ ,
\end{aligned} \tag{59}$$

where, letting $I = I_1 + iI_2$ and $P = P_1 + iP_2$:

$$\Omega^\pm := \Big\{(P_1, I_1) \in \mathbf{R}^2 \quad \text{s.t.} \quad P_1 \in D^\pm(I_1), \ I_1 \in \Delta^0\Big\},$$

$$\begin{aligned}
&D^+(I_1) := \Big(P^+(\eta, I_1), P^+(R_0^2/2, I_1)\Big), \\
&D^-(I_1) := \Big(P^-(-2\varepsilon k(I_1) + \eta, I_1), P^-(-\eta, I_1)\Big), \\
&P^+ := \frac{\sqrt{2}}{\pi} \int_0^\pi \sqrt{g(E, I, \theta)} d\theta, \qquad P^- := \frac{2\sqrt{2}}{\pi} \int_0^{\psi_0(E,I)} \sqrt{g(E, I, \theta)} d\theta,
\end{aligned}$$

and

$$g(E, I, \theta) := E + \varepsilon k(I)(1 + \cos\theta), \qquad \psi_0(E, I) := \arccos(-1 - E/\varepsilon k(I)).$$

Moreover, the following estimates hold for $(P, I) \in \Omega^\pm_{(\rho_1,\rho_2)}$:

$$\Big|\partial_P E^\pm(P, I)\Big| \quad \ge \quad \frac{1}{2\sqrt{2}\pi} \frac{\sqrt{\varepsilon k_1(I)}}{\ln\Big(1 + \sqrt{\frac{\varepsilon k_1(I)}{|E_1^\pm(P,I)|}}\Big)} \ge \frac{1}{4\pi} \frac{\sqrt{\varepsilon \hat{k}}}{\ln\Big(1 + \sqrt{\frac{\varepsilon \hat{k}}{\eta}}\Big)}, \tag{60}$$

$$\Big|\partial_I E^\pm(P, I)\Big| \quad \le \quad \frac{\sqrt{4\sqrt{6}\pi |k'(I)|\varepsilon}}{\ln\Big(1 + \sqrt{\frac{\varepsilon k_1(I)}{|E_1^\pm(P,I)|}}\Big)}, \tag{61}$$

where $E = E_1 + iE_2$. Finally, since

$$\partial_I E^\pm(P, I) = -\partial_I P^\pm(E^\pm(P, I), I)[\partial_E P^\pm(E^\pm(P, I), I)]^{-1} \ ,$$

we can write, for real P and I,

$$\partial_I E^\pm(P, I) = -\varepsilon k(I)[1 + Y^\pm(P, I)] \quad \text{with} \quad |Y^\pm(P, I)| < 1 \tag{62}$$

where

$$Y^\pm(P, I) := \left[\int_0^{\psi^\pm} \frac{1}{\sqrt{g(E^\pm(P, I), I, \theta)}} d\theta\right]^{-1} \int_0^{\psi^\pm} \frac{\cos\theta}{\sqrt{g(E^\pm(P, I), I, \theta)}} d\theta$$

with $\psi^+ := \pi$ and $\psi^- := \psi_0(E^-(P, I), I)$.

To apply Proposition 3.1 to the pendulum (57), we set $(p, I, q, \varphi) := (\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2)$, $(P, J, Q, \psi) = (\check{I}_1, \check{I}_2, \check{\varphi}_1, \check{\varphi}_2)$, $k := k_2$, $\Delta^0 := [\bar{I}_2 - \frac{5}{4}r, \bar{I}_2 + \frac{5}{4}r]$, $r_2 := r/4$, $\eta := \varepsilon^\lambda$, $R_0 := \frac{5}{4}r\varepsilon^\ell$, $r_1 := r\varepsilon^\ell/4$, $s_1 := s_2 := S/6$. If ε is sufficiently small we can apply Proposition 3.1 transforming the Hamiltonian

$$\hat{H}^{(1)} := \hat{H}^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon) := \frac{\hat{I}_1^2}{2} - \varepsilon k(\hat{I}_2)(1 + \cos \hat{\varphi}_1) + \varepsilon h(\hat{I}_2) + \varepsilon^a H_1^{(1)}$$

into the new Hamiltonians

$$\check{H}^{(1)\pm} := \hat{H}^{(1)} \circ \phi^\pm(\check{I}_1, \check{I}_2, \check{\varphi}_1, \check{\varphi}_2; \varepsilon) = E^\pm(\check{I}_1, \check{I}_2) + \varepsilon h(\check{I}_2) + \varepsilon^a H_1^{(1)\pm}(\check{I}_1, \check{I}_2, \check{\varphi}_1, \check{\varphi}_2)$$

which belongs to $\mathcal{H}_{\mathbf{R}}(\Omega^\pm_{(\rho_1,\rho_2)} \times \mathbf{T}_{\sigma_1} \times \mathbf{T}_{\sigma_2})$, where[14]

$$\rho_1 := \xi_8 \varepsilon^{\lambda - 1/2}, \quad \rho_2 := \xi_9 \varepsilon^{\lambda - 1} \ln^{-1}(1/\varepsilon), \quad \sigma_1 := \xi_{10} S \ln^{-1}(1/\varepsilon), \quad \sigma_2 := \xi_{11} S.$$

We now perform the analogous of step 3 in § 3.4. In order to apply the Normal Form Lemma, we take ε sufficiently small and we set[15] $d := 2$, $(I, \varphi) := (\check{I}_1, \check{I}_2, \check{\varphi}_1, \check{\varphi}_2)$, $h := E^\pm(\check{I}_1, \check{I}_2) + \varepsilon h(\check{I}_2)$, $f := \varepsilon^a H_1^{(1)\pm}$, $D := \Omega^\pm$, $\rho := (\rho_1, \rho_2)$, $\rho_0 := \rho_1$, $\sigma := \sigma_1$, $\Lambda := \{(0, k_2) \text{ s.t. } k_2 \in \mathbf{Z}\}$, $\alpha := \xi_{12}\sqrt{\varepsilon} \ln^{-1} \varepsilon^{-1}$, $K := \xi_{13}\varepsilon^{-C_0}$. So we find two real–analytic symplectic transformations $\hat{\phi}_1^\pm$ such that

$$\tilde{H}^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) := \check{H}^{(1)\pm} \circ \hat{\phi}_2^\pm(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) :=$$
$$E^\pm(\tilde{I}_1, \tilde{I}_2) + \varepsilon h(\tilde{I}_2) + \varepsilon^a H_1^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_2; \varepsilon) + \tilde{H}_2^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon)$$

belongs to $\mathcal{H}_{\mathbf{R}}(\Omega^\pm_{(\rho_1,\rho_2)/2} \times \mathbf{T}_{\sigma_1/6} \times \mathbf{T}_{\sigma_2/6})$ and

$$|\tilde{I}_1 - \check{I}_1|, |\tilde{I}_2 - \check{I}_2| \leq \rho_1/2^7 \ , \qquad \|\tilde{H}_2^{(2)\pm}\| \leq \xi_{14}\varepsilon^a \exp(-\xi_{15} S/\varepsilon^{C_0}) \ . \tag{63}$$

Now we complete our two symplectic transformations defining

$$\phi_2^\pm(\tilde{I}_1, \tilde{I}_2, \tilde{I}_3, \tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\varphi}_3; \varepsilon) := (\hat{\phi}_1^\pm \circ \phi^\pm(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon), \tilde{I}_3, \tilde{\varphi}_3)$$

so that

$$H^{(2)\pm}(\tilde{I}, \tilde{\varphi}; \varepsilon) := H^{(1)} \circ \phi_2^\pm(\tilde{I}, \tilde{\varphi}; \varepsilon) := \tag{64}$$
$$E^\pm(\tilde{I}_1, \tilde{I}_2) + \varepsilon h(\tilde{I}_2) + \omega \tilde{I}_3 + \varepsilon^a H_1^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_2; \varepsilon) + H_2^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}; \varepsilon).$$

belongs to $\mathcal{H}_{\mathbf{R}}(\Omega^\pm_{(\rho_1,\rho_2)/2} \times \mathbf{C} \times \mathbf{T}_{\sigma_1/6} \times \mathbf{T}_{\sigma_2/6} \times \mathbf{T}_{S/6})$ with

$$\|H_2^{(2)\pm}\| \leq \xi_{16}\varepsilon^a \exp(-\xi_{17} S/\varepsilon^{C_0}) \ .$$

We now perform the analogous of Step 4 in § 3.5. Let

$$\tilde{\Omega}^\pm := \Omega^\pm_{(1+2^{-7})(\rho_1,\rho_2)} \cap \mathbf{R}^2 \ .$$

From the form of the Hamiltonian (64) we deduce that $\forall\, (\tilde{I}_1(t_0), \tilde{I}_2(t_0)) \in \tilde{\Omega}^\pm$

$$|\tilde{I}_1(t) - \tilde{I}_1(t_0)| \leq \xi_{18}\varepsilon^{\xi_{19}} \ , \quad \forall\ 0 \leq t_0 \leq t \leq T_2(\varepsilon) := \xi_{20} \exp(-\xi_{21}/\varepsilon^{C_0}) < T_1(\varepsilon) \ . \tag{65}$$

Since $\tilde{I}_3(t) - \tilde{I}_3(t_0) = \hat{I}_3(t) - \hat{I}_3(t_0)$, from (56) and (65) we deduce, using the energy conservation, that

$$\left| [E^\pm(\tilde{I}_1(t_0), \tilde{I}_2(t)) + \varepsilon h(\tilde{I}_2(t))] - [E^\pm(\tilde{I}_1(t_0), \tilde{I}_2(t_0)) + \varepsilon h(\tilde{I}_2(t_0))] \right| \leq \xi_{22}\varepsilon^{\xi_{23}} \ . \tag{66}$$

[14] We have $\xi_8 = c_4$, $\xi_9 = c_5/(\lambda - 1)$, $\xi_{10} = c_6/6(\lambda - 1)$, $\xi_{11} = 1/12$.

[15] We can make such a choice of α, using (60) and (61).

We now prove that $G^{\pm}(\cdot) := E^{\pm}(\tilde I_1(t_0),\cdot) + \varepsilon h(\cdot)$ are non constant analytic functions. From (16) and (55), it follows that

$$\frac{dG^{\pm}}{dy}(y) = \varepsilon\Big[c_{02} y - k'(y) Y^{\pm}(\tilde I_1(t_0), y)\Big] . \tag{67}$$

Now we observe that by (55), (17), (18), $k(y)$ is effectively defined and analytic for all $|y| < \bar I_1$ and the same is true for $Y^{\pm}$. Thus, from the fact that $\lim_{y\to(-\bar I_1)^+} k'(y) = 0$ (as it follows differentiating (17)) and that $|Y^{\pm}| \le 1$, by (67) we deduce that

$$\lim_{y\to(-\bar I_1)^+} \frac{dG^{\pm}}{dy}(y) = -\varepsilon \bar I_1 c_{02},$$

which is different from 0 by (16) and the non-degeneracy assumption (6). This proves that $G^{\pm}$ are non constant analytic functions. Finally, using (66), we can apply Lemma 3.3 and find $\xi_{24}, \xi_{25} > 0$ such that

$$|\tilde I_2(t) - \tilde I_2(0)| \le \xi_{25}\varepsilon^{\xi_{24}}. \tag{68}$$

We remark that, in principle, ξ_{24}, ξ_{25} found with Lemma 3.3, depend on $\tilde I_1(t_0)$ but, since we work in compact sets $\tilde\Omega^{\pm}$, we can take them independent on $\tilde I_1(t_0)$.

We have proved stability for $0 \le t_0 \le t \le T_2(\varepsilon)$ and $(\tilde I(t_0), \tilde\varphi(t_0)) \in \tilde\Omega^{\pm} \times \mathbf{R} \times \mathbf{T}^3$. By (63) this implies stability for $(\check I(t_0), \check\varphi(t_0)) \in \Omega^{\pm} \times \mathbf{R} \times \mathbf{T}^3$. By (59) this is equivalent to prove stability for $(\hat I(t_0), \hat\varphi(t_0)) \in M^{\pm} \times \mathbf{R} \times \mathbf{T}^3$ where

$$\begin{aligned} M^+ &:= \Big\{(\hat I, \hat\varphi) \text{ s.t. } \varepsilon^{\lambda} \le E(\hat I_1, \hat I_2, \hat\varphi_1) \le R_0^2/2,\ \hat I_2 \in \Delta^0\Big\} \\ M^- &:= \Big\{(\hat I, \hat\varphi) \text{ s.t.} -2\varepsilon k(\hat I_2) + \varepsilon^{\lambda} \le E(\hat I_1, \hat I_2, \hat\varphi_1) \le -\varepsilon^{\lambda},\ \hat I_2 \in \Delta^0\Big\} . \end{aligned}$$

Using (58) and (56) it is immediate to prove stability for $(\hat I(0), \hat\varphi(0)) \in M \times \mathbf{R} \times \mathbf{T}^3$ and $0 \le t \le T_2(\varepsilon)$, where

$$M := \big\{(\hat I, \hat\varphi) \text{ s.t.} E(\hat I_1, \hat I_2, \hat\varphi_1) \le R_0^2/2,\ \hat I_2 \in \Delta^0\big\} .$$

Observing that $M \supset \Big(A_{\frac54 R} \cap \mathbf{R}^3\Big) \times \mathbf{T}^3$, by (22) and the fact that ϕ_0 is linear, we finally obtain (7). This finishes the proof in the case $(p, q) = (2, 1)$.

It remains to consider the case $(p, q) = (1, 1)$. The Hamiltonian (54) becomes

$$H^{(1)} = \frac{\hat I_1^2}{2} - \varepsilon k(\hat I_2)(1 + \cos 2\hat\varphi_1) + F \tag{69}$$

where $F := F(\hat I, \hat\varphi; \varepsilon) := \omega \hat I_3 + \varepsilon h_p(\hat I_2) + \varepsilon^a H_1^{(1)}(\hat I_1, \hat I_2, \hat\varphi_1, \hat\varphi_2; \varepsilon) + H_*^{(1)}(\hat I_1, \hat I_2, \hat\varphi; \varepsilon)$. Next, we perform the following linear change of variables $\hat I_1^\star := \hat I_1/2$, $\hat I_2^\star := \hat I_2$, $\hat I_3^\star := \hat I_3$, $\hat\varphi_1^\star := 2\hat\varphi_1$, $\hat\varphi_2^\star := \hat\varphi_2$, $\hat\varphi_3^\star := \hat\varphi_3$, casting the Hamiltonian (69) into the form

$$H_\star^{(1)} := H_\star^{(1)}(\hat I^\star, \hat\varphi^\star; \varepsilon) := 4\Big[E(\hat I_1^\star, \hat I_2^\star, \hat\varphi_1^\star; \varepsilon) + \frac14 F(2\eta_1^\star, \hat\varphi_1^\star/2)\Big] \tag{70}$$

with

$$E(\hat I_1^\star, \hat I_2^\star, \hat\varphi_1^\star; \varepsilon) := \frac{(\hat I_1^\star)^2}{2} - \varepsilon \frac{k(\hat I_2^\star)}{4}(1 + \cos\hat\varphi_1^\star) . \tag{71}$$

For ease of notation, we have omitted in F the $(\hat I_2^\star, \hat I_3^\star, \hat\varphi_2^\star, \hat\varphi_3^\star; \varepsilon)$–dependence, which, here, plays no rôle. We now apply[16] Proposition 3.1 to E defined in (71), finding

[16] Again we will omit the dependence on the variables $(\hat I_2^\star, \hat I_3^\star, \hat\varphi_2^\star, \hat\varphi_3^\star; \varepsilon)$.

two symplectic change of variables $\hat{I}_1^\star := p^\pm(P,Q)$, $\hat{\varphi}_1^\star := q^\pm(P,Q)$, putting the Hamiltonian (70) in the form

$$H^{(1)}_{\star\star} := 4\big[E^\pm(P) + F^\pm_\star(P,Q)\big] \ , \tag{72}$$

where $F^\pm_*(P,Q) := \frac{1}{4}F(2p^\pm(P,Q), q^\pm(P,Q)/2)$. We note that the functions $p^\pm$ are both 2π-periodic in Q. The function q^-, and hence F^-, is 2π-periodic in Q too; so we can define $(\check{I}_1, \check{\varphi}_1) := (P,Q)$ and proceed exactly as for $(p,q) = (2,1)$, applying the Normal Form Lemma and the subsequent arguments.

The positive energy case is different: in fact $q^+(P, Q+2\pi) = q^+(P,Q) + 2\pi$ so that F^+_* is only 4π-periodic in Q, but in order to apply the Normal Form Lemma we need a 2π-periodic function. We, therefore, define another linear change of variables $P_0 := 2P$, $Q_0 := Q/2$, so that (72) becomes

$$H^{(1)}_{\star\star\star} := 4\big[E^+(P_0/2) + F^+_{\star\star}(P_0,Q_0)\big] \ . \tag{73}$$

where $F^+_{\star\star}(P_0,Q_0) := F^+_\star(P_0/2, 2Q_0)$ which is 2π-periodic in Q_0. Therefore, we may define $(\check{I}_1, \check{\varphi}_1) := (P_0, Q_0)$ and proceed again as in the case $(p,q) = (2,1)$. The proof of Theorem 2.1 is, now, complete .

Appendix A. **Complex action-angle variables for the pendulum.** In this appendix we give a detailed proof of Proposition 3.1. Throughout this section, we shall denote z_1 and z_2, respectively, the real and imaginary part of a complex number $z = z_1 + iz_2$.

- *First step: estimates on the action domains*

Let $\bar{k}_1 := \sup_{I\in\Delta^0_{\rho_2}} |k_1(I)|$, $\bar{k}_2 := \sup_{I\in\Delta^0_{\rho_2}} |k_2(I)|$, $\bar{a}_1 := \bar{k}_1(1+\cosh s_1) + \bar{k}_2 \sinh s_1$ and $\bar{a}_2 := \bar{k}_2(1+\cosh s_1) + \bar{k}_1 \sinh s_1$.

For suitable $c_7, c_8 > 0$ small enough, we define $E_2^*(E_1) := c_7\eta \ln^{-1}(1+\sqrt{\frac{\varepsilon}{|E_1|}})$, $\overline{E} := c_8 r_1(R_0+r_1) + R_0^2/2$, and, for $\bar{I}_1 \in \Delta^0$, we define the domains $\mathcal{E}^+$, $\mathcal{E}^- := \mathcal{E}^-(\bar{I}_1)$

$$\begin{aligned}\mathcal{E}^+ &:= \{E_1 + iE_2 \ , \text{ s.t. } \eta/2 \le E_1 \le \overline{E} \ ; \ |E_2| \le E_2^*(E_1)\},\\ \mathcal{E}^- &:= \{E_1 + iE_2 \ , \text{ s.t. } -2\varepsilon k(\bar{I}_1) + \eta/2 \le E_1 \le -\eta/2 \ ; \ |E_2| \le E_2^*(E_1)\}.\end{aligned}$$

Let $F(p) := p^2/2$. We claim that

$$|I - \bar{I}_1| \le \rho_2, \ \bar{I}_1 \in \Delta^0, \ E \in \mathcal{E}^+ \cup \mathcal{E}^-(\bar{I}_1), \ \theta \in \mathbf{T}_{s_1} \quad \Longrightarrow \quad g(E,I,\theta) \in F(D^0_{r_1}). \tag{A.1}$$

It is immediate to see that $F(D^0_{r_1}) \supseteq \hat{\mathcal{E}}$, where

$$\begin{aligned}\hat{\mathcal{E}} &:= \{-r_1^2 \le 2E_1 \le R_0^2 - r_1^2, \ |E_2| \le r_1\sqrt{2E_1 + r_1^2}\}\\ &\quad \bigcup \ \{R_0^2 - r_1^2 \le 2E_1 \le (R_0+r_1)^2, \ |E_2| \le \hat{E}_2(E_1)\}\end{aligned}$$

and $\hat{E}_2(E_1) := [-E_1 + (R_0+r_1)^2/2]R_0/(R_0+r_1)$.

Next, we define

$$\widetilde{\mathcal{E}} := \{-2\varepsilon\bar{k}_1 \le E_1 \le 2\varepsilon, \ |E_2| \le E_2^0\} \cup \{2\varepsilon < E_1 \le \overline{E}, \ |E_2| \le 2c_7\eta\sqrt{E_1/\varepsilon}\} \ ,$$

where $E_2^0 := \max\{E_2^*(2\varepsilon), E_2^*(2\varepsilon\bar{k}_1)\}$, and $\bar{\mathcal{E}} := \bar{\mathcal{E}}^{(1)} \cup \bar{\mathcal{E}}^{(2)} \cup \bar{\mathcal{E}}^{(3)}$ with:

$$\begin{aligned}\bar{\mathcal{E}}^{(1)} &:= \{-2\varepsilon\bar{k}_1 - \varepsilon\bar{a}_1 \le E_1 \le 2\varepsilon - \varepsilon\bar{a}_1, \ |E_2| \le E_2^0 + \varepsilon\bar{a}_2\},\\ \bar{\mathcal{E}}^{(2)} &:= \{2\varepsilon - \varepsilon\bar{a}_1 < E_1 \le \overline{E} - \varepsilon\bar{a}_2, \ |E_2| \le 2c_7\eta\sqrt{(E_1+\varepsilon\bar{a}_1)/\varepsilon} + \varepsilon\bar{a}_1\},\\ \bar{\mathcal{E}}^{(3)} &:= \{\overline{E} - \varepsilon\bar{a}_1 < E_1 \le \overline{E} + \varepsilon\bar{a}_1, \ |E_2| \le 2c_7\eta\sqrt{\overline{E}/\varepsilon} + \varepsilon\bar{a}_2\}.\end{aligned}$$

We observe that obviously $\widetilde{\mathcal{E}} \subseteq \bar{\mathcal{E}}$; moreover (recalling the definitions of $\bar{a}_1, \bar{a}_2$)

$$|I - \bar{I}_1| \le \rho_2, \ \bar{I}_1 \in \Delta^0, \ E \in \widetilde{\mathcal{E}}, \ \theta \in \mathbf{T}_{s_1} \quad \Longrightarrow \quad g(E,I,\theta) \in \bar{\mathcal{E}}.$$

We observe that $E_1 \geq 2\varepsilon$ implies $\ln^{-1}(1+\sqrt{\varepsilon/E_1}) \leq 2\sqrt{E_1/\varepsilon}$ and hence $\widetilde{\mathcal{E}} \supseteq \mathcal{E}^+ \cup \mathcal{E}^-(\bar{I}_1)$. Now we prove $\bar{\mathcal{E}} \subseteq \hat{\mathcal{E}}$ which will imply (A.1), since $F(D^0_{r_1}) \supseteq \hat{\mathcal{E}} \supseteq \bar{\mathcal{E}} \supseteq \widetilde{\mathcal{E}} \supseteq \mathcal{E}^+ \cup \mathcal{E}^-(\bar{I}_1)$. It is simple to see that $\bar{\mathcal{E}} \subseteq \hat{\mathcal{E}}$ is implied by the following conditions:

(1) $(-2\varepsilon\bar{k}_1 - \varepsilon\bar{a}_1) + i(E^0_2 + \varepsilon\bar{a}_2) \in \hat{\mathcal{E}}$, which is implied by $r_1\sqrt{2(-2\varepsilon\bar{k}_1 - \varepsilon\bar{a}_1) + r_1^2} \geq E^0_2 + \varepsilon\bar{a}_2$

(2) $(\overline{E}+\varepsilon\bar{a}_1)+i(2c_7\eta\sqrt{\overline{E}/\varepsilon}+\varepsilon\bar{a}_2)$, which is implied by $\hat{E}_2(\overline{E}+\varepsilon\bar{a}_1) \geq 2c_7\eta\sqrt{\overline{E}/\varepsilon}+\varepsilon\bar{a}_2$

(3) if $(R_0^2 - r_1^2)/2 > 2\varepsilon - \bar{a}_1$ then $r_1\sqrt{2E_1 + r_1^2} \geq 2c_7\eta\sqrt{\overline{E}/\varepsilon} + \varepsilon\bar{a}_2$ for all $2\varepsilon - \bar{a}_1 \leq E_1 \leq \min\{(R_0^2 - r_1^2)/2, \overline{E} - \varepsilon\bar{a}_1\}$.

Defining $\tilde{k}_1 := \max\{1, \bar{k}_1\}$, one sees that (1) holds provided

$$c_2 \geq \max\{\sqrt{4c_7c_3\tilde{k}_1 + \sqrt{2}\bar{a}_2}\ ,\ 2\sqrt{2\bar{k}_1 + \bar{a}_1}\}.$$

If $c_8 \leq 1/2$, we have that $\hat{E}_2(\overline{E}+\varepsilon\bar{a}_1) \geq R_0(R_0r_1/2 - \varepsilon\bar{a}_1)/(R_0+r_1)$ and hence (2) holds provided $c_2 \geq 16\sqrt{2}c_7c_3$ and $c_1c_2 \geq 4\bar{a}_1 + 8\bar{a}_2$.
Finally, conditions $c_2 \geq \sqrt{2\bar{a}_2}$ and $c_2 \geq 2c_7c_3$ imply (3).

- In the following we will choose the positive branch of the square root i.e. if $z = |z|e^{i\alpha}$ with $\alpha \in (-\pi, \pi)$, then we define the analytic function $\sqrt{z} := \sqrt{|z|}e^{i\alpha/2}$. We also define $\ln z := \ln|z| + i\alpha$ and $\arccos z := -i\ln(z + i\sqrt{1-z^2})$.

We need the following elementary lemma (whose obvious proof is left to the reader):

Lemma A.1. *Let $x_1, x_2 \geq 0$. Then $\sqrt{x_1 \pm ix_2} = w_1 \pm iw_2$ and $(x_1 \pm ix_2)^{-1/2} = y_1 \mp iy_2$ where*

$$w_1(x_1,x_2) := \tfrac{1}{\sqrt{2}}\sqrt{x_1 + \sqrt{x_1^2+x_2^2}}\ , \qquad w_2(x_1,x_2) := \tfrac{1}{\sqrt{2}}\sqrt{-x_1 + \sqrt{x_1^2+x_2^2}}\ ,$$
$$y_1(x_1,x_2) := \frac{\sqrt{x_1+\sqrt{x_1^2+x_2^2}}}{\sqrt{2}\sqrt{x_1^2+x_2^2}}\ , \qquad y_2(x_1,x_2) := \frac{\sqrt{-x_1+\sqrt{x_1^2+x_2^2}}}{\sqrt{2}\sqrt{x_1^2+x_2^2}}\ .$$

We observe that, for x_1 fixed, w_1 (resp. y_1) is increasing (resp. decreasing) for $x_2 \geq 0$; y_2 is also increasing but only for $x_2 \leq \sqrt{3}x_1$. Moreover if $x_1 \geq x_2$ the following estimates hold

$$\sqrt{x_1} \leq w_1 \leq \sqrt{x_1+x_2} \leq \sqrt{2}\sqrt{x_1}\ , \qquad \frac{1}{3}\frac{x_2}{\sqrt{x_1}} \leq w_2 \leq \frac{1}{\sqrt{2}}\frac{x_2}{\sqrt{x_1}},$$
$$\frac{1}{\sqrt{2}}\frac{1}{\sqrt{x_1}} \leq \frac{1}{\sqrt{x_1+x_2}} \leq y_1 \leq \frac{\sqrt{2}}{\sqrt{x_1}}, \qquad \frac{1}{4}\frac{x_2}{x_1^{3/2}} \leq y_2 \leq \frac{1}{\sqrt{2}}\frac{x_2}{x_1^{3/2}}\ .$$

- *Second step: estimates on the action derivatives in the positive energy case*

In the following we put $\epsilon := \varepsilon k_1(I)$. We observe that, for $\theta \in [0,\pi]$, $2\tilde{g}/\pi^2 \leq g_1 \leq \tilde{g}$, where $\tilde{g}(E_1, I; \theta) := E_1 + \varepsilon k_1(I)(\pi-\theta)^2$. The following estimates hold[17]

$$\begin{aligned}
\frac{1}{\sqrt{\epsilon}}\ln\ 1+\sqrt{\frac{\epsilon}{E_1}} &\leq \int_0^\pi \frac{d\psi}{\sqrt{\tilde{g}(\psi)}} = \int_0^\pi \frac{d\psi}{\sqrt{E_1+\epsilon\psi^2}} = \frac{1}{\sqrt{\epsilon}}\int_0^a \frac{dy}{\sqrt{1+y^2}} \\
&= \frac{1}{\sqrt{\epsilon}}\operatorname{arcsinh}(a) \leq \frac{2\pi}{\sqrt{\epsilon}}\ln\ 1+\sqrt{\frac{\epsilon}{E_1}}\quad ; \\
\frac{1}{E_1\sqrt{E_1+\epsilon}} &\leq \int_0^\pi \frac{d\psi}{(\tilde{g}(\psi))^{3/2}} = \int_0^\pi \frac{d\psi}{(E_1+\epsilon\psi^2)^{3/2}} \\
&= \frac{1}{E_1\sqrt{\epsilon}}\int_0^a \frac{dy}{(1+y^2)^{3/2}} = \frac{\pi}{E_1^{3/2}\sqrt{1+\pi^2\epsilon/E_1}} \\
&\leq \frac{\pi}{E_1\sqrt{E_1+\epsilon}} \qquad \text{where}\quad a := \pi\sqrt{\epsilon/E_1}\ .
\end{aligned} \tag{A.2}$$

[17] Use $\ln(1+t) \leq \operatorname{arcsinh}(t) = \ln(t+\sqrt{1+t^2}) \leq 2\ln(1+t)$

Next, we prove that $\forall\, |I - \bar{I}_1| \le \rho_2,\ \bar{I}_1 \in \Delta^0,\ E \in \mathcal{E}^+\ g_1(E, I, \theta) \ge 2|g_2(E, I, \theta)|$. In fact we first have that if $c_5 \le 1/8\bar{k}'$ then $g_1(E, I, \theta) \ge E_1/2$ and, hence, we have only to prove that $E_1/4 \ge E_2^*(E_1) + 2\bar{k}' c_5 \eta \ln^{-1}(\varepsilon/\eta)$. Taking $c_5 \le c_7/4\bar{k}'$ we have only to verify that

$$|E_1| \ge 6E_2^*(E_1)\ . \tag{A.3}$$

It is easy to see that the previous inequality is verified for $c_7 \le 1/36$.
Consider, now, $I = \bar{I}_1 \in \Delta^0$ real and $g_2 = E_2 \ge 0$. Using Lemma A.1, $g_1 \ge \tilde{g}$ and (A.2) we have

$$P_2^+(E, \bar{I}_1) \ge \frac{\sqrt{2}}{3\pi} E_2 \int_0^\pi \frac{d\psi}{\sqrt{\tilde{g}(\psi)}} \ge \frac{\sqrt{2}}{3\pi} E_2 \frac{1}{\sqrt{\epsilon}} \ln \quad 1 + \sqrt{\frac{\epsilon}{E_1}} \quad . \tag{A.4}$$

Using Lemma A.1, $2\tilde{g}/\pi^2 \le g_1 \le \tilde{g}$ and (A.2) we have

$$\frac{1}{2\pi\sqrt{\epsilon}} \ln \quad 1 + \sqrt{\frac{\epsilon}{E_1}} \quad \le \partial_E P_1^+(E, I) \le \frac{\sqrt{2}\pi}{\sqrt{\epsilon}} \ln \quad 1 + \sqrt{\frac{\epsilon}{E_1}} \quad . \tag{A.5}$$

Using Lemma A.1, $2\tilde{g}/\pi^2 \le g_1$, (A.2) and the fact that $|g_2| \le |E_2| + 2\varepsilon\bar{k}_2$ we have

$$|\partial_E P_2^+(E, I)| \le (|E_2| + 2\varepsilon\bar{k}_2) \frac{\pi^4}{4\sqrt{2}} \frac{1}{E_1\sqrt{E_1 + \epsilon}}\ . \tag{A.6}$$

We observe that[18]

$$\frac{|E_2|}{E_1} \le \frac{c_7\eta}{E_1 \ln(1 + \sqrt{\varepsilon/E_1})} \le c_7 \ln(1 + \sqrt{\varepsilon/E_1}) \le \frac{c_7}{\sqrt{k_1(I)}} \ln(1 + \sqrt{\epsilon/E_1})\ . \tag{A.7}$$

Let us proceed by proving that

$$\varepsilon\bar{k}'\rho_2\pi^3 = \bar{k}' c_5 \pi^3 \eta \le \eta \ln(1 + \sqrt{2\epsilon/\eta}) \le 2E_1 \ln(1 + \sqrt{\epsilon/E_1})\ . \tag{A.8}$$

In fact, last inequality holds because the function $E_1 \ln(1 + \sqrt{\epsilon/E_1})$ is increasing and attains minimum for $E_1 = \eta/2$; the first inequality is proved, if $c_3 \le 1/8$ and using $k_1(I) \ge \hat{k}/2$, if $\bar{k}' c_5 \pi^3 \le \ln(1 + 2\sqrt{2}\sqrt{\hat{k}})$, which is verified if $c_5 \le (\pi^3\sqrt{2}\bar{k}')^{-1} \min\{1, \hat{k}\}$.
Using (A.5), (A.6), (A.7), (A.8) we have

$$|\partial_E P^+(E, I)| \le \frac{2\sqrt{2}\pi}{\sqrt{\epsilon}} \ln \quad 1 + \sqrt{\frac{\epsilon}{E_1}} \quad . \tag{A.9}$$

It remains to estimate

$$|\partial_I P^+(E, I)| = \frac{\sqrt{2}}{2\pi} \varepsilon |k'(I)| |\int_0^\pi \frac{1 + \cos\theta}{\sqrt{g(E, I; \theta)}} d\theta|\ .$$

We observe that

$$\int_0^\pi \frac{1 + \cos\theta}{\sqrt{g(E, I; \theta)}} d\theta \ = \ \int_0^2 \frac{\sqrt{x}}{\sqrt{2 - x}\sqrt{E + \varepsilon k(I) x}} dx \ \le \int_0^2 F_1(x) dx\ ,$$

where $F_1(x) := \sqrt{x}(\sqrt{2 - x}\sqrt{E_1 + \epsilon x})^{-1}$. In order to estimate last integral we split it as

$$\int_0^2 F_1(x) dx = \int_0^1 F_1(x) dx + \int_1^2 F_1(x) dx.$$

We have

$$\int_0^1 F_1(x) dx \le \int_0^1 \frac{\sqrt{x}}{\sqrt{E_1 + \epsilon x}} dx \le \int_0^1 \frac{\sqrt{x}}{\sqrt{E_1 x + \epsilon x}} dx = \frac{1}{\sqrt{E_1 + \epsilon x}}\ ,$$

$$\int_1^2 F_1(x) dx \le \frac{\sqrt{2}}{\sqrt{E_1 + \epsilon}} \int_1^2 \frac{1}{\sqrt{2 - x}} dx = \frac{2\sqrt{2}}{\sqrt{E_1 + \epsilon}}\ ,$$

which implies

$$|\partial_I P^+(E, I)| \le \frac{2\sqrt{2}}{\pi} |k'(I)| \frac{\varepsilon}{\sqrt{E_1 + \epsilon}} \le 2\sqrt{6} \frac{|k'(I)|}{\sqrt{k_1(I)}} \sqrt{\varepsilon}\ . \tag{A.10}$$

[18] We use the fact that $x \ln^2(1 + \sqrt{\varepsilon/E_1}) \ge \eta$ if $x \ge \eta/2$ and $c_3 \le 1/8$.

We, now, prove that

$$P^+(\mathcal{E}^+) \supseteq (D^+(\bar I_1))_{2\rho_1} \qquad \forall\, \bar I_1 \in \Delta^0. \tag{A.11}$$

Since $P_1^+(E_1+iE_2,\bar I_1) = \frac{\sqrt{2}}{\pi}\int_0^\pi w_1(E_1+\varepsilon k(\bar I_1)(1+\cos\theta),E_2)d\theta$, we have by Lemma A.1 that P_1^+ is an increasing function for $E_2\geq 0$. Hence, in order to prove (A.11), we have to prove the following estimates, $\forall\, E_1+iE_2\in\mathcal{E}^+$, $\bar I_1\in\Delta^0$:

$$\begin{array}{ll} \text{(i)} & P_2^+(E_1+iE_2^*(E_1),\bar I_1)\geq 2\rho_1,\\ \text{(ii)} & P_1^+(\eta/2+iE_2^*(\eta/2),\bar I_1)\leq P_1^+(3\eta/4,\bar I_1),\\ \text{(iii)} & P_1^+(\eta,\bar I_1)-P_1^+(3\eta/4,\bar I_1)\geq 2\rho_1,\\ \text{(iv)} & P_1^+(\overline{E},\bar I_1)-P_1^+(R_0^2/2,\bar I_1)\geq 2\rho_1.\end{array}$$

If $c_4\leq \frac{\sqrt{2}}{6\pi}c_7\min\{1,\frac{1}{\bar k}\}$ we obtain (i), since, from (A.4),

$$P_2^+(E_1+iE_2^*(E_1),\bar I_1)\geq c_7\frac{\sqrt{2}\eta}{3\pi\sqrt{\varepsilon}}\frac{\ln(1+\sqrt{k(\bar I_1)}\sqrt{\varepsilon/E_1})}{\sqrt{k(\bar I_1)}\ln(1+\sqrt{\varepsilon/E_1})}\geq c_7\frac{\sqrt{2}\eta}{3\pi\sqrt{\varepsilon}}\min\{1,\frac{1}{\bar k}\}.$$

Inequality (ii) follows from

$$\begin{aligned} P_1^+(\eta/2+iE_2^*(\eta/2),\bar I_1) &\leq (\sqrt{2}/\pi)\int_0^\pi\sqrt{\eta/2+E_2^*(E_1)+\varepsilon k(\bar I_1)(1+\cos\theta)}d\theta\\ &\leq (\sqrt{2}/\pi)\int_0^\pi\sqrt{3\eta/4+\varepsilon k(\bar I_1)(1+\cos\theta)}d\theta = P_1^+(3\eta/4,\bar I_1).\end{aligned}$$

Using (A.5) and the fact that $\eta\leq\varepsilon/8$, we have[19]

$$P_1^+(\eta,\bar I_1)-P_1^+(3\eta/4,\bar I_1)\geq \frac{1}{8\pi}\frac{\eta}{\sqrt{\varepsilon}}\Big[\frac{1}{\sqrt{\bar k}}\ln(1+2\sqrt{\bar k})\Big]\geq \frac{1}{8\pi}\frac{\eta}{\sqrt{\varepsilon}}\min\{1,\frac{1}{\bar k}\},$$

which implies (iii), provided $c_4\leq\frac{1}{16\pi}\min\{1,\frac{1}{\bar k}\}$.
Again, from (A.5), we have

$$P_1^+(\overline{E},\bar I_1)-P_1^+(R_0^2/2,\bar I_1)\geq \frac{c_8}{\pi}r_1(R_0+r_1)\frac{1}{\sqrt{\overline{E}}}\min\left\{1,2\sqrt{\frac{\overline{E}}{\varepsilon k(\bar I_1)}}\right\}.$$

Distinguishing the two cases for the minimum, we see that (iv) holds, provided $c_8c_2\geq\sqrt{2}\pi c_4c_3$ and $c_8c_2c_1\geq\sqrt{\bar k}\pi c_4c_3$.

- *Third step: estimates on the action derivatives in the negative energy case*

Defining $\tilde E := E+2\varepsilon k(I)$ and using the substitution $\xi := 1+\varepsilon k(I)\,(\cos\theta-1)/\tilde E$ we obtain

$$\begin{aligned} P^-(E,I) &= \frac{2\sqrt{2}}{\pi}\int_0^1\frac{\tilde E\sqrt{\xi}}{\sqrt{1-\xi}\sqrt{\tilde E\xi-E}}d\xi,\\ \partial_E P^-(E,I) &= \frac{\sqrt{2}}{\pi}\int_0^1\frac{1}{\sqrt{\xi}\sqrt{1-\xi}\sqrt{\tilde E\xi-E}}d\xi,\\ \partial_I P^-(E,I) &= \frac{\sqrt{2}k'(I)}{\pi k(I)}\int_0^1\frac{\sqrt{\tilde E\xi-E}}{\sqrt{\xi}\sqrt{1-\xi}}d\xi.\end{aligned}$$

We define $\tilde E\xi - E = x_1+ix_2$ where $x_1 := 2\epsilon\xi - E_1(1-\xi)$ and $x_2 := 2\varepsilon k_2(I)\xi - E_2(1-\xi)$. Using that $|E_1|\geq|E_2|$ and that, if $c_5c_3\leq \hat k/(2\bar k')$, we have $k_1(I)\geq|k_2(I)|$, we obtain that $x_1\geq|x_2|$.

We observe also that in order to perform the previous change of variables $\theta=\arccos(-1+(\xi-1)\tilde E/\varepsilon k(I))$ we have to verify that the argument of arccos is well defined[20]. For any

[19] Use $\ln(1+2x)/x\geq\min\{1,1/x\}$.
[20] We define $\arccos(z_1+iz_2)$ in the complementary of the set $\{z_1\in(-\infty,-1]\cup[1+\infty)\}$

$I \in \Delta^0_{\rho_2}$, we take $\bar{I}_1 \in \Delta^0$ with $|I - \bar{I}_1| \le \rho_2$. We have $-E_1 \le 2\varepsilon k(\bar{I}_1) - \eta/2$. We have to prove that, defining $y := k_2^2(I)/k_1^2(I)$ with $0 \le y \le 1$, if $E_2 = E_1 k_2(I)/k_1(I)$ then $[2\varepsilon(k(\bar{I}_1) + k_2(I)) - \eta/2]x + [-2\varepsilon k_1 + \eta/2] \ge 0$, which is verified provided $c_5 \le 1/(4\bar{k}')$.

In the following, in order to estimate the derivatives of P^-, we set $b := 2\epsilon/|E_1| \ge 1$ and we will use Lemma A.1.
For $I = \bar{I}_1 \in \Delta^0$, we have

$$\begin{aligned} P_2^-(E, \bar{I}_1) &= \frac{2\sqrt{2}}{\pi} \int_0^1 (\tilde{E}_1 y_2 + E_2 y_1) \frac{\sqrt{\xi}}{\sqrt{1-\xi}}\, d\xi \\ &\ge \int_0^1 \frac{\sqrt{2}\epsilon E_2}{\pi(2\epsilon\xi + |E_1|(1-\xi))^{3/2}} \frac{\sqrt{\xi}}{\sqrt{1-\xi}}\, d\xi \\ &\ge \frac{E_2}{2\pi\sqrt{\mu}} \int_0^b \frac{\sqrt{\xi}}{(2\epsilon\xi + |E_1|)^{3/2}}\, d\xi \ge \frac{E_2}{6\pi}\frac{1}{\sqrt{\mu}} \ln\ 1 + \sqrt{\frac{\epsilon}{|E_1|}}\ . \end{aligned} \tag{A.12}$$

Furthermore,[21]

$$\begin{aligned} \partial_E P_1^-(E, I) &= \frac{\sqrt{2}}{\pi} \int_0^1 \frac{y_1(x_1, x_2)}{\sqrt{\xi}\sqrt{1-\xi}}\, d\xi \\ &\le \frac{2}{\pi} \int_0^1 \frac{1}{\sqrt{\xi}\sqrt{2\epsilon\xi + |E_1|/2}}\, d\xi + \frac{2}{\pi}\int_0^1 \frac{1}{\sqrt{\xi}\sqrt{\epsilon}}\, d\xi \\ &= \frac{\sqrt{2}}{\pi\sqrt{\mu}} \int_0^b \frac{1}{\sqrt{t}\sqrt{1+t}}\, dt + \frac{\sqrt{2}}{\pi\sqrt{\mu}} \\ &\le \frac{12}{\pi\sqrt{\mu}} \ln\ 1 + \sqrt{\frac{\epsilon}{|E_1|}}\ . \end{aligned} \tag{A.13}$$

On the other hand,

$$\begin{aligned} \partial_E P_1^-(E, I) &\ge \frac{1}{\pi} \int_0^{1/2} \frac{1}{\sqrt{\xi}\sqrt{2\epsilon\xi + |E_1|}}\, d\xi = \frac{1}{\sqrt{2}\pi\sqrt{\mu}} \int_0^{b/2} \frac{1}{\sqrt{t}\sqrt{1+t}}\, dt \\ &\ge \frac{1}{2\pi\sqrt{\mu}} \ln\ 1 + \sqrt{\frac{\epsilon}{|E_1|}}\ . \end{aligned} \tag{A.14}$$

Using the fact that $|x_2| \le 2\varepsilon\bar{k}_2 + |E_2|$ and the estimate

$$|\partial_E P_2^-(E, I)| \le \frac{1}{\pi} \int_0^1 \frac{|x_2|}{\sqrt{\xi}\sqrt{1-\xi}\, x_1^{3/2}}\, d\xi \le \frac{6}{\pi|E_1|\sqrt{\mu}} + \frac{2}{\epsilon^{3/2}},$$

we obtain that, as in the positive energy case,

$$|\partial_E P^-(E, I)| \le \frac{2\sqrt{2}\pi}{\sqrt{\epsilon}} \ln\ 1 + \sqrt{\frac{\epsilon}{E_1}}\ . \tag{A.15}$$

Since $x_1 \ge |x_2|$, one has $\sqrt{|x_1 + ix_2|} \le \sqrt{2}\sqrt{x_1}$. From the previous inequality we conclude that

$$|\partial_I P^-(E, I)| \le \frac{2|k'(I)|}{\pi|k_1(I)|} \int_0^1 \frac{\sqrt{x_1}}{\sqrt{\xi}\sqrt{1-\xi}}\, d\xi \le 2\sqrt{6}\frac{|k'(I)|}{\sqrt{k_1(I)}}\sqrt{\varepsilon}\ . \tag{A.16}$$

Finally differentiating the equality $E^\pm(P^\pm(E, I), I) = E$ with respect to E and I we obtain respectively $\partial_P E^\pm(P, I) = [\partial_E P^\pm(E^\pm(P, I), I)]^{-1}$ and

$$\partial_I E^\pm(P, I) = -\partial_I P^\pm(E^\pm(P, I), I)[\partial_E P^\pm(E^\pm(P, I), I)]^{-1},$$

which, by (A.5), (A.9), (A.10), (A.14), (A.15), (A.16), imply (60) and (61) .

Next, we prove that

$$P^-(\mathcal{E}^-(\bar{I}_1)) \supseteq (D^-(\bar{I}_1))_{2\rho_1} \qquad \forall\, \bar{I}_1 \in \Delta^0. \tag{A.17}$$

[21] Use $\int_0^b \frac{1}{\sqrt{t}\sqrt{1+t}}\, dt \le 4\ln(1 + \sqrt{b})$.

Since $P_1^-(E_1+iE_2,\bar I_1)=\frac{2\sqrt2}{\pi}\int_0^1(\tilde E_1y_1-E_2y_2)\sqrt{\xi/1-\xi}d\xi$, we have by Lemma A.1 that (being y_1, $-y_2$ decreasing) P_1^- is a decreasing function (for $E_2\geq 0$). Hence, in order to prove (A.17), it is enough to prove the following estimates, $\forall\,\bar I_1\in\Delta^0$, $E_1+iE_2\in\mathcal{E}^-(\bar I_1)$:

$$\begin{aligned}
&\text{(i)} && P_2^-(E_1+iE_2^*(E_1),\bar I_1)\geq 2\rho_1,\\
&\text{(ii)} && P_1^-(-\eta/2+iE_2^*(\eta/2),\bar I_1)-P_1^-(-\eta,\bar I_1)\geq 2\rho_1,\\
&\text{(iii)} && P_1^-(-2\varepsilon k(\bar I_1)+\eta,\bar I_1)-P_1^-(-2\varepsilon k(\bar I_1)+\eta/2,\bar I_1)\geq 2\rho_1.
\end{aligned}$$

If $c_4\leq\frac{1}{12\pi}c_7\min\{1,\frac1k\}$ we obtain (i) because of:

$$P_2^-(E_1+iE_2^*(E_1),\bar I_1)\geq c_7\frac{\eta}{6\pi\sqrt\varepsilon}\frac{\ln(1+\sqrt{k(\bar I_1)}\sqrt{\varepsilon/E_1})}{\sqrt{k(\bar I_1)}\ln(1+\sqrt{\varepsilon/E_1})}\geq c_7\frac{\eta}{6\pi\sqrt\varepsilon}\min\{1,\frac1k\}.$$

Since

$$\begin{aligned}
&P_1^-(-\eta/2+iE_2^*(\eta/2),\bar I_1)-P_1^-(-\eta,\bar I_1)\\
\geq\ &P_1^-(-\eta/2,\bar I_1)-P_1^-(-\eta,\bar I_1)-|P_1^-(-\eta/2+iE_2^*(\eta/2),\bar I_1)-P_1^-(-\eta/2,\bar I_1)|,
\end{aligned}$$

using (A.14) and (A.15), if $c_7\leq 1/(32\pi^2)$, we have

$$\begin{aligned}
P_1^-(-\eta/2+iE_2^*(\eta/2),\bar I_1)-P_1^-(-\eta,\bar I_1) &\geq \frac{1}{4\pi}-c_7 4\pi\ \frac{\eta}{\sqrt\epsilon}\ln\ \ 1+\sqrt{\frac{\epsilon}{E_1}}\\
&\geq \frac{1}{8\pi}\frac{\eta}{\sqrt\varepsilon}\min\{1,\frac1k\},
\end{aligned}$$

which (exactly as in the positive energy case) yields (ii).
From (A.14) we have

$$P_1^-(-2\varepsilon k(\bar I_1)+\eta,\bar I_1)-P_1^-(-2\varepsilon k(\bar I_1)+\eta/2,\bar I_1)\geq\frac{\eta}{4\pi\sqrt\mu}\ln(1+1/\sqrt2)\geq\frac{1}{8\sqrt2\pi\sqrt k},$$

which yields (iii), provided $c_4\leq 1/(16\sqrt2\pi\sqrt k)$.

- *Fourth step: construction of the symplectic transformation*

We will find our symplectic transformation using the generating function

$$S(E,I,q):=\sqrt2\int_0^q\sqrt{g(E,I,\theta)}\,d\theta\ .$$

We note that in order to well define S we have to take into account the presence of the square root. In particular we are interested in the definition of the functions

$$\chi^\pm(E,I,q):=\frac{\partial_E S(E,I,q)}{\partial_E P^\pm(E,I)}\ ,\qquad \xi^\pm(E,I,q):=\frac{\partial_I S(E,I,q)}{\partial_I P^\pm(E,I)}\ .$$

Let $\mathcal{T}^+:=\mathbf{C}$ and $\mathcal{T}^-:=\{q\in\mathbf{C}\text{ s.t. }|q_1|<\pi\}$ and define[22]

$$\tilde{\mathcal{D}}^\pm(E,I):=\{q\in\mathcal{T}^\pm\text{ s.t. }g(E,I,q)\notin(-\infty,0]\}.$$

For any $E\in\mathcal{E}^\pm$, $I\in\Delta^0_{\rho_2}$, the functions

$$S(E,I,q)\ ,\quad \partial_E S(E,I,q)\ ,\quad \partial_I S(E,I,q)\ ,\quad \chi^\pm(E,I,q)\ ,\quad \xi^\pm(E,I,q)\ ,$$

are analytic in q on $\tilde{\mathcal{D}}^\pm(E,J)$. So the functions $\chi^\pm(E(p,I,q),I,q)$ and $\xi^\pm(E(p,I,q),I,q)$ are analytic in p,q on the disconnected set[23]

$$\begin{aligned}
\tilde{\mathcal{D}}^\pm\ &:=\ \tilde{\mathcal{D}}^\pm(I):=\{(p,q)\in\mathbf{C}\times\mathcal{T}^\pm\text{ s.t. }g(E(p,I,q),I,q)\notin(-\infty,0]\}\\
&=\ \{(p,q)\in\mathbf{C}\times\mathcal{T}^\pm\text{ s.t. }p^2/2\notin(-\infty,0]\}=\{(p,q)\in\mathbf{C}\times\mathcal{T}^\pm\text{ s.t. }p_1\neq 0\}\ .
\end{aligned}$$

[22] If $a,b\in\mathbf{C}$ we denote $(a,b):=\{z=a+t(b-a),\text{ with }t\in(0,1)\}$ (and, analogously, for $[a,b)$, $(a,b]$, $[a,b]$); symbols like $(a,\alpha\infty)$, with $\alpha\in\mathbf{C}$ and $|\alpha|=1$, (or $[a,\alpha\infty)$, $(\alpha\infty,\beta\infty)$, etc) denote lines: $(a,\alpha\infty):=\{z=a+\alpha t,\text{ with }t>0\}$.

[23] We see that $\tilde{\mathcal{D}}^\pm$ does not really depend on I.

Our next step will be to define both $\bar{\chi}^{\pm}(p,I,q) := \chi^{\pm}(E(p,I,q),I,q)$ and $\bar{\xi}^{\pm}(p,I,q) := \xi^{\pm}(E(p,I,q),I,q)$ for all $(p,q) \in \mathbf{C} \times \mathcal{T}^{\pm}$. We set $\tilde{\chi}^{\pm} := \tilde{\chi}^{\pm}(p,I,q)$ and $\tilde{\xi}^{\pm} := \tilde{\xi}^{\pm}(p,I,q)$ where

$$\tilde{\chi}^{\pm} := \begin{cases} \dfrac{1}{\sqrt{2}\, \partial_E P^{\pm}(E(p,I,q),I)} \displaystyle\int_0^q \dfrac{d\theta}{\sqrt{g(E(p,I,q),I,\theta)}} \ , & \text{if } \ p_1 > 0, \\[2ex] \pi - \dfrac{1}{\sqrt{2}\, \partial_E P^{\pm}(E(p,I,q),I)} \displaystyle\int_0^q \dfrac{d\theta}{\sqrt{g(E(p,I,q),I,\theta)}} \ , & \text{if } \ p_1 < 0, \end{cases}$$

and

$$\tilde{\xi}^{\pm} := \begin{cases} \dfrac{\epsilon}{\sqrt{2}\, \partial_I P^{\pm}(E(p,I,q),I)} \displaystyle\int_0^q \dfrac{1+\cos\theta}{\sqrt{g(E(p,I,q),I,\theta)}}\, d\theta \ , & \text{if } \ p_1 > 0, \\[2ex] \pi - \dfrac{\epsilon}{\sqrt{2}\, \partial_I P^{\pm}(E(p,I,q),I)} \displaystyle\int_0^q \dfrac{1+\cos\theta}{\sqrt{g(E(p,I,q),I,\theta)}}\, d\theta \ , & \text{if } \ p_1 < 0, \end{cases}$$

which are well defined and analytic for $(p,q) \in \tilde{\mathcal{D}}^{\pm}$. Notice that, in the positive energy case, there are no problems with the definition of $\bar{\chi}^{+}$ and $\bar{\xi}^{+}$, and we note that

$$\bar{\chi}^{+}(p,I,q+2\pi) = \bar{\chi}^{+}(p,I,q) + 2\pi \quad \text{and} \quad \bar{\xi}^{+}(p,I,q+2\pi) = \bar{\xi}^{+}(p,I,q) + 2\pi \ . \tag{A.18}$$

In the negative energy case we proceed differently. We define

$$F_E(E,I,p) := -\int_0^p \frac{1}{\epsilon\sqrt{\hat{g}(E,I,z)}}\, dz \ , \qquad F_I(E,I,p) := -\int_0^p \frac{-E+z^2}{\epsilon\sqrt{\hat{g}(E,I,z)}}\, dz \ ,$$

where $\hat{g}(E,I,p) := 1-(-1-E/\epsilon+p^2/2\epsilon)^2$ is analytic on the complex domain $\hat{\mathcal{D}}^{-}(E,I) := \{p \in \mathbf{C} \text{ s.t. } \hat{g}(E,I,p) \notin (-\infty,0]\}$. Then, $F_E(E(p,I,q),I,p)$ and $F_I(E(p,I,q),I,p)$ are well defined and analytic on[24]

$$\hat{\mathcal{D}}^{-} := \hat{\mathcal{D}}^{-}(I) := \{(p,q) \in \mathbf{C} \times \mathcal{T}^{-} \text{ s.t. } \hat{g}(E(p,I,q),I,p) \notin (-\infty,0]\} =$$
$$\{(p,q) \in \mathbf{C} \times \mathcal{T}^{-} \text{ s.t. } 1-\cos^2 q \notin (-\infty,0]\} = \{(p,q) \in \mathbf{C} \times \mathcal{T}^{-} \text{ s.t. } q_1 \neq 0\} \ .$$

We now split the integral in the definition of S, $\partial_E S$ and $\partial_I S$ as $\int_0^q = \int_0^{\psi_0} + \int_{\psi_0}^q$ and in the second integral we let $\theta = \arccos(-1-E/\epsilon+z^2/2\epsilon)$. Then, defining[25]

$$\hat{\chi}^{-}(p,I,q) := \begin{cases} \pi/2 + \big(\partial_E P^{-}(E(p,I,q),I)\big)^{-1} F_E(E(p,I,q),I,p) \ , & \text{if } \ q_1 > 0, \\[2ex] -\pi/2 - \big(\partial_E P^{-}(E(p,I,q),I)\big)^{-1} F_E(E(p,I,q),I,p) \ , & \text{if } \ q_1 < 0, \end{cases}$$

and

$$\hat{\xi}^{-}(p,I,q) := \begin{cases} \pi/2 + \big(\partial_I P^{-}(E(p,I,q),I)\big)^{-1} F_I(E(p,I,q),I,p) \ , & \text{if } \ q_1 > 0, \\[2ex] -\pi/2 - \big(\partial_I P^{-}(E(p,I,q),I)\big)^{-1} F_I(E(p,I,q),I,p) \ , & \text{if } \ q_1 < 0, \end{cases}$$

we have $\forall\, I \in \Delta^0_{\rho_2}$, $\forall\ (p,q) \in \tilde{\mathcal{D}}^{-} \cap \hat{\mathcal{D}}^{-}$

$$\tilde{\chi}^{-}(p,I,q) \equiv \hat{\chi}^{-}(p,I,q) \mod 2\pi \quad \text{and} \quad \tilde{\xi}^{-}(p,I,q) \equiv \hat{\xi}^{-}(p,I,q) \mod 2\pi \ . \tag{A.19}$$

---

[24] Also in this case $\hat{\mathcal{D}}^{-}$ does not really depend on I.

[25] We note that $\hat{\chi}^{-}$ and $\hat{\xi}^{-}$ are analytic on $\hat{\mathcal{D}}^{-}$.

Using (A.19), we can finally define[26] $\bar\chi^-, \bar\xi^- : (p,q) \in \tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^- \to \mathbf{C}/2\pi\mathbf{Z}$

$$\bar\chi^-(p,I,q) := \begin{cases} \tilde\chi^-(p,I,q), & \text{if } (p,q) \in \tilde{\mathcal{D}}^-, \\ \hat\chi^-(p,I,q), & \text{if } (p,q) \in \hat{\mathcal{D}}^-, \end{cases}$$

and

$$\bar\xi^-(p,I,q) := \begin{cases} \tilde\xi^-(p,I,q), & \text{if } (p,q) \in \tilde{\mathcal{D}}^-, \\ \hat\xi^-(p,I,q), & \text{if } (p,q) \in \hat{\mathcal{D}}^- , \end{cases}$$

where, on $\tilde{\mathcal{D}}^-$,

$$\chi^-(E(p,I,q),I,q) \equiv \bar\chi^-(p,I,q) \quad \text{and} \quad \xi^-(E(p,I,q),I,q) \equiv \bar\xi^-(p,I,q) \ .$$

Moreover we finally extend by periodicity the definition of $\bar\chi^-(p,I,q)$ and $\bar\xi^-(p,I,q)$ on all $\{q \in \mathbf{C} \text{ s.t. } q_1 \neq \pi + 2k\pi,\ k \in \mathbf{Z}\} = \cup_{k\in\mathbf{Z}}\ 2k\pi + \mathcal{T}^-$ in the following way: if $q \in 2k\pi + \mathcal{T}^-$ we define $\bar\chi^-(p,I,q) := \bar\chi^-(p,I,q-2k\pi)$ and $\bar\xi^-(p,I,q) := \bar\xi^-(p,I,q-2k\pi)$.

Now we are able to construct our symplectic transformation. Since $\partial P^\pm_E \neq 0$, by the Implicit Function Theorem, there exists $E^\pm = E^\pm(P,J)$ such that

$$P^\pm(E^\pm(P,J),J) \equiv P. \tag{A.20}$$

Let $S^\pm(P,J,q) := S(E^\pm(P,J),J,q)$; we , then, define the following generating functions (which depend on the new actions and on the old angles): $G^\pm(P,J,q,\varphi) := J\varphi + S^\pm(P,J,q)$. Our symplectic transformation $\phi^\pm$ is implicitly defined by

$$\begin{cases} p = \partial_q G^\pm = \partial_q S^\pm(P,J,q), & Q = \partial_P G^\pm = \partial_P S^\pm(P,J,q), \\ I = \partial_\varphi G^\pm = J, & \psi = \partial_J G^\pm = \varphi + \partial_J S^\pm(P,J,q), \end{cases}$$

We want to express $(\phi^\pm)^{-1}$ as a function of the old variables (p,I,q,φ). We immediately have $J = I$ and $P = P^\pm(E(p,I,q),I)$. Differentiating (A.20) with respect to J and P we have, respectively, $\partial_E P^\pm \partial_J E^\pm + \partial_J P^\pm = 0$ and $\partial_E P^\pm \partial_P E^\pm = 1$. Now, we can express the new angles as functions of the old variables:

$$\begin{cases} Q = Q^\pm(p,I,q) := \bar\chi^\pm(p,I,q) \ , \\ \psi = \psi^\pm(p,I,q,\varphi) := \varphi - \partial_I P^\pm(E(p,I,q),I)\Big[Q^\pm(p,I,q) - \bar\xi^\pm(p,I,q)\Big] \ . \end{cases}$$

We observe that Q^- and ψ^- are 2π-periodic in q by definition of $\bar\chi^-$ and ξ^-; by (A.18) we deduce that ψ^+ is 2π-periodic in q too and $Q^+(p,I,q+2\pi) = Q^+(p,I,q) + 2\pi$.

- *Fifth step: estimate on the angle analyticity radius*

We first study the analyticity radius in Q. Fix $I \in \Delta^0_{\rho_2}$ and $\bar I_1 \in \Delta^0$ with $|I - \bar I_1| \le \rho_2$. We must prove that $\forall P_* \in D^\pm_{\rho_1}$ and $\forall Q_* \in \mathbf{T}_{\sigma_1}$ there exist $p^\pm_*$ and $q^\pm_*$ such that $Q^\pm(p^\pm_*, I, q^\pm_*) = Q_*$. So it is sufficient to prove that $\forall E_* \in \mathcal{E}^\pm(\bar I_1)$ we have $\chi^\pm(E_*, I, \mathbf{T}_{s_1}) \supseteq \mathbf{T}_{\sigma_1}$.

We first consider the positive energy case.
We observe that we have $\chi^+(E,I,0) = 0$, $\chi^+(E,I,\pm\pi) = \pm\pi$. Let us first consider the case $I = \bar I_1 \in \Delta^0$, $E = E_1 \in \mathcal{E}^+$. In such a case

$$\chi^+(E_1,\bar I_1,(-\pi,\pi)) = (-\pi,\pi), \quad \chi^+(E_1,\bar I_1,(0,\pm i\infty)) = (0,\pm i s^+(E_1,\bar I_1)) \ ,$$

[26] We observe that $\tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^-$ is an open set and that its complementary set $(\tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^-)^c = \{(p,q) \in \mathbf{C}\times\mathcal{T}^-$, s.t. $p_1 = 0,\ q_1 = 0\}$ does not interest our analysis. In fact, if $(p,q) \in (\tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^-)^c$, then $p = ip_2$ and $q = iq_2$ and, hence, we have $E(ip_2,I,iq_2) = -p^2/2 - \varepsilon k(I)(1+\cosh q_2)$ and $E_1(ip_2,I,iq_2) = -p_2^2/2 - \varepsilon k_1(I)(1+\cosh q_2) \le -2\varepsilon k_1(I) < -2\varepsilon k(\bar I_1) + \eta/2$ where $|I - \bar I_1| \le \rho_2$ (and we have used the fact that $c_5 \le 1/4\bar k'$). We conclude that $E(ip_2,I,iq_2) \notin \mathcal{E}^-(\bar I_1)$.

and

$$\chi^+(E_1,\bar I_1,(\pm\pi,\pm\pi\pm i\psi_0(E_1,\bar I_1)))=(\pm\pi,\pm\pi\pm is^+(E_1,\bar I_1)),$$

where

$$s^\pm(E_1,\bar I_1):=\sqrt{2}(\partial_E P^\pm)^{-1}\int_0^\infty \frac{d\theta}{\sqrt{E_1+\varepsilon k(\bar I_1)(1+\cosh\theta)}}\ . \tag{A.21}$$

In fact, it is $\chi^+(E_1,\bar I_1,\mathcal{D}(E_1,\bar I_1))=\mathbf{T}_{s^+(E_1,\bar I_1)}$. We will prove that

$$\chi^+(E_1,\bar I_1,\mathbf{T}_{s_1}\cap\tilde{\mathcal{D}}^+(E_1,\bar I_1))\supseteq\mathbf{T}_{\sigma^+(E_1,\bar I_1,s_1)} \tag{A.22}$$

where, for $s>0$,

$$\sigma^\pm(E,I,s):=\inf_{t\in(-\pi,\pi)}\chi_2^\pm(E,I,t+is)\ . \tag{A.23}$$

Observe that $\partial_E P^+\in\mathbf{R}$ and $g=g_1-ig_2$, where[27] $g_1=E_1+\varepsilon k(\bar I_1)(1+\cos q_1\cosh q_2)$ and $g_2=\varepsilon k(\bar I_1)\sin q_1\sinh q_2$. Splitting the integral $\int_0^{t+is}=\int_0^{is}+\int_{is}^{is+t}$, we have $\int_0^{is}1/\sqrt{g}=i\int_0^s 1/\sqrt{g_1}$ and, using the notation of Lemma A.1, we obtain $\mathrm{Im}\int_{is}^{is+t}1/\sqrt{g}=\int_0^t y_2(g_1,g_2)$, which, since $y_2>0$, attains its minimum at $t=0$. Collecting all these informations we have that

$$\sigma^+(E_1,\bar I_1,s)=\chi_2^+(E_1,\bar I_1,is):=\sqrt{2}(\partial_E P^+)^{-1}\int_0^s\frac{d\theta}{\sqrt{E_1+\varepsilon k(\bar I_1)(1+\cosh\theta)}}.$$

It is easy to see that

$$\int_0^s\frac{d\theta}{\sqrt{E_1+\varepsilon k(\bar I_1)(1+\cosh\theta)}}\geq c_9\frac{1}{\sqrt{\varepsilon k(\bar I_1)}}\ln\left(1+s\sqrt{\frac{\varepsilon k(\bar I_1)}{E_1+2\varepsilon k(\bar I_1)}}\right)\ .$$

Thus, by (A.9), we get

$$\sigma^+(E_1,\bar I_1,s)\geq c_{10}\ln\left(1+s\sqrt{\frac{\varepsilon k(\bar I_1)}{E_1+2\varepsilon k(\bar I_1)}}\right)\ln^{-1}\left(1+\sqrt{\frac{\varepsilon k(\bar I_1)}{E_1}}\right)\ ,$$

which implies that, $\forall\,\eta/2\leq E_1\leq\overline{E}$ and $\bar I_1\in\Delta^0$,

$$\sigma^+(E_1,\bar I_1,s)\geq c_{11}\frac{s}{\ln(\varepsilon/\eta)}.$$

In the general case, using the estimates on χ^+ and its derivatives[28], one has that, if $E=E_1+iE_2\in\mathcal{E}^+$ and $I\in\Delta^0_{\rho_2}$, then $\sigma^+(E,I,s)\geq c_{12}\sigma^+(E_1,\bar I_1,s)\geq c_{13}\frac{s}{\ln(\varepsilon/\eta)}$. Taking $s:=s_1$ and $c_6\leq c_{13}$, we have the claim concerning the form of σ_1.

We now pass to the negative energy case. As before fix E and I and observe that $\chi^-(E,I,0)=0$ and $\chi^-(E,I,\pm\psi_0(E,I))=\pm\pi/2$. Consider first the case $I=\bar I_1\in\Delta^0$ and $E=E_1\in\mathcal{E}^-(\bar I_1)$. We find $\chi^-(E_1,\bar I_1,(-\psi_0(E_1,\bar I_1),\psi_0(E_1,\bar I_1)))=(-\pi/2,\pi/2)$ and $\chi^-(E_1,\bar I_1,(0,\pm i\infty))=(0,\pm is^-(E_1,\bar I_1))$, where $s^-(E_1,\bar I_1)$ was defined in (A.21). It is simple to see that we have

$$\left\{\bar\chi^-(p,\bar I_1,q)\ \text{s.t.}\ (p,q)\in\tilde{\mathcal{D}}^-\cup\hat{\mathcal{D}}^-,\ E(p,\bar I_1,q)=E_1,\ |q_2|<s_1\right\}\supseteq$$
$$\supseteq\left\{|Q_2|<\sigma^-(E_1,\bar I_1,s_1)\right\}$$

which is analogous to (A.22). The estimate on σ_1 for the general case $E\in\mathcal{E}^-(\bar I_1)$ and $I\in\Delta^0_{\rho_2}$ with $\bar I_1\in\Delta^0$, $|I-\bar I_1|\leq\rho_2$, follows exactly as in the positive energy case.

We now briefly discuss the analyticity radius in the angle ψ. Observing that, as it is simple to see, $|\bar\chi^\pm|,|\bar\xi^\pm|\leq c_{14}$ and remembering (A.10) and (A.16), we see that, $|\psi^\pm(P,I,q,\varphi)-\varphi|\leq c_{15}\sqrt{\varepsilon}$. Hence, if ε is sufficiently small, we can take $\sigma_2=s_2/2$,. The proof of Proposition 3.1 is now complete.

[27] For symmetry reasons we can consider $t,q_1\geq 0$.

[28] See Appendix B of [4].

REFERENCES

[1] Arnold V. I. (editor): *Encyclopedia of Mathematical Sciences*, Dynamical Systems III, Springer–Verlag **3** (1988)

[2] Benettin, G., Fassò, F., Guzzo, M.: *Fast rotations of the rigid body: a study by Hamiltonian perturbation theory. II. Gyroscopic rotations*, Nonlinearity **10** (1997), no. 6, 1695–1717

[3] Biasco L., Chierchia L.: *Nekhoroshev stability for the D'Alembert problem of Celestial Mechanics*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. 13 (2002), no. 2, 85–89

[4] Biasco L., Chierchia L.: *On the stability of some properly–degenerate Hamiltonian systems*, Discrete Contin. Dyn. Syst. A, **9**, n. 2, (2003) 233-262

[5] Biasco L., Chierchia L.: *Effective Hamiltonian for the D'Alembert Planetary model near a spin/orbit resonance*, Celestial Mechanics & Dynamical Astronomy, **83**, (2002) 223-237

[6] Biasco L., Chierchia L., Treschev D.: *Absense of diffusion in properly degenerate systems with two degrees of freedom*, preprint 2003, 29 pages

[7] Chierchia L., Gallavotti G.: *Drift and diffusion in phase space*, Ann. Inst. Henri Poincaré, Phys. Théor., **60**, 1–144 (1994). *Erratum*, Ann. Inst. Henri Poincaré, Phys. Théor., **68**, no. 1, 135 (1998)

[8] Gallavotti G., Gentile G., Mastropietro V.: *Separatrix splitting for systems with three time scales*, Commun. Math. Phys. **202**, 197–236 (1999)

[9] Gallavotti G., Gentile G., Mastropietro V.: *Hamilton–Jacobi equation and existence of heteroclinic chains in three time scale systems*, Nonlinearity, **13**, 323–340 (2000)

[10] de LaPlace P. S.: *Mécanique Céleste*, tome II, livre 5, ch. I, 1799, English Translation by N. Bowditch, reprinted by AMS Chelsea Publishing, 1966

[11] Lochak, P. :*Canonical perturbation theory: an approach based on joint approximations*, (Russian) Uspekhi Mat. Nauk 47 (1992), no. 6(288), 59–140; translation in Russian Math. Surveys 47 (1992), no. 6, 57–133

[12] Neishtadt, A. I.: *The separation of motions in systems with rapidly rotating phase*, J. Appl. Math. Mech. **48** (1984), no. 2, 133–139 (1985); translated from Prikl. Mat. Mekh. 48 (1984), no. 2, 197–204 (Russian)

[13] Nekhoroshev N. N.: *An exponential estimate of the time of stability of nearly–integrable Hamiltonian systems*, I, Usp. Mat. Nauk. **32** 5–66 (1977); Russ. Math. Surv. **32**, 1–65 (1977)

[14] Pöschel J.: *Nekhoroshev estimates for quasi–convex Hamiltonian Systems* Math. Zeitschrift **213**, 187–216 (1993)

[15] Procesi M.: *Estimates on Hamiltonian splittings. Tree techniques in the theory of homoclinic splitting and Arnold diffusion for a–priori stable systems*, PhD. Thesis, Università degli Studi "La Sapienza", Roma (Italy), (2001)

Received (revised version) from Editor for publication May 2004.

E-mail address: biasco@mat.uniroma3.it
E-mail address: luigi@mat.uniroma3.it