ON THE STABILITY OF SOME PROPERLY–DEGENERATE HAMILTONIAN SYSTEMS WITH TWO DEGREES OF FREEDOM

Luca Biasco

Settore di Analisi Funzionale
SISSA/ISAS
Via Beirut 2–4, 34013 Trieste (Italy)

Luigi Chierchia

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

(Communicated by Antonio Ambrosetti)

Abstract. Properly degenerate nearly–integrable Hamiltonian systems with two degrees of freedom such that the "intermediate system" depend explicitly upon the angle–variable conjugated to the non–degenerate action–variable are considered and, in particular, model problems motivated by classical examples of Celestial Mechanics, are investigated. Under suitable "convexity" assumptions on the intermediate Hamiltonian, it is proved that, in every energy surface, the action variables stay forever close to their initial values. In "non convex" cases, stability holds up to a small set where, in principle, the degenerate action–variable might (in exponentially long times) drift away from its initial value by a quantity independent of the perturbation. Proofs are based on a "blow up" (complex) analysis near separatrices, KAM techniques and energy conservation arguments.

1. **Introduction and results.** As pointed out with particular emphasis by H. Poincaré [8], one of the main problem in Dynamical Systems concerns the stability of action variables in nearly–integrable Hamiltonian systems. Notwithstanding the efforts of Poincaré himself and the great success of powerful, more modern techniques such as averaging theory, KAM and Nekhoroshev theory (see [1] for general information), the "action–stability" problem for general nearly–integrable Hamiltonian systems remains essentially open. By "action–stability problem" we mean the following. Consider a (real–analytic) nearly–integrable, one–parameter family of Hamiltonian functions $H(I, \varphi; \varepsilon) = h(I) + \varepsilon f(I, \varphi)$ where (I, φ) are standard symplectic "action–angle" variables in a $2d$–dimensional phase space (the angles φ_i are defined modulus 2π) and ε is a small parameter. The problem is, then, to give upper bounds on the quantity $|I(t) - I_0|$, where $(I(t), \varphi(t))$ denotes the H–flow at time t of the initial datum (I_0, φ_0), and "stability" means that $\sup_t |I(t) - I_0|$ goes to zero when ε goes to zero.

V.I. Arnold, in 1963, in one of the fundamental paper on KAM theory [2], conjectured that the general feature of nearly–integrable Hamiltonian dynamics with more than two degrees–of–freedom (i.e., with phase space of dimension greater than four) is action–instability[1]. On the other hand, KAM theory yields, in general, "metric" stability (i.e., stability for the majority of initial data) and this implies "total stability" in systems with two degrees of freedom: in such a case, under suitable non–degeneracy assumptions, the three–dimensional energy surfaces are separated by a multitude of two–dimensional invariant KAM tori and trajectories are trapped in–between these tori allowing only for a small (with ε) variation of the action variables (*for any time and for any initial data*). The non–degeneracy KAM assumption, at fixed energy, is that the (unperturbed) map between action variables on fixed energy surface and the frequency map viewed in projective space is a diffeomorphysm.

The main motivation for Poincaré to look up at the action–stability problem came from Celestial Mechanics. Now, a typical feature in Celestial Mechanics is that the unperturbed system is *properly degenerate*, i.e., the unperturbed Hamiltonian function does not depend upon all action variables[2]. In such a case the above mentioned non–degeneracy condition is obviously strongly violated. However, in [2], Arnold proved the following result (compare also [1], Chapter 5, Section 3). Consider a nearly–integrable (real–analytic) Hamiltonian system with two degrees of freedom governed by

$$H(I,\varphi;\varepsilon) := H_0(I;\varepsilon) + \varepsilon^2 H_1(I,\varphi) := H_{00}(I_1) + \varepsilon H_{01}(I) + \varepsilon^2 H_1(I,\varphi) \ , \quad (1.1)$$

$(I,\varphi) = (I_1, I_2, \varphi_1, \varphi_2) \in U \times \mathbb{T}^2$ (where $U \subset \mathbb{R}^2$ and $\mathbb{T}^2$ denotes the standard two–dimensional torus $\mathbb{R}^2/(2\pi\mathbb{Z}^2)$). We say that the "perturbation removes the degeneracy[3]" on the energy level $H^{-1}(E)$, if

$$\frac{\partial H_{00}}{\partial I_1}(I) \neq 0 \ , \qquad \frac{\partial^2 H_{01}}{\partial I_2^2}(I) \neq 0 \ , \qquad \forall I \in H_0^{-1}(E) \ . \qquad (1.2)$$

THEOREM 1.1. ([2])*If, in a (real–analytic) properly degenerate system with two degrees of freedom, the perturbation removes the degeneracy (i.e., condition (1.2) holds), then, for all ε small enough, total stability holds (i.e., for all initial data on the given energy level, the values of the action–variables stay forever near their initial values).*

REMARK 1. (i) If condition (1.2) is violated "instability channels" may appear as suggested by the following example (which is a trivial modification of an example due to N.N. Nekhoroshev [7]). Let

$$H_{00}(I_1) + \varepsilon H_{01}(I_2) := \frac{I_1^2}{2} - \varepsilon \frac{I_2^2}{2} \ , \qquad (1.3)$$

and notice that (the first inequality in) condition (1.2) is violated on each energy level crossing the axis $\{I_1 = 0\}$ (in particular is violated at $E = 0$). Then, one

---

[1] Arnold calls it "topological instability" and the conjecture is the following: near an arbitrary point in phase space there are trajectories along which the action variables undergo a displacement of order one (i.e., independently of ε) in a finite (albeit exponentially long) time.

[2] This happens, for example, in three body problems or in the D'Alembert planetary model. In Appendix A below the D'Alembert planetary model is discussed in detail.

[3] Or, more precisely, that "the intermediate term H_{01} removes the degeneracy".

can construct a sequence $\varepsilon_j \downarrow 0$ and a sequence of perturbations $H_{1,j}(\varphi)$ with $\sup\limits_{|\mathrm{Im}\varphi_i| \le 1} |H_{1,j}(\varphi)|$ uniformly bounded such that

$$I_\varepsilon(t) := e^{-1/\sqrt{\varepsilon}} \left(-\varepsilon^2 \, t, \varepsilon^{3/2} \, t \right) , \quad \varphi_\varepsilon(t) := e^{-1/\sqrt{\varepsilon}} \left(-\varepsilon^2 \, \frac{t^2}{2}, -\varepsilon^{5/2} \, \frac{t^2}{2} \right) , \quad (1.4)$$

is a solution of the Hamilton equation associated to $H_{00}(I_1) + \varepsilon H_{01}(I_2) + \varepsilon^2 H_{1,j}(\varphi)$ when $\varepsilon = \varepsilon_j$. In fact, it is enough to take

$$\varepsilon_j := \frac{1}{j^2} , \qquad H_{1,j}(\varphi) := e^{-j} \sin(\varphi_1 - j\varphi_2) .$$

Notice that a displacement of order one of the action variables $I_\varepsilon(t)$ with respect to their initial value $I_\varepsilon(0) = (0,0)$ occurs in the exponentially long time $\sim \exp(1/\sqrt{\varepsilon_j})/\varepsilon_j^2$.

(ii) Condition (1.2) is violated, at $E = 0$, also by the "convex" Hamiltonian $H_0 := \frac{I_1^2}{2} + \varepsilon \frac{I_2^2}{2}$, $(\varepsilon > 0)$. However, in such a case, $H_0^{-1}(0)$ consists only of one point and exploiting convexity (and using energy conservation arguments), it is not difficult to show that, also on the energy level $E = 0$, total stability holds for $\varepsilon > 0$ small enough. It is therefore clear that "convexity" (or, more in general, "steepness") should play a fundamental role in this business.

Properly degenerate systems with two degrees of freedom of the form (1.1), are, in general, "more integrable" than non–degenerate systems, as A.I. Nejshtadt proved in 1981:

THEOREM 1.2. ([6]) *Assume that a (real–analytic) properly degenerate system with two degrees of freedom satisfies condition (1.2) together with $\frac{\partial H_{01}}{\partial I_2} \ne 0$. Then the measure of the set of unperturbed tori that disappear when $\varepsilon > 0$ is exponentially small (i.e. $O(\exp(-\mathrm{const}/\varepsilon)$ rather than $O(\sqrt{\varepsilon})$ as in general nondegenerate systems). Furthermore the deviation of a perturbed torus from the unperturbed one is of $O(\varepsilon)$ (rather than $O(\sqrt{\varepsilon})$).*

In this paper we take up the action–stability problem for properly degenerate Hamiltonian system with two degrees of freedom *allowing the intermediate system H_{01} to depend also on the angle φ_1.* Thus, we shall consider real–analytic, properly–degenerate systems with two degrees of freedom described by nearly–integrable, real–analytic Hamiltonians given by

$$H(I,\varphi;\varepsilon) := H_{00}(I_1) + \varepsilon H_{01}(I,\varphi_1) + \varepsilon^a H_1(I,\varphi) , \qquad 0 < \varepsilon \ll 1 , \qquad a > 1 . \quad (1.5)$$

The interest for such systems stems again from Celestial Mechanics. For example, the "planetary D'Alembert model" describing the motion of a nearly spherical planet subject to the gravitational attraction of a fixed star occupying a focus of a Keplerian nearly circular ellipse along which the centre of mass of the planet revolves, is governed,*up to an exponentially small term*, by a Hamiltonian of the form

$$\begin{aligned} H_D(I_1,I_2,\varphi_1,\varphi_2;\varepsilon,\mu) := \frac{I_1^2}{2} + \varepsilon \Big(\hat{c}_0(I_1,I_2) + \hat{d}_1(I_1,I_2)\cos\varphi_1 \Big) \\ + \varepsilon^a G(I_1,I_2,\varphi_1,\varphi_2;\varepsilon,\mu) , \end{aligned} \quad (1.6)$$

where: ε and $\mu \le \varepsilon^c$ (with $c > 1/2$) are perturbation parameters (related, respectively, to the "oblateness" of the planet and to the eccentricity of the Keplerian orbit); $a \in (3/2, 2]$; $\hat{c}_0$, $\hat{d}_1$ and G are given real–analytic functions uniformly bounded in suitable analytic norms; see Appendix A for a full description of this model.

REMARK 2. (i) In the above D'Alembert model, the "intermediate" system is given by $H_0 := \frac{I_1^2}{2} + \varepsilon \hat{c}_0(I_1, I_2)$. It turns out that in physically interesting phase space regions[4] H_0 *is non convex*. For this reason, below, we shall consider also non convex models.

(ii) The planetary D'Alembert model motivated in [4] new investigations about action–instability ("Arnold Diffusion"). For such studies, of course, exponentially small terms cannot be disregarded. In relation with the (full) D'Alembert problem, the results presented here go in the direction of giving action–stability bounds for *exponentially long times*. Such bounds would not immediately follow from standard Nekhoroshev techniques because of the strong degeneracies of the model[5].

(iii) The dependence of H_{01} upon the angle φ_1 (that is, on the angle conjugated to the non–degenerate action I_1), besides being motivated by classical examples, is the only significative angle–dependence one wants to take into account in connection with the problems considered here. In general, in fact, a Hamiltonian function of the form $H_{00}(I_1) + \varepsilon H_{01}(I, \varphi_2) + \varepsilon^2 H_1(I, \varphi)$ will be trivially unstable as the following example shows. Let $H_{01} = \frac{I_2^2}{2} - (1 + \cos \varphi_2)$ and $H_1 = 0$. Then, one has $\sup_t |I_2(t) - I_2(0)| = 2$, for any $\varepsilon > 0$ and for any motion with $(I_2(0), \varphi_2(0))$ belonging to the (open) separatrix of the pendulum H_{01}. Moreover, these hyperbolic motions would be persistent under non–vanishing perturbations H_1.

The Hamiltonian

$$H_0(I, \varphi_1; \varepsilon) := H_{00}(I_1) + \varepsilon H_{01}(I, \varphi_1) , \qquad (1.7)$$

regarded as a one–degree–of–freedom system in the (I_1, φ_1) variables, is still integrable exhibiting, in general, the typical features of a one–degree–of–freedom dimensional system (phase space regions foliated by invariant circles of possibly different homotopy, stable/unstable equilibria, separatrices, etc.). A natural approach (which we shall, in fact, follow) is to introduce action–angle variables for the one–degree–of–freedom Hamiltonian $H_0(I, \varphi_1; \varepsilon)$ (regarding I_2 as a dumb parameter) and then to apply KAM techniques trying to confine all motions among KAM tori (as in the non–degenerate case). The problem with this approach is that the action–angle variable for the (I_1, φ_1) system are singular in any neighbourhood of the separatrix (and stable equilibria) and is exactly near separatrices where one expects the motion to become "chaotic" and where, in principle, drift of order one in the I_2 variable is conceivable[6] even in the two–degrees–of–freedom (properly degenerate) case considered here. Therefore a careful analysis near these "singular phase space regions" is needed and arguments different from KAM theory have to be used to control the displacement of the action variable in such singular regions. Clearly, as discussed in Remark 1, regions where the non–degeneracy assumption fails need a separate discussion: in fact, in such zones (and in the non convex case), we can not exclude a "possibly non–chaotic–drift" of the I_2 action.

---

[4]Such regions correspond to unperturbed situations in which the spin axis of the planet is nearly orthogonal to ecliptic plane (i.e., to the plane containing the Keplerian ellipse): this is the observed situation for most planets in the Solar system.

[5]For other investigations on exponential (Nekhoroshev) stability in Celestial Mechanics we refer, also, to [3].

[6]Better: "compatible with energy conservation".

To avoid "extra" technical difficulties, we shall consider, in this paper, *model problems*, namely, we shall let

$$H_{00} := \frac{I_1^2}{2} , \qquad H_{01} := H_{01}^{(\sigma)} := \sigma \frac{I_2^2}{2} - (1 + \cos \varphi_1) , \qquad (1.8)$$

with σ equal either $+1$ or -1; the phase space will be taken to be $\mathcal{M}_{R_0} := B_{R_0}^2 \times \mathbb{T}^2$ where $B_{R_0}^2$ denotes a ball of radius R_0 around the origin.

REMARK 3. These model problems are intended to capture the main features of "general" properly degenerate systems with two degrees of freedom and, in particular, the features of the exponential approximation (1.6) to the D'Alembert Hamiltonian. This is the reason for considering both the convex and the non convex case in (1.8), corresponding, respectively to $\sigma = 1$ and $\sigma = -1$ (compare, also, point (i) of Remark 2).

We can now state our main results. Denote, as above, by $(I(t), \varphi(t)) := \phi_H^t(I_0, \varphi_0)$ the time t evolution of the initial data $(I(0), \varphi(0)) := (I_0, \varphi_0)$ governed by the Hamiltonian H. We shall prove the following

THEOREM 1.3. *Let $H^{(\sigma)}(I, \varphi; \varepsilon) := H(I, \varphi; \varepsilon)$ and $\mathcal{M}_{R_0}$ be as in (1.5), (1.8). Assume $a > 3/2$ and choose*

$$0 < R < R_0 \qquad \text{and} \qquad 0 < b < \min \left\{ \frac{1}{4} , \frac{a-1}{4} , \frac{1}{3} \left(a - \frac{3}{2} \right) \right\} . \qquad (1.9)$$

Then, there exists $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, the ϕ_H^t–evolution $(I(t), \varphi(t))$ of an initial datum (I_0, φ_0) satisfies

$$|I(t)| < R_0 , \qquad |I(t) - I_0| < \varepsilon^b , \qquad \forall\, t \in \mathbb{R} , \qquad (1.10)$$

where, in the case $\sigma = 1$, (I_0, φ_0) is an arbitrary point in the phase space $\mathcal{M}_R$, while, in the "non–convex" case $\sigma = -1$, (I_0, φ_0) belongs to $\mathcal{M}_R \backslash \mathcal{N}_*$, $\mathcal{N}_*$ being an open region whose measure does not exceed $\varepsilon^{2/3}$.

This theorem will be a simple corollary of the following result, which describes the distribution and density of KAM tori. Let $H_{\rm p}$ denote the pendulum Hamiltonian[7]

$$H_{\rm p} := H_{\rm p}(I_1, \varphi_1; \varepsilon) := \frac{I_1^2}{2} - \varepsilon(1 + \cos \varphi_1) . \qquad (1.11)$$

THEOREM 1.4. *Let the hypotheses and choices of Theorem 1.3 hold and let $\mathcal{M}^{(\sigma)} := \mathcal{M}_R \backslash \mathcal{N}^{(\sigma)}$ where the sets $\mathcal{N}^{(\sigma)} := \mathcal{N}^{(\sigma)}(\varepsilon, b)$ are defined by*

$$\mathcal{N}^{(1)} := \left\{ (I, \varphi): \ |H_{\rm p}| < \varepsilon^{1+2b} \ \text{or} \ H_{\rm p} < -2\varepsilon + \varepsilon^{1+2b} \right\} \cup \left\{ (I, \varphi): \ |I_2| < R\varepsilon^b \right\}$$

$$\mathcal{N}_* := \left\{ (I, \varphi): \ c\, \varepsilon^{\frac{2}{3} + \frac{4}{3}b} < H_{\rm p} < \frac{\varepsilon^{\frac{2}{3}}}{c} \right\} ,$$

$$\mathcal{N}^{(-1)} := \mathcal{N}^{(1)} \cup \mathcal{N}_* , \qquad (1.12)$$

$0 < c < 1$ being a suitable constant. Fix q such that

$$0 < q < a - \frac{3}{2} - 3b . \qquad (1.13)$$

---

[7]$H_{\rm p}$ is a standard mathematical pendulum having the stable equilibrium in $(0,0)$ with energy -2ε, the unstable equilibrium in $(0, \pm\pi)$ with energy 0 (hence the separatrix as well has energy 0).

Then, there exists $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, the following holds. Apart from a small dense subset of measure $O(\exp(-1/\varepsilon^q))$, the region $\mathcal{M}^{(\sigma)}$, is filled up by two–dimensional, real–analytic $H^{(\sigma)}$–invariant tori; each of these tori is $O(\exp(-1/\varepsilon^q))$–close to an unperturbed torus $\{(I_1, \varphi_1) : H_{\mathrm{p}} = E\} \times \{(I_2, \varphi_2)$ s.t. $I_2 = const\}$ in $\mathcal{M}^{(\sigma)}$. Furthermore, for any motion $(I(t), \varphi(t))$ in $\mathcal{M}^{(\sigma)}$, the displacement of $I(t)$ from its initial value I_0 is bounded, for all times t, by $\sqrt{\varepsilon}$.

REMARK 4. (i) By simple energy–conservation argument one sees immediately that $|I_1(t) - I_1(0)| < \mathrm{const}\ \sqrt{\varepsilon}$ for any motion $(I(t), \varphi(t))$ in $\mathcal{M}_R$; thus the "stability" statement in (1.10) concerns actually only the I_2 action variable.

(ii) The discarded region $\mathcal{N}^{(\sigma)}$ is a ("elementary") set small with ε. If we replace $\mathcal{N}^{(\sigma)}$ by a small set of order one (say, $\{(I_1, \varphi_1) : |H_{\mathrm{p}}| < \delta\} \times \{(I_2, \varphi_2) : |I_2| < \delta\}$ for a fixed $0 < \delta \ll 1$), then the displacement of $I(t)$ from its initial value I_0 is bounded by ε.

(iii) In the two–degrees–of–freedom case considered here, as mentioned above, the 2-dimensional KAM tori constructed in Theorem 1.4 (which fill, up to an exponentially small set, the region $\mathcal{M}^{(\sigma)}$) *separate* the three–dimensional energy levels. Thus, the topological "trapping" argument may be applied leading to stability, for all times, of the action variables in $\mathcal{M}^{(\sigma)}$. Then, an elementary *energy–conservation argument* implies action stability in $\mathcal{M}_R$ or in $\mathcal{M}_R \backslash \mathcal{N}_*$ (according to whether $\sigma = 1$ or $\sigma = -1$).

(iv) In the case $a = 2$ one can take any $0 < b < 1/6$ and $q < \frac{1}{2} - 3b$.

(v) Theorem 1.3 and Theorem 1.4 may be viewed as extensions, in the model cases considered here, of, respectively, Theorem 1.1 and 1.2.

We close this introduction with a list of problems.

(1) Generalise Theorem 1.3 and 1.4 to the Hamiltonian H_D (see Appendix A for a full description of H_D) and deduce exponential stability estimates for the full D'Alembert planetary Hamiltonian.

(2) Find general conditions on H_0 under which Theorems 1.3 and 1.4 hold.

(3) Extend the example in Remark 1 to ε–independent perturbations H_1. Extend the example in Remark 1 to H_0 dependent also on φ_1.

(4) The examples in (3) may indicate a possible route to $O(1)$–drift of action variables, in properly degenerate systems, *different* from Arnold Diffusion.

The paper is organized as follows. In §2 we list the technical tools we need in order to prove Theorems 1.3 and 1.4, namely: a quantitative, accurate discussion of the real–analytic extension of action–angle variable for the pendulum with particular care to singular regions; an "averaging" or "normal form" lemma (standard in Nekhoroshev theory); a quantitative iso–energetically KAM theorem. In §3, the proofs of the Theorem 1.3 and 1.4 are given. In Appendix A we discuss the D'Alembert planetary model and show how averaging theory may be used to reduce it, up to an exponentially small term, to the form in (1.6). In Appendix B the (lengthy but elementary) details for the construction of the real–analytic action–angle variables for the pendulum are provided.

2. **Preliminaries.** The construction of KAM tori in $\mathcal{M}^{(\sigma)}$ is based on the following three lemmata: the first lemma provides (real–analytic) action–angle variable for the pendulum slightly away from the separatrix and the stable equilibrium; the second lemma is a "normal form lemma"; the third lemma is a "iso–energetic" KAM theorem. For general information about normal forms, KAM theory, etc, we refer to, e.g., [1] and references therein.

In the following we shall use the following notations: if $A \subset \mathbb{R}^d$ and $r > 0$, we denote by A_r the subset of points in $\mathbb{C}^d$ at distance less than r from A; $\mathbb{T}_s^d$ denotes the complex set $\{z \in \mathbb{C}^d : |\mathrm{Im} z_j| < s \text{ for all } j\}$ (thought of as a complex neighborhood of $\mathbb{T}^d$). If $f(I, \varphi)$ is a real analytic function on $A_r \times \mathbb{T}_s^d$ we let $\|f\|_{r,s}$ denote the following norm[8]

$$\|f\|_{r,s} := \sum_{k \in \mathbb{Z}^d} \sup_{I \in A_r} |\hat{f}_k(I)| e^{|k|s} , \qquad (2.14)$$

$\hat{f}_k(I)$ being the Fourier coefficients of the periodic function $\varphi \to f(I, \varphi)$.

LEMMA 2.1 (Real–analytic action–angle variables for the pendulum).
Let $D^0 := [-R_0, R_0]$, let $E_0 := H_{\mathrm{p}}(R_0, 0) = R_0^2/2$, let $0 < \eta < \varepsilon/32$ and define

$$\mathcal{M}_p^+ := \mathcal{M}_p^+(\eta, \varepsilon) := \left\{ (I_1, \varphi_1) \in D^0 \times \mathbb{T} : I_1 > 0 , \eta < H_{\mathrm{p}}(I_1, \varphi_1) < E_0 - \varepsilon \right\}$$

$$\mathcal{M}_p^- := \mathcal{M}_p^-(\eta, \varepsilon) := \left\{ (I_1, \varphi_1) \in D^0 \times \mathbb{T} : -2\varepsilon + \eta < H_{\mathrm{p}}(I_1, \varphi_1) < -\eta \right\} . \quad (2.15)$$

Then, for all $r_* < R_0/2$ and s_* positive, there exist positive numbers r_0 and s_0, closed intervals $D^\pm \subset \mathbb{R}$, symplectic transformations $\phi^\pm$ real–analytic on $D^\pm \times \mathbb{T}$ and functions $h^\pm$ real–analytic on $D^\pm$ such that

$$\phi^\pm : (\hat{I}_1, \hat{\varphi}_1) \in D_{r_0}^\pm \times \mathbb{T}_{s_0} \to \phi^\pm(\hat{I}_1, \hat{\varphi}_1) \in D_{r_*}^0 \times \mathbb{T}_{s_*} , \qquad (2.16)$$

$$\phi^\pm(D^\pm \times \mathbb{T}) = \mathcal{M}^\pm(\eta, \varepsilon) , \qquad (2.17)$$

$$H_{\mathrm{p}} \circ \phi^\pm(\hat{I}_1, \hat{\varphi}_1) = h^\pm(\hat{I}_1) , \quad \forall (\hat{I}_1, \hat{\varphi}_1) \in D_{r_0}^\pm \times \mathbb{T}_{s_0} . \qquad (2.18)$$

The analyticity radii r_0 and s_0 may be taken to be

$$r_0 := c \, r_* \, \frac{\eta}{\sqrt{\varepsilon}} , \qquad s_0 := c \, s_* \, \frac{1}{\ln(\varepsilon/\eta)} , \qquad (2.19)$$

where $0 < c < 1$ is a suitable (universal) constant. Furthermore, the functions $h^\pm$ satisfy, for all $\hat{I}_1 \in D_{r_0}^\pm$, the following bounds

$$\eta \le \mathrm{Re}\, h^+(\hat{I}_1) \le E_0 - \varepsilon , \qquad -2\varepsilon + \eta \le \mathrm{Re}\, h^-(\hat{I}_1) \le -\eta , \quad (2.20)$$

$$\frac{dh^\pm}{d\hat{I}_1}(\hat{I}_1) = \frac{\alpha^\pm}{\pi_1^\pm} , \qquad (2.21)$$

$$\frac{d^2 h^\pm}{d\hat{I}_1^2}(\hat{I}_1) = \pm \beta^\pm \, \frac{\pi_2^\pm}{(\pi_1^\pm)^3} , \qquad (2.22)$$

---

[8] The specific choice of norm will play no role in the sequel; obviously if f is a real–analytic function on $\mathbb{T}_s^d$, $\|f\|_s$ stands for $\sum_{k \in \mathbb{Z}^d} |\hat{f}_k| e^{|k|s}$, $\hat{f}_k$ being the Fourier coefficients of f, while, if f is a real–analytic function on A_r, then $\|f\|_r = \sup_{I \in A_r} |f(I)|$.

where

$$\pi_1^{\pm} \quad := \quad \pi_1^{\pm}(\hat{I}_1) := \frac{1}{\sqrt{\varepsilon}} \, \ln \left(1 + \sqrt{\frac{\varepsilon}{|\operatorname{Re} h^{\pm}(\hat{I}_1)|}} \right) \, ,$$

$$\pi_2^{\pm} \quad := \quad \pi_2^{\pm}(\hat{I}_1) := \frac{1}{|\operatorname{Re} h^{\pm}(\hat{I}_1)|} \, \frac{1}{\sqrt{|\operatorname{Re} h^{\pm}(\hat{I}_1)| + \varepsilon}} \, , \qquad (2.23)$$

$\alpha^{\pm} := \alpha^{\pm}(\hat{I}_1)$, $\beta^{\pm} := \beta^{\pm}(\hat{I}_1)$ *are real–analytic functions such that*

$$d_1 \leq \frac{\operatorname{Re}(\alpha^{\pm})}{R_0}, \, \operatorname{Re}(\beta^{\pm}) \leq d_2 \, ; \qquad \left| \frac{\operatorname{Im}(\alpha^{\pm})}{R_0} \right|, |\operatorname{Im}(\beta^{\pm})| \leq d_0 \qquad (2.24)$$

for suitable (universal) constants $0 < d_1 < d_2$ and $0 < d_0 < d_1/10$. An identical statement holds if, in the definition of $\mathcal{M}_p^+$, one replaces "$I_1 > 0$" with[9] "$I_1 < 0$".

The next two lemmata are typical statements from KAM theory (see [1] for generalities). The first one is a "normal form lemma" common, also, in averaging theory. The second one is an iso–energetic KAM theorem (i.e. a KAM theorem on fixed energy levels).

The only (technical) difference in the statement of the normal form lemma is that we have to allow *different* radii of analyticity in the action–variables (a fact that is convenient for our application of the KAM theorem; see, also, point (ii) of Remark 6 below). For notational simplicity we state the normal form lemma for $d = 2$ (which suffices for our applications).

LEMMA 2.2 (Normal forms). *Let $\widehat{D}$ and $\widehat{D}'$ be two subsets of $\mathbb{R}$ and consider a Hamiltonian function $H(\hat{I}, \hat{\varphi}) := h(\hat{I}) + f(\hat{I}, \hat{\varphi})$ real–analytic on $\widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}} := (\widehat{D}_{\hat{r}_1} \times \widehat{D}'_{\hat{r}_2}) \times \mathbb{T}_{\hat{s}}^2$ for some $\hat{r}_2 \geq \hat{r}_1 > 0$ and $\hat{s} > 0$. Assume that there exist $K \geq 6/\hat{s}$ and $\alpha > 0$ such that*

$$|\omega(\hat{I}) \cdot k| \geq \alpha \, , \qquad \forall \, k \in \mathbb{Z}^2 \, , \, 0 < |k| \leq K \, , \quad \forall \, \hat{I} \in \widehat{D}_{\hat{r}_1} \times \widehat{D}'_{\hat{r}_2} \, , \qquad (2.25)$$

where $\omega(\hat{I}) := \nabla h(\hat{I})$. Assume also that[10]

$$\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \leq \frac{\alpha \hat{r}_1}{2^8 K} \, . \qquad (2.26)$$

Then, there exist a real–analytic symplectic transformation

$$\Phi : (J, \psi) \in \widehat{W}_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \rightarrow \Phi(J, \psi) \in \widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}}$$

such that

$$H \circ \Phi(J, \psi) = h(J) + g(J) + f_*(J, \psi) \qquad (2.27)$$

with[11]

$$\|g - f_0\|_{\hat{r}_1/2, \hat{r}_2/2} \leq \frac{2^6 K}{\alpha \hat{r}_1} \, (\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}})^2 \leq \frac{1}{4} \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \, ,$$

$$\|f_*\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \, \exp(-K\hat{s}/6) \, ,$$

$$\|\Phi(J, \psi) - (J, \psi)\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \hat{c} \, \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \, , \qquad (2.28)$$

---

[9] By symmetry, the interval D^+ in the case $I_1 < 0$ is just the opposite of the interval D^+ in the case $I_1 > 0$.

[10] Adapt the norms in (2.14) and in the footnote 8 in the obvious way replacing A_r by $\widehat{D}_{r_1} \times \widehat{D}'_{r_2}$ (and replacing the subscript "r" in the norms by "$\hat{r}_1, \hat{r}_2$").

[11] f_0 is the zero–Fourier coefficient of f, i.e., the average of $f(\hat{I}, \hat{\varphi})$ over $\mathbb{T}^2$.

where $\hat{c} > 0$ is a suitable constant.

Let, as above, $\omega(J)$ denote the gradient $\nabla h(J)$, let $h''(J)$ denote the Hessian matrix of h. We recall that a vector $\omega \in \mathbb{R}^d$ is said to be (γ, τ)–Diophantine if

$$|\omega \cdot k| \geq \frac{\gamma}{|k|^\tau} , \qquad \forall \ k \in \mathbb{Z}^d \backslash \{0\} , \qquad (2.29)$$

for some $\gamma > 0$ and[12] $\tau > 0$.

LEMMA 2.3 (Iso–energetic KAM theorem). *Let $D \subset \mathbb{R}^d$ be a bounded domain and consider a Hamiltonian $H(J, \psi) := h(J) + f(J, \psi)$ real–analytic on the domain $W_{r,s}$ $:= D_r \times \mathbb{T}_s^d$ for some $r > 0$ and $s > 0$. Assume that $\|h''\|_r > 0$ and that the $(d + 1) \times (d + 1)$ matrix*

$$U := \begin{pmatrix} h''(J) & \omega(J) \\ \omega(J) & 0 \end{pmatrix} \qquad (2.30)$$

is invertible on D_r. Given $E \in \mathbb{R}$ (such that $h^{-1}(E) \neq \emptyset$) and given

$$0 < \gamma < \min_{i, J \in D} |\omega_i(J)| \qquad \text{and} \quad \tau \geq d - 1 , \qquad (2.31)$$

denote

$$\mathcal{D} = \left\{ J \in D : \ h(J) = E \ \text{and} \ \omega(J) \ \text{is} \ (\gamma, \tau) - \text{Diophantine} \right\} . \qquad (2.32)$$

Then, if $\|f\|_{r,s}$ is small enough, for each $J \in \mathcal{D}$, there exists a unique d–dimensional, real–analytic, invariant torus $\mathcal{T} \subset H^{-1}(E)$ which is a graph over the angle ψ, which is close to the torus $\{J\} \times \mathbb{T}^d$ and on which the H–flow is analytically conjugated to the translation $\theta \to \theta + \omega(J)(1 + \kappa)t$, κ being a small real number. More precisely, let A, F and G be positive numbers such that

$$A \geq \|h''\|_r , \qquad F \geq A \ \|f\|_{r,s} \ \gamma^{-2} , \qquad G \geq \max \left\{ A \ \|U^{-1}\|_r , \ 1 \right\} , \qquad (2.33)$$

let $0 < \bar{s} < s$ and let

$$C := \max \left\{ 1 , \ \frac{\gamma \ (s - \bar{s})^{c_1}}{c_2 A \ r \ |\ln F|^{c_3}} \right\} , \qquad \widehat{F} := C \ c_4 \ \frac{1}{(s - \bar{s})^{c_5}} \ G^{c_6} F |\ln F|^{c_7} , \qquad (2.34)$$

where the $c_i > 1$ are suitable constants depending only upon τ and d. If $\widehat{F} \leq 1$, then, for each $J \in \mathcal{D}$, there exists a unique invariant torus $\mathcal{T} \subset H^{-1}(E)$ satisfying the following properties:

(i) $\mathcal{T} = \left\{ (\mathcal{J}(\psi), \psi) : \ \psi \in \mathbb{T}^d \right\}$ with $\mathcal{J}$ real–analytic on $\mathbb{T}_{\bar{s}}^d$ and $|\mathcal{J}(\psi) - J| \leq r \widehat{F}$ for all $\psi \in \mathbb{T}_{\bar{s}}^d$;

(ii) there exist real–analytic functions on $\mathbb{T}_{\bar{s}}^d$, u, v and a smooth function $\kappa : D_r \to \mathbb{C}$ (real for real J) such that

$$\max\{r^{-1}\|v - J\|_{\bar{s}}, \ \|u\|_{\bar{s}}, \ |\kappa|\} \leq \widehat{F} ;$$

the map $\theta \in \mathbb{T}_{\bar{s}}^d \to \left(v(\theta), \theta + u(\theta) \right)$ is a real–analytic embedding whose real image is the torus $\mathcal{T}$: $\mathcal{T} = \left\{ \left(v(\theta), \theta + u(\theta) \right), \ \theta \in \mathbb{T}^d \right\}$; on the torus $\mathcal{T}$ the H–flow, ϕ^t, linearizes: denoting $\omega_* := \left(1 + \kappa(J) \right) \omega(J)$, one has

$$\phi^t \left(v(\theta), \theta + u(\theta) \right) = \left(v(\theta + \omega_* t), \theta + \omega_* t + u(\theta + \omega_* t) \right) ;$$

---

[12]Necessarily $\tau \geq d - 1$ by a theorem of Liouville. Also, (2.29) (with $|k| = 1$) implies that $\gamma \leq \min_i |\omega_i|$.

(iii) if $\tau > d - 1$ and $\hat{\gamma} := \left(\text{const.}\dfrac{\|h'\|_r^d}{\min\limits_{D_r} |\det U|}\right) \gamma$, then

$$\text{meas}\Big(H^{-1}(E)\backslash\Big\{\text{tori satisfying } (i) \text{ and } (ii)\Big\}\Big) \leq \hat{\gamma} \ .$$

REMARK 5. As mentioned above, in the case of two degrees–of–freedom ($d = 2$) considered in this paper, the above KAM tori *separate* the three–dimensional energy levels forming *barriers* for the motion; any two KAM tori (with equal energy) bound an invariant region in corresponding energy level. More precisely, let $[a_1, b_1] \times [a_2, b_2] \subset D$ with $a_i < b_i$. Then, because of (2.31), we can take as coordinates for the three–dimensional energy level $H^{-1}(E)$ either of the action variables[13] plus the angles ψ. Take first as coordinates (J_1, ψ_1, ψ_2) and fix $\bar{J}_1 \in [a_1 + \delta, a_2 - \delta]$ where $\delta := 2\max\{r\widehat{F}, \hat{\gamma}\}$ ($\hat{\gamma}$ measures the complement of the surviving KAM tori and $r\widehat{F}$ the maximal oscillation of the graph of each KAM torus). Then, by (i) and (iii) in Lemma 2.3, it follows that there exist two tori $\mathcal{T}'$ and $\mathcal{T}''$ so that $\sup_\psi \mathcal{J}_1' < \bar{J}_2 < \inf_\psi \mathcal{J}_1''$ and $0 < \inf_\psi \mathcal{J}_1'' - \sup_\psi \mathcal{J}_1' \leq O(\delta)$. The same reasoning applies to $\bar{J}_2$. Hence, if $(J(t), \psi(t)) := \phi^t(\bar{J}, \bar{\psi})$ (for any $\bar{\psi}$) one has that $\sup_t |J(t) - \bar{J}| \leq O(\delta)$.

REMARK 6. **(On the proofs of the lemmata)**

(i) The action–angle variables for the pendulum $H_{\mathrm{p}}(I_1, \varphi_1; \varepsilon) = \frac{I_1^2}{2} - \varepsilon(1 + \cos\varphi_1)$ are produced by the generating function $\displaystyle\int_{\Gamma(\varepsilon, E)} I_1 d\varphi_1$, where $\Gamma(\varepsilon, E)$ denotes the positively oriented circle $H_{\mathrm{p}}^{-1}(E)$ (the homotopy of $\Gamma(\varepsilon, E)$ depends on whether $E > 0$ or $-2\varepsilon < E < 0$). The point is that we need a very detailed and quantitative analysis for Γ *very* close to the separatrix (i.e. E close to 0) and for Γ close to the stable equilibrium (i.e. E close to -2ε) regions where the action–angle variables become singular; "very close" meaning, here, "at a distance of order ε^β with $\beta > 1$". Therefore, in such "singular" regions, a careful "blow–up" analysis is needed. Furthermore we also need to study the *complex* analytic continuation of the action-angle variables since we want to apply a KAM theorem in real-analytic class. To perform this blow–up in analytic class a certain amount of straightforward (although rather lengthy) computations are needed: we provide details in Appendix B.

We mention also that for our main purpose (i.e., total stability of action variables) it would be enough to apply a iso–energetic KAM theorem in *smooth* class (since all we need is a topological "trapping argument"); however a *quantitative* version of such a theorem (necessary for our task) is not available in literature and providing the details for its proof would be certainly much longer (and far less elementary) than the proof of Lemma 2.1.

(ii) Lemma 2.2, as mentioned above, is a standard "normal form lemma"; a proof may be found, e.g., in [9], pag 192. Keeping track of different radii (going into the proof in, e.g., [9]) is routine (notice that in the "smallness condition" (2.26) there appears the *smallest* radius). We add only a technical comment: in [9] there appears the condition $r \leq \alpha/(\text{const}K)$; such a condition is needed to control the

---

[13]Furthermore, the map $J_1 \rightarrow \alpha_1(J_1) = \omega_1(J_1, J_2^0(J_1))/\omega_2(J_1, J_2^0(J_1))$, where J_2^0 is such that $h(J_1, J_2^0(J_1)) = E$, is a diffeomorphysm:

$$\frac{d\alpha_1}{dJ_1} = -\left.\frac{\det U}{\omega_2^3}\right|_{(J_1, J_2^0(J_1))} \ ;$$

(and a completely symmetric statement holds interchanging the indices 1 and 2).

small divisor bounds on complex domains. Since *we are assuming* the small divisor bounds *directly* on complex domains such a condition is not needed in our case.

(iii) Also Lemma 2.3 is by now rather standard. In fact it is easy, under an extra "nondegeneracy condition" satisfied in our application[14], to derive the iso–energetic KAM theorem directly by the standard one by means of a standard Implicit Function Theorem. Alternatively, one can find a very detailed version, e.g., in [5]. For these reasons we shall omit the proof of Lemma 2.3. In our application the exact values of the constants c_i are not needed; however we can prove Lemma 2.3 with the following constants:

$$c_1 = \tau + 1 \ , \quad c_2 = 2 \cdot 6^{\tau+1} \ , \quad c_3 = c_1 \ , \quad c_4 = d \, 2^{10} \ ,$$
$$c_5 = 2(\tau + 1) \ , \quad c_6 = 2 \ , \quad c_7 = 2(\tau + 1) \ .$$

Also, in our case, it will be $C = 1$.

3. **Proofs of the Theorems.** We first prove Theorem 1.4 (Theorem 1.3 will be a simple corollary of it). Since most of the arguments are identical for both models $\sigma = 1$ and $\sigma = -1$, we shall usually do not indicate explicitly the dependence upon σ. The only point where the two models differ is in the estimates regarding the iso–energetical non–degeneracy (see Lemma 3.1 below).

Proof of Theorem 1.4

The *first step* is to use Lemma 2.1 to put H_0 in (1.7) into action–angle variables. Let R be as in (1.9) and *assume that H_1 in (1.5) is analytic on $B_{r_1} \times \mathbb{T}_{s_1}$ where B denotes here $B_{R_0}^2(0)$, and $0 < r_1 < R/2$, $s_1 > 0$.* Since, in our case, H_{p} is an entire function we can choose, in Lemma 2.1 the parameters

$$r_* := r_1 \ , \qquad s_* := s_1 \ . \tag{3.35}$$

Let b and q be as in (1.9) and (respectively) (1.13), let

$$\lambda = 1 + 2b \ , \tag{3.36}$$

and let q_0 be a number such that

$$q < q_0 < a - \frac{3}{2} - 3b \ . \tag{3.37}$$

Notice that with such choices the following relations hold:

$$\lambda > 1 \ , \qquad 0 < b < \lambda - \frac{1}{2} \ , \qquad b + \lambda + q_0 + \frac{1}{2} < a \ . \tag{3.38}$$

We also set

$$\eta := \varepsilon^\lambda \tag{3.39}$$

so that r_0 and s_0 in Lemma 2.1 become

$$r_0 = c \, r_1 \, \varepsilon^{\lambda - 1/2} \ , \qquad s_0 = \frac{c}{\lambda - 1} \, \frac{1}{\ln \varepsilon^{-1}} \, s_1 \ . \tag{3.40}$$

Let D^0, $D^\pm$ and $\phi^\pm$ be as in Lemma 2.1 and let

$$D := [-R_1, R_1] \subset D^0 \ , \qquad R_1 := \frac{R_0 + R}{2} \ . \tag{3.41}$$

Now, define $D^\pm(\sigma) \subset \mathbb{R}$ as follows:

$$D^-(\sigma) := D^- \ , \qquad D^+(1) := D^+ \ , \qquad D^+(-1) \times \mathbb{T} := (\phi^+)^{-1}(\mathcal{M}_*^+) \ , \tag{3.42}$$

---

[14]Namely, the invertibility of the Hessian h'' on D_r, which is the usual nondegeneracy condition in the standard KAM theorem.

where

$$\mathcal{M}_*^+ := \mathcal{M}_*^+(\eta,\varepsilon) := \mathcal{M}_p^+ \backslash \mathcal{R}_* := \mathcal{M}_p^+ \backslash \left\{ (I_1,\varphi_1): \ c\,\varepsilon^{\frac{2}{3}+\frac{4}{3}b} \le H_{\mathrm p} \le \frac{\varepsilon^{\frac{2}{3}}}{c} \right\}, \quad (3.43)$$

with a suitable small positive constant c to be fixed later. Denoting $J = (J_1, J_2)$, $\psi = (\psi_1, \psi_2)$, $\hat{I} = (\hat{I}_1, \hat{I}_2)$, $\hat{\varphi} = (\hat{\varphi}_1, \hat{\varphi}_2)$, then, by Lemma 2.1, we have

$$\hat{\phi}^\pm : (\hat{I}, \hat{\varphi}) \in (D^\pm(\sigma)_{r_0} \times D_{r_1}) \times \mathbb{T}_{s_0}^2 \to (I,\varphi) \in B_{r_1} \times \mathbb{T}_{s_1}^2 \ ,$$

$$\text{where } (I_1,\varphi_1) := \phi^\pm(\hat{I}_1, \hat{\varphi}_1) \ , \qquad (I_2,\varphi_2) := (\hat{I}_2, \hat{\varphi}_2) \ . \qquad (3.44)$$

In the symplectic coordinates $(\hat{I}, \hat{\varphi})$ the Hamiltonian H in (1.5) takes the form

$$H^\pm(\hat{I}, \hat{\varphi}; \varepsilon) := H \circ \hat{\phi}^\pm(\hat{I}, \hat{\varphi}) = h^\pm(\hat{I}_1) + \varepsilon\sigma \frac{\hat{I}_2^2}{2} + \varepsilon^a H_1^\pm(\hat{I}, \hat{\varphi}; \varepsilon) \ , \qquad (3.45)$$

where $h^\pm$ is as in Lemma 2.1 and $H_1^\pm := H_1 \circ \hat{\phi}^\pm$; hence

$$\|H_1^\pm\|_{r_0, r_1, s_0} \le \|H_1\|_{r_1, s_1} \ . \qquad (3.46)$$

The *second step* is to apply the normal form lemma (Lemma 2.2), in a suitable phase space region, to the Hamiltonian $H^\pm$: in such a way we shall be able to to put $H^\pm$ in a normal form of the type appearing in (2.27)–(2.28), to meet the (stringent) KAM condition, $\widehat{F} \le 1$, in the KAM theorem (Lemma 2.3) and to give a "good" estimates on the measure of the KAM tori. We therefore set[15]

$$h(\hat{I}) := h^\pm(\hat{I}_1) + \varepsilon\sigma \frac{\hat{I}_2^2}{2} \ , \quad f := \varepsilon^a H_1^\pm \ ,$$

$$\hat{r}_1 := r_0 = c r_1 \varepsilon^{\lambda - \frac{1}{2}} \ , \quad \hat{r}_2 := \frac{R_1}{10}\varepsilon^b \ , \quad \hat{s} := s_0 = \frac{c}{\lambda - 1}\,\frac{1}{\ln\varepsilon^{-1}}\,s_1 \ ,$$

$$\widehat{D} := D^\pm(\sigma) \ , \quad \widehat{D}' := \{ \hat{I}_2 \in \mathbb{R}: \ R_1\varepsilon^b \le |\hat{I}_2| \le R_1 \} \ ,$$

$$\widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}} := \left\{ \hat{I} \in D^\pm(\sigma)_{\hat{r}_1} \times \widehat{D}'_{\hat{r}_2} \right\} \times \mathbb{T}_{\hat{s}}^2 \ . \qquad (3.47)$$

Notice that the second relation in (3.38) implies that $\hat{r}_1 \ll \hat{r}_2$ for ε small. Define also

$$K := \frac{1}{\varepsilon^{q_0}\ln\varepsilon^{-1}} \ , \qquad (3.48)$$

where q_0 is as in (3.38). Let us, now, estimate α in (2.25). Denote by $\omega^\pm(\hat{I}) := \left((h^\pm)'(\hat{I}_1), \varepsilon\sigma\hat{I}_2\right)$. Then, for any $k \in \mathbb{Z}^2\backslash\{0\}$ with $|k| \le K$, by (2.20)–(2.24) and the choice of η, we find

$$|\omega^\pm(\hat{I}) \cdot k| \ge \begin{cases} |(h^\pm)'| - 2\varepsilon R_1 K \ge \kappa_1 R_1 \dfrac{\sqrt{\varepsilon}}{\ln\varepsilon^{-1}} \ , & \text{if } k_1 \ne 0 \ , \\[2ex] \dfrac{R_1}{2}\varepsilon^{b+1} \ , & \text{if } k_1 = 0 \ , \end{cases} \qquad (3.49)$$

for a suitable constant[16] κ_1 and provided $\varepsilon > 0$ is small enough . We can therefore take

$$\alpha := \frac{R_1}{2}\,\varepsilon^{1+b} \ . \qquad (3.50)$$

We can now check (2.26). Since, by (3.46),

$$\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} := \varepsilon^a \|H_1^\pm\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \le \varepsilon^a \|H_1\|_{r_1, s_1} \ , \qquad (3.51)$$

---

[15]$I_2 = 0$ is a singularity (resonance): we therefore have to stay a bit away from it.

[16]From here on, κ_i denote suitable constants depending, possibly, on λ, a, b, c, q_i, s_1 and r_1.

because of the choices of α, $\hat{r}_1$, $\hat{s}$ and K (see (3.50), (3.40), we find (2.19), (3.39) and (3.48)),

$$\frac{\alpha \hat{r}_1}{2^8 K} = \frac{c \, R_1 r_1 \, \varepsilon^{b+\lambda+q_0+1/2} \ln \varepsilon^{-1}}{2^9} \ . \tag{3.52}$$

Thus, in view of the choice of the various parameters made in (3.38), (2.26) is satisfied for $\varepsilon > 0$ small enough. Thus, by Lemma 2.2, there exist a real–analytic symplectic transformation

$$\Phi^{\pm} : (J, \psi) \in \widehat{W}_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \rightarrow \Phi^{\pm}(J, \psi) \in \widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}} \tag{3.53}$$

such that

$$H^{\pm} \circ \Phi^{\pm}(J, \psi) = h^{\pm}(J_1) + \varepsilon \sigma \, \frac{J_2^2}{2} + g^{\pm}(J) + H_*^{\pm}(J, \psi) \tag{3.54}$$

with (recall (2.28), (3.51), (3.48))

$$\|g^{\pm} - \varepsilon^a (H_1^{\pm})_0\|_{\hat{r}_1/2, \hat{r}_2/2} \leq \frac{1}{4} \, \varepsilon^a \, \|H_1\|_{r_1, s_1} \ ,$$

$$\|H_*^{\pm}\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \, \exp(-K\hat{s}/6) \ \leq \|H_1\|_{r_1, s_1} \, \exp\left(\frac{-\kappa_2}{\varepsilon^{q_0} (\ln \varepsilon^{-1})^2}\right) \ ,$$

$$\|\Phi^{\pm}(J, \psi) - (J, \psi)\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \hat{c} \, \varepsilon^a \, \|H_1\|_{r_1, s_1} \ , \tag{3.55}$$

for a suitable $\kappa_2 > 0$ (and ε small enough). Thus, if we pick a q_1 so that

$$q < q_1 < q_0 \ , \tag{3.56}$$

we have that, for all $\varepsilon > 0$ small enough,

$$\|H_*^{\pm}\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \|H_1\|_{r_1, s_1} \, \exp\left(-\frac{1}{\varepsilon^{q_1}}\right) \ . \tag{3.57}$$

Third step. In order to apply the KAM theorem (Lemma 2.3) we set:

$$h(J) = h^{\pm}(J_1) + \varepsilon \sigma \frac{J_2^2}{2} + g^{\pm}(J) := h_*^{\pm}(J) \ , \qquad f(J, \psi) = H_*^{\pm}(J, \psi) \ ,$$

$$r = \kappa_3 \, r_1 \, \varepsilon^{\lambda - \frac{1}{2}} \ , \qquad s = \kappa_3 \, s_1 \, \frac{1}{\ln \varepsilon^{-1}} \ , \qquad \bar{s} = \frac{s}{2} \ ,$$

$$D = D^{\pm}(\sigma) \times \widehat{D}' \ , \qquad W_{r,s} = D_r \times \mathbb{T}_s^2 \ , \tag{3.58}$$

where κ_3 is a suitable constant such that[17]

$$r \leq \frac{\hat{r}_1}{4} \ , \qquad s \leq \frac{\hat{s}}{6} \ .$$

Obviously the norm relative to the domain $W_{r,s}$ will again be denoted $\|\cdot\|_{r,s}$ but *beware that the sup–norms in the action variables are taken on different domains according to whether $\sigma = 1$ or $\sigma = -1$ (recall (3.42) and (3.43): in the case $\sigma = -1$ the set $\mathcal{R}_*$ has to be discarded)*. The estimates on $\|(h_*^{\pm})''\|$ and on[18] $\|U^{-1}\|$ require

---

[17]Recall (3.53) and that $\hat{r}_1 < \hat{r}_2$. The factor $1/4$ is included in order to bound derivatives of $g^{\pm}$ (and hence of h'') via Cauchy estimates. We recall the statement concerning Cauchy estimates in our context: if $g(J)$ is a function analytic on $D_r \times D_{r'}'$ then for any integers p_1, p_2 and for any $0 < c < 1$

$$\left\| \frac{\partial^{p_1 + p_2} g}{\partial J_1^{p_1} \partial J_2^{p_2}} \right\|_{cr, cr'} \leq \text{const.} \ (p_1! p_2!) \ \frac{\|g\|_{r,r'}}{r^{p_1} r'^{p_2}} \ \frac{1}{(1-c)^{p_1 + p_2}} \ .$$

[18]Recall the definition of the matrix U in Lemma 2.3.

some computations, which we collect in the following lemma. Recall that from (3.36) and (1.9) there follows that b and λ satisfy

$$b < \frac{1}{4} \ , \qquad \lambda < \frac{a+1}{2} \ . \tag{3.59}$$

LEMMA 3.1. *There exists*[19] $C_0 > 0$ *such that, for all $\varepsilon > 0$ small enough,*

$$\|(h_*^\pm)''\|_r \leq \frac{C_0}{\varepsilon^{\lambda-1}(\ln \varepsilon^{-1})^3} \ , \qquad \|U^{-1}\|_r \leq \frac{C_0}{\varepsilon^\lambda \ln \varepsilon^{-1}} \ , \tag{3.60}$$

where $\|\cdot\|_r$ denotes the sup–norm on D_r defined in (3.58), (3.42), (3.43), (3.47).

Proof. First, we need estimates on the derivatives of $g^\pm$. From (3.55) there follows $\|g^\pm\|_{\hat{r}_1/2,\hat{r}_2/2} \leq \frac{5}{4}\varepsilon^a \|H_1\|_{r_1,s_1}$; whence, by Cauchy estimates[20],

$$\left\|\frac{\partial g^\pm}{\partial J_1}\right\|_r \leq \kappa_5 \frac{\|H_1\|_{r_1,s_1}}{r_1} \varepsilon^{a-\lambda+\frac{1}{2}} \ , \quad \left\|\frac{\partial g^\pm}{\partial J_2}\right\|_r \leq \kappa_5 \frac{\|H_1\|_{r_1,s_1}}{r_1} \varepsilon^{a-b} \ ,$$

$$\left\|\frac{\partial^2 g^\pm}{\partial J_1^2}\right\|_r \leq \kappa_5 \frac{\|H_1\|_{r_1,s_1}}{r_1^2} \varepsilon^{a-2\lambda+1} \ , \quad \left\|\frac{\partial^2 g^\pm}{\partial J_2^2}\right\|_r \leq \kappa_5 \frac{\|H_1\|_{r_1,s_1}}{r_1^2} \varepsilon^{a-2b} \ ,$$

$$\left\|\frac{\partial^2 g^\pm}{\partial J_1 \partial J_2}\right\|_r \leq \kappa_5 \frac{\|H_1\|_{r_1,s_1}}{r_1^2} \varepsilon^{a-\lambda-b+\frac{1}{2}} \ , \tag{3.61}$$

with a suitable constant $\kappa_5 > 0$. By (3.38) and (3.59), one has

$$a - \lambda + \frac{1}{2} > 1 \ , \qquad a - b > \frac{3}{2} \ , \qquad a - 2\lambda + 1 > 0 \ ,$$

$$a - 2b > \frac{5}{4} \ , \qquad a - \lambda - b + \frac{1}{2} > 1 \ . \tag{3.62}$$

The symmetric matrix U has the form

$$U = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{12} & u_{22} & u_{23} \\ u_{13} & u_{23} & 0 \end{pmatrix} \tag{3.63}$$

where (recall (3.58), (3.54) and (2.18))

$$\begin{aligned}
u_{11} &= \frac{\partial^2 h_*^\pm}{\partial J_1^2} = (h^\pm)'' + \frac{\partial^2 g^\pm}{\partial J_1^2} \ , & u_{12} &= \frac{\partial^2 h_*^\pm}{\partial J_1 \partial J_2} = \frac{\partial^2 g^\pm}{\partial J_1 \partial J_2} \ , \\
u_{13} &= \frac{\partial h_*^\pm}{\partial J_1} = \frac{\partial h^\pm}{\partial J_1} + \frac{\partial g^\pm}{\partial J_1} \ , & u_{22} &= \frac{\partial^2 h_*^\pm}{\partial J_2^2} = \varepsilon\sigma + \frac{\partial^2 g^\pm}{\partial J_2^2} \ , \\
u_{23} &= \frac{\partial h_*^\pm}{\partial J_2} = \varepsilon\sigma J_2 + \frac{\partial g^\pm}{\partial J_2} \ .
\end{aligned} \tag{3.64}$$

Since (recall the estimates in Lemma 2.1)

$$C_1 \varepsilon^{\lambda-1} (\ln \varepsilon^{-1})^3 \leq \frac{(\pi_1^\pm)^3}{\pi_2^\pm} \leq C_2 \ , \quad C_1 \leq \pi_1^\pm \leq C_2 \frac{\ln \varepsilon^{-1}}{\sqrt{\varepsilon}} \ , \quad \frac{C_1}{\sqrt{\varepsilon}} \leq \pi_2^\pm \leq \frac{C_2}{\varepsilon^{\lambda+\frac{1}{2}}} \ , \tag{3.65}$$

---

[19]From here on, C_i denote suitable constants depending, possibly, on λ, n, a, b, q_i, E_0, R_1, r_1 and $\|H_1\|_{r_1,s_1}$.

[20]It is exactly in order to get the estimates (3.61) that we kept track of the different complex extension sizes in the variables J_1 and J_2.

for suitable constants $C_i > 0$, by (3.61) and (3.62), we see that there exists a $\bar{q}_1 > 0$ such that, for all $J_2 \in \widehat{D}'_r$, the following asymptotics hold[21]

$$u_{11} = \pm\beta^{\pm} \frac{\pi_2^{\pm}}{(\pi_1^{\pm})^3} \left(1 + O(\varepsilon^{\bar{q}_1})\right) , \quad u_{12} = O(\varepsilon^{\frac{1}{2}+\bar{q}_1}) , \quad u_{22} = \varepsilon\sigma \left(1 + O(\varepsilon^{\bar{q}_1})\right) ,$$

$$u_{13} = \frac{\alpha^{\pm}}{\pi_1^{\pm}} \left(1 + O(\varepsilon^{\bar{q}_1})\right) , \qquad\qquad u_{23} = \varepsilon\sigma J_2 \left(1 + O(\varepsilon^{\bar{q}_1})\right) . \tag{3.66}$$

From these relations there follows immediately that

$$\|(h_*)''\|_r \leq \frac{C_3}{\varepsilon^{\lambda-1}(\ln\varepsilon^{-1})^3} . \tag{3.67}$$

Let us, now, write the matrix U^{-1} as follows

$$U^{-1} = \frac{1}{\delta} \begin{pmatrix} u_1 & u_2 & u_3 + u_4 \\ u_2 & 1 & u_5 + u_6 \\ u_3 + u_4 & u_5 + u_6 & u_7 + u_8 \end{pmatrix} \tag{3.68}$$

where

$$\delta \quad := \quad \frac{u_{11}}{u_{13}^2}u_{23}^2 + u_{22} - 2\frac{u_{12}u_{23}}{u_{13}} , \qquad u_1 := \left(\frac{u_{23}}{u_{13}}\right)^2 ,$$

$$u_2 \quad := \quad -\frac{u_{23}}{u_{13}} , \qquad u_3 := \frac{u_{22}}{u_{13}} , \qquad u_4 := -\frac{u_{12}u_{23}}{u_{13}^2} , \qquad u_5 := \frac{u_{11}u_{23}}{u_{13}^2} ,$$

$$u_6 \quad := \quad -\frac{u_{12}}{u_{13}} , \qquad u_7 := \left(\frac{u_{12}}{u_{13}}\right)^2 , \qquad u_8 := -\frac{u_{11}u_{22}}{u_{13}^2} . \tag{3.69}$$

Observe that from the above asymptotics (3.66) it follows

$$\frac{u_{11}}{u_{13}^2} = \pm\frac{\beta^{\pm}}{(\alpha^{\pm})^2} \pi_2^{\pm} \left(1 + O(\varepsilon^{\bar{q}_2})\right) , \tag{3.70}$$

for some $\bar{q}_2 > 0$; we also recall that

$$\pi_2^{\pm} = \frac{1}{|E_{\pm}| \sqrt{\varepsilon + |E_{\pm}|}} , \qquad E_{\pm} := \text{Re}\left(h^{\pm}(J_1)\right) , \tag{3.71}$$

where $-2\varepsilon + \varepsilon^{\lambda} \leq E_- \leq -\varepsilon^{\lambda}$, $\varepsilon^{\lambda} \leq E_+ \leq E_0$. Notice that, from (3.65) and (3.66), it follows also that

$$\sup_{i, J \in D_r} |u_i| \leq \frac{C_3}{\varepsilon^{\lambda-1} \ln\varepsilon^{-1}} . \tag{3.72}$$

Thus, it remains to estimate $1/|\delta|$. From (3.69), (3.66) and (3.70), one sees that there exist a complex number $z := z_1 + iz_2$ with $z_1 > 0$ and $|z_2| < z_1/10$ such that[22]

$$\delta = \varepsilon\left(\pm z\pi_2^{\pm} \varepsilon \left(\text{Re } J_2\right)^2 + \sigma + O(\varepsilon^{\bar{q}_3})\right) ,$$

$$\text{Re }\delta = \varepsilon\left(\pm z_1 \pi_2^{\pm}\varepsilon \left(\text{Re } J_2\right)^2 + \sigma + O(\varepsilon^{\bar{q}_3})\right) \tag{3.73}$$

for a suitable $\bar{q}_3 > 0$. Let us consider the two different signs separately. In the "plus" case, we have to distinguish whether $\sigma = 1$ or $\sigma = -1$. When $\sigma = 1$, since $z_1\pi_2^+ \varepsilon \left(\text{Re } J_2\right)^2 > 0$, one has

$$|\delta| \geq |\text{Re }\delta| = \varepsilon \left(z_1\pi_2^+ \varepsilon \left(\text{Re } J_2\right)^2 + 1 + O(\varepsilon^{\bar{q}_3})\right) \geq \frac{\varepsilon}{2} \tag{3.74}$$

---

[21]Obviously, $x = O(\varepsilon^c)$ means that there exists a positive constant d such that, for all ε small enough, $|x| \leq d\varepsilon^c$.

[22]Use also that, for $J_2 \in \widehat{D}'_r$, $|\text{Im}J_2|/|\text{Re }J_2| \leq \text{const. } \varepsilon^{\lambda-\frac{1}{2}-b} < \sqrt{\varepsilon}$ by (3.38).

for $\varepsilon > 0$ small enough. Let now $\sigma = -1$ and notice that [23]

$$\pi_2^+(E_+) \leq \pi_2^+(\varepsilon^{\frac{2}{3}}/c) \leq \frac{c}{\varepsilon} \ , \qquad\qquad \forall\, E_+ \geq \frac{\varepsilon^{\frac{2}{3}}}{c} \ ,$$

$$\pi_2^+(E_+) \geq \pi_2^+(c\varepsilon^{\frac{2}{3}+\frac{4}{3}b}) \geq \frac{1}{2c\varepsilon^{1+2b}} \ , \qquad \forall\, E_+ \leq c\varepsilon^{\frac{2}{3}+\frac{4}{3}b} \ . \qquad (3.75)$$

Choose

$$c := \frac{1}{16} \ \min\left\{ z_1 R_1^2 \ , \ \frac{1}{z_1 R_1^2} \right\} \ . \qquad (3.76)$$

Thus, in the region $E_+ \geq \varepsilon^{2/3}/c$, one has

$$|\delta| \geq |\operatorname{Re} \delta| \geq \varepsilon\Big(1 - 4z_1 c R_1^2 + O(\varepsilon^{\bar{q}_3})\Big) \geq \frac{\varepsilon}{2} \ ;$$

in the region $E_+ \leq c\varepsilon^{\frac{2}{3}+\frac{4}{3}b}$, one has

$$|\delta| \geq |\operatorname{Re} \delta| \geq \Big(\frac{z_1 R_1^2}{8c} - 1 + O(\varepsilon^{\bar{q}_3})\Big) \geq \frac{\varepsilon}{2} \ .$$

Let us turn now to the "minus" sign case and notice that $\varepsilon^\lambda \leq |E_-| \leq 2\varepsilon$ and $\pi_2^- \geq \kappa_6/\varepsilon^{3/2}$ with a suitable $\kappa_6 > 0$. Hence (recalling (3.66) and the assumption $b < 1/4$)

$$|\delta| \geq C_4 \ \pi_2^- \ \varepsilon^{2(1+b)} - C_5\varepsilon \geq C_4\kappa_6\varepsilon^{\frac{1}{2}+2b} - C_5\varepsilon \geq C_6 \ \varepsilon^{\frac{1}{2}+2b} \ , \qquad (3.77)$$

where C_4, C_5 and C_6 are suitable positive constants. Thus, since $\frac{1}{2} + 2b < 1$, we see that (in all cases)

$$|\delta^{-1}| \leq \frac{C_7}{\varepsilon} \ , \qquad (3.78)$$

with a suitable $C_7 > 0$. This bound together with (3.72) leads to the estimates on $\|U^{-1}\|$ given in (3.60), completing the proof of the lemma. $\qquad\square$

We proceed to estimating the parameters appearing in the statement of Lemma 2.3. From (3.64), (3.66) and (2.23) there follows that

$$\left|\frac{\partial h_*^\pm}{\partial J_1}\right| := |u_{13}| \geq C_8\frac{\sqrt{\varepsilon}}{\ln\varepsilon^{-1}} \ ,$$

$$\left|\frac{\partial h_*^\pm}{\partial J_2}\right| := |u_{23}| \geq C_8\varepsilon^{1+b} \ , \qquad (3.79)$$

for a suitable $C_8 > 0$ so that $\min_{i,J}\left|\frac{\partial h_*^\pm}{\partial J_i}\right| \geq C_8\varepsilon^{1+b}$. We next choose $\gamma \ll C_8\varepsilon^{1+b}$. Since the norm of $H_*^\pm$ is exponentially small with ε, we can choose also γ exponentially small with ε: we let, in fact, for a suitable $\gamma_0 > 0$,

$$\gamma := \gamma_0 \exp\Big(-\frac{1}{\varepsilon^{q_2}}\Big) \ , \qquad \text{with}\quad q < q_2 < q_1 \ . \qquad (3.80)$$

Therefore, in view of (3.57), (3.60) and (3.80), we can take [24]

$$A := \frac{C_0}{\varepsilon^{\lambda-1}(\ln\varepsilon^{-1})^3} \ , \qquad F := \exp\Big(-\frac{1}{2\varepsilon^{q_1}}\Big) \ , \qquad G := \frac{C_9}{\varepsilon^{2\lambda-1}(\ln\varepsilon^{-1})^4} \ , \quad (3.81)$$

---

[23]By (3.71) π_2^+ is a decreasing function of E_+. Recall also (3.43), that $c < 1$ and that ε is small.

[24]Recall the definitions of F and G given in (2.33).

for a suitable $C_9 > 0$. Next, we show that C in (2.34) is one in our case. By (3.58), (3.60), (3.80) and (3.81), we see that (for a suitable $C_{10} > 0$)

$$\frac{\gamma(s - \bar{s})^{c_1}}{c_2 A \, r \, |\ln F|^{c_3}} = C_{10} \frac{\varepsilon^{c_3 q_1} \exp\left(- \frac{1}{\varepsilon^{q_2}} \right)}{(\ln \varepsilon^{-1})^{c_1 - 2}} \; ;$$

which implies that $C = 1$ for ε small enough. Therefore, recalling the definition (2.34) of $\widehat{F}$, we can take, for a suitable $C_{11} > 0$ (see (3.81) and (3.58)) and for $\varepsilon > 0$ small enough,

$$\widehat{F} \leq C_{11} \exp\left(- \frac{1}{\varepsilon^{q_2}} \right) , \qquad (3.82)$$

which obviously will be smaller than one for any $\varepsilon > 0$ small enough. Thus, under conditions (3.38), (3.59) and (3.80), Lemma 2.3 can be applied to the Hamiltonian (3.54) showing the existence of KAM tori in each energy level of $W_{r,s}$ apart from a small set of measure bounded by[25] $O(\hat{\gamma}) \leq O(\exp(-1/\varepsilon^q))$. Thus (recall Remark 5), the motions starting in $W_{r,s}$ have action variables $O(\exp(-1/q))$–close to their initial values for all times. In the original coordinates (I, φ), the measure of the complementary of the KAM tori is again bounded by $O(\exp(-1/q))$; the KAM tori fill up the region $\mathcal{M}^{(\sigma)}$ with the exception of a set of measure $O(\exp(-1/q))$. In view of (3.55), the displacement of the KAM tori from the corresponding unperturbed ones is $O(\varepsilon^a)$ while the oscillation of the graph of the tori may be bounded by $O(\sqrt{\varepsilon})$. Repeating the argument in Remark 5 we find that, denoting $(I(t), \varphi(t))$ the ϕ^t evolution of (I_0, φ_0) with[26] $(I_0, \varphi_0) \in \mathcal{M}^{(\sigma)}$,

$$|I(t) - I_0| < C_{12}\sqrt{\varepsilon} , \qquad \forall \, t , \qquad (3.83)$$

(provided $\varepsilon > 0$ is small enough).

This concludes the proof of Theorem 1.4.   □

Proof of Theorem 1.3

We proceed to show that Theorem 1.4 and energy conservation imply (1.10) in $\mathcal{M}_R$ when $\sigma = 1$ and in $\mathcal{M}_R \backslash \mathcal{N}_*$ when $\sigma = -1$ (recall the definition of $\mathcal{N}_*$ in (1.12)).

In view of the oscillations of the KAM tori in the region $\mathcal{M}^{(\sigma)}$ we shall consider slightly smaller sets $\widetilde{\mathcal{M}}^{(\sigma)} \subset \mathcal{M}^{(\sigma)}$. To define such sets we let $\widetilde{\mathcal{N}}_* := \mathcal{N}_*$ and:

$$\widehat{\mathcal{N}}_* := \left\{ (I, \varphi) : \; 2c \, \varepsilon^{\frac{2}{3} + \frac{4}{3}b} < H_{\mathrm{p}} < \frac{\varepsilon^{\frac{2}{3}}}{2c} \right\} ,$$

$$\widetilde{\mathcal{N}}^{(1)} := \left\{ (I, \varphi) : \; |H_{\mathrm{p}}| < 2\varepsilon^{1+2b} \;\; \text{or} \;\; H_{\mathrm{p}} < -2\varepsilon + 2\varepsilon^{1+2b} \right\} ,$$

$$\widetilde{\mathcal{N}}^{(2)} := \left\{ (I, \varphi) : \; |I_2| < 2R\varepsilon^b \right\} ,$$

$$\widetilde{\mathcal{M}}^{(1)} := \mathcal{M}_R \backslash (\widetilde{\mathcal{N}}^{(1)} \cup \widetilde{\mathcal{N}}^{(2)}) , \qquad \widetilde{\mathcal{M}}^{(-1)} := \widetilde{\mathcal{M}}^{(1)} \backslash \widetilde{\mathcal{N}}_* . \qquad (3.84)$$

REMARK 7. Because of Theorem 1.4 (and, hence, because of the confinement due to the presence of two–dimensional KAM tori in three–dimensional energy levels), the smaller sets $\widetilde{\mathcal{M}}^{(\sigma)}$ have the property that $\bigcup_{t \in \mathbb{R}} \phi_\sigma^t(\widetilde{\mathcal{M}}^{(\sigma)}) \subset \mathcal{M}^{(\sigma)}$ (where ϕ_σ^t denotes the $H^{(\sigma)}$–flow). In particular, in the case $\sigma = -1$, a trajectory cannot cross

---

[25] Provided τ is choosen striclty larger than one; the constant $\hat{\gamma}$ is defined in (iii) of Lemma 2.3 and, in view of Lemma 3.1, is related to γ by a power of ε.

[26] Recall that $R < R_1 < R_0$ and that ε will be small compared also to $(R_0 - R)$.

the region $\widehat{\mathcal{N}}_*$ (a fact that could also be checked directly by energy conservation since $\frac{2}{3} + \frac{4}{3}b < 1$).

Denote by $z(t) := (I(t), \varphi(t))$ the motion with initial data $z_0 := (I_0, \varphi_0)$ governed by $H^{(\sigma)}$ in $\mathcal{M}_R$ (if $\sigma = 1$) or $\mathcal{M}_R \backslash \mathcal{N}_*$ (if $\sigma = -1$). Let us consider the different cases which may occur.

(i) If $z_0 \in \widetilde{\mathcal{M}}^{(\sigma)}$ then, as remarked above, $z(t)$ doest not leave $\mathcal{M}^{(\sigma)}$ where (3.83) (and hence (1.10)) holds.

(ii) If $z(t) \in \widetilde{\mathcal{N}}^{(1)}$ for $|t| < T$ for some $T > 0$, then, by energy conservation, (1.10) holds[27] for $|t| < T$.

(iii) If $z(t) \in \widetilde{\mathcal{N}}^{(2)}$ for $|t| < T$ then (1.10) (trivially) holds for $|t| < T$.

(iv) By (ii) and (iii) (1.10) holds until $z(t) \in \widetilde{\mathcal{N}}^{(1)} \cup \widetilde{\mathcal{N}}^{(2)}$. But if $z(t)$ leaves $\widetilde{\mathcal{N}}^{(1)} \cup \widetilde{\mathcal{N}}^{(2)}$ and enters the region $\widetilde{\mathcal{M}}^{(\sigma)}$, then, by (i), (1.10) holds again. $\qquad\square$

Appendix A. The Planetary D'Alembert Hamiltonian. In this appendix we revisit briefly the Hamiltonian version of the planetary D'Alembert model as presented in [4] and discuss a connection with the result presented in this paper.

In [4], Section 12, it is shown that the motion of a planet modelled by a rotational ellipsoid with flatness $\varepsilon > 0$ whose center of mass revolves on a Keplerian ellipse of eccentricity $\mu > 0$, subject to the gravitational attraction of a fixed star occupying one of the foci of the ellipse, is governed (in suitable units) by a Hamiltonian function given by

$$
\begin{aligned}
H_{\varepsilon,\mu}(J, \psi) \;&= \frac{J_1^2}{2} + \bar{J}_1 J_1 + \omega(J_3 - J_2) \\
&\quad + \varepsilon F_0(J_1, J_2, \psi_1, \psi_2) + \varepsilon\mu F_1(J_1, J_2, \psi_1, \psi_2, \psi_3) \;,
\end{aligned}
\tag{A.1}
$$

where:

- $(J, \psi) \in A \times \mathbb{T}^3$ are standard symplectic coordinate; the domain $A \subset \mathbb{R}^3$ is given by

$$
A := \left\{ |J_1| < c\varepsilon^\ell \;, \quad |J_2 - \bar{J}_2| < d \;, \quad J_3 \in \mathbb{R} \right\} \;,
\tag{A.2}
$$

with $0 \le \ell < 1$, $0 < c\varepsilon^\ell \ll d \ll 1$, $(\bar{J}_1, \bar{J}_2)$ fixed "reference data" (verifying certain assumptions spelled out below);
- $2\pi/\omega$ is the period of the Keplerian motion ("year of the planet");
- the functions F_i are trigonometric polynomial given by

$$
F_0 = \sum_{\substack{j \in \mathbb{Z} \\ |j| \le 2}} c_j \cos(j\psi_1) + d_j \cos(j\psi_1 + 2\psi_2)
$$

$$
F_1 = \sum_{\substack{j \in \mathbb{Z} \\ |j| \le 2}} (-3)c_j \cos(j\psi_1 + \psi_3) + \frac{d_j}{2} \left\{ \cos(j\psi_1 + 2\psi_2 + \psi_3) - 7\cos(j\psi_1 + 2\psi_2 - \psi_3) \right\}
$$

---

[27] In fact, calling $E_p(t) = H_p(I_1(t), \varphi_1(t))$, if $z(t) \in \widetilde{\mathcal{N}}^{(1)}$ for $|t| < T$, then $|E_p(t) - E_p(0)| \le O(\varepsilon^\lambda)$ for all $|t| \le T$ (recall that $\lambda = 1 + 2b$ and that $a > \lambda$). Thus, by energy conservation, there follows that

$$
\frac{I_2(t)^2 - I_2(0)^2}{2} + \frac{E_p(t) - E_p(0)}{\varepsilon} = O(\varepsilon^{a-1}) \;,
$$

for all $|t| < T$. Therefore, $I_2(t)^2 - I_2(0)^2 = O(\varepsilon^{\lambda-1})$ and (1.10) follows.

where c_j and d_j are functions of (J_1, J_2) listed in the following item;
• let

$$\kappa_1 := \kappa_1(J_1) := \frac{L}{\bar{J}_1 + J_1} \ , \qquad \kappa_2 := \kappa_2(J_1, J_2) := \frac{J_2}{\bar{J}_1 + J_1} \ ,$$

$$\nu_1 := \nu_1(J_1) := \sqrt{1 - \kappa_1^2} \ , \qquad \nu_2 := \nu_2(J_1, J_2) := \sqrt{1 - \kappa_2^2} \ ;$$

where L is a real parameter; the parameters $\bar{J}_i$, L, ε and the constants c and d are assumed to satisfy

$$|\bar{J}_2| + d + c\varepsilon^\ell < |\bar{J}_1| \ , \qquad 0 < |L| + c\varepsilon^\ell < |\bar{J}_1| \ , \qquad (A.3)$$

so that $0 < \kappa_i < 1$ (and the ν_i's are well defined). Then, the functions c_j and d_j are defined by

$$c_0(J_1, J_2) := \frac{1}{4}\left(2\kappa_1^2\nu_2^2 + \nu_1^2(1 + \kappa_2^2)\right) \ , \qquad d_0(J_1, J_2) := -\frac{\nu_2^2}{4}(2\kappa_1^2 - \nu_1^2) \ ,$$

$$c_{\pm 1}(J_1, J_2) := \frac{\kappa_1\kappa_2\nu_1\nu_2}{2} \ , \qquad d_{\pm 1}(J_1, J_2) := \mp\frac{(1 \pm \kappa_2)\kappa_1\nu_1\nu_2}{2} \ ,$$

$$c_{\pm 2}(J_1, J_2) := -\frac{\nu_1^2\nu_2^2}{8} \ , \qquad d_{\pm 2}(J_1, J_2) := -\frac{\nu_1^2(1 \pm \kappa_2)^2}{8} \ . \qquad (A.4)$$

REMARK 8. (i) We recall that, actually, the above model is a "first order μ–truncation" of the full D'Alembert model, which in place of $F_0 + \mu F_1$ has a series $\sum_{j\geq 0}\mu^j F_j$ with F_j trigonometric polynomials.

(ii) Since J_3 appears only linearly with coefficient ω, the angle ψ_3 corresponds to time t and $H_{\varepsilon,\mu}$ is actually a two–degrees–of–freedom Hamiltonian depending explicitly on time in a periodic way (with period $2\pi/\omega$).

(iii) The physical interpretation of the action–variables J_1, J_2 and the parameter L is the following. The action variable $\bar{J}_1 + J_1$ is (in suitable units) the absolute value of the angular momentum of the planet; the variable J_2 is the absolute value of the projection of the angular momentum of the planet onto the direction orthogonal to the *ecliptic* plane (i.e., the plane containing the Keplerian ellipse) and L is the absolute value of the projection of the angular momentum of the planet in the direction of the polar axis of the planet (and is a constant of the motion).

(iv) Under our assumptions (i.e., that $0 < c\varepsilon^\ell \ll d \ll 1$), the *average over the angle of* $H_{\varepsilon,0}$ is given by

$$\frac{J_1^2}{2} + \bar{J}_1 J_1 + \omega(J_3 - J_2) + \varepsilon\frac{1}{4}\left\{(2 - \bar{\nu}_1^2) - (2 - 3\bar{\nu}_1^2)\frac{J_2^2}{\bar{J}_1^2} + O(d)\right\} \ , \qquad (A.5)$$

where $\bar{\nu}_1 := \sqrt{1 - (L/\bar{J}_1)^2}$. By (iii) we see that $\bar{\nu}_1 \ll 1$ corresponds to rotations of the planet with spin axis nearly orthogonal to the ecliptic plane (a case common, for example, in the Solar System). In such a case the *average over the angle of* $H_{\varepsilon,0}$ *is not a convex function of the action variable* J_2. This lack of convexity (for the "effective" Hamiltonian) is quite a common feature in Celestial Mechanics and is exhibited, for example, also in three–body–problems. This is the reason why, in our model problem, we considered also non convex cases (corresponding above to $\sigma = -1$).

We proceed now to show how the D'Alembert model relates to the model (1.5)–(1.8) investigated in this paper.

We are interested (as in [4]) to "reference data" corresponding to day/year resonances (as the one often observed in the Solar system). We let, therefore,

$$\bar{J}_1 = 2\omega \; , \qquad\qquad\qquad (A.6)$$

corresponding to a 2:1 day/year–resonance. Then, the linear symplectic change of variables $\hat{\Phi} : (\hat{J}, \hat{\psi}) \to (J, \psi)$ given by

$$J = (\hat{J}_1, 2\hat{J}_1 + \hat{J}_2, \hat{J}_2 + \hat{J}_3) \; , \qquad \psi = (\hat{\psi}_1 - 2\hat{\psi}_2 + 2\hat{\psi}_3, \hat{\psi}_2 - \hat{\psi}_3, \hat{\psi}_3) \; , \qquad (A.7)$$

casts the Hamiltonian $H_{\varepsilon,\mu}$ into the form $\widehat{H}_{\varepsilon,\mu}(\hat{J}, \hat{\psi}) := H_{\varepsilon,\mu} \circ \hat{\Phi}(\hat{J}, \hat{\psi})$ with

$$\widehat{H}_{\varepsilon,\mu}(\hat{J}, \hat{\psi}) := \frac{\hat{J}_1^2}{2} + \omega\hat{J}_3 + \varepsilon\widehat{F}_0(\hat{J}_1, \hat{J}_2, \hat{\psi}) + \varepsilon\mu\widehat{F}_1(\hat{J}_1, \hat{J}_2, \hat{\psi}) \; , \qquad (A.8)$$

where

$$\widehat{F}_0(\hat{J}_1, \hat{J}_2, \hat{\psi}) := F_0(\hat{J}_1, 2\hat{J}_1 + \hat{J}_2, \hat{\psi}_1 - 2\hat{\psi}_2 + 2\hat{\psi}_3, \hat{\psi}_2 - \hat{\psi}_3) \; ,$$
$$\widehat{F}_1(\hat{J}_1, \hat{J}_2, \hat{\psi}) := F_0(\hat{J}_1, 2\hat{J}_1 + \hat{J}_2, \hat{\psi}_1 - 2\hat{\psi}_2 + 2\hat{\psi}_3, \hat{\psi}_2 - \hat{\psi}_3, \hat{\psi}_3) \; . \quad (A.9)$$

For $0 < \varepsilon \ll 1$, the angle $\hat{\psi}_3$ (i.e., the time) is a "fast variable[28]" and we may apply averaging theory (or normal form theory). We shall apply the "resonant version" (in three degrees–of–freedom) of Lemma 2.2, which, for the sake of clarity we reformulate[29]:

LEMMA A.1. *Let $\widehat{D} \subset \mathbb{R}$ and $\widehat{D}' \subset \mathbb{R}^2$ and consider a Hamiltonian $H(\hat{J}, \hat{\psi}) := h(\hat{J}) + f(\hat{J}, \hat{\psi})$ real–analytic on $\widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}} := (\widehat{D}_{\hat{r}_1} \times \widehat{D}'_{\hat{r}_2}) \times \mathbb{T}^3_{\hat{s}}$ for some $\hat{r}_2 \geq \hat{r}_1 > 0$ and $\hat{s} > 0$. Assume that there exist $K \geq 6/\hat{s}$ and $\alpha > 0$ such that*

$$|\omega(\hat{J}) \cdot k| \geq \alpha \; , \qquad \forall \; k \in \mathbb{Z}^3 \; , \; |k| \leq K \; , \quad k_3 \neq 0 \; , \quad \forall \; \hat{J} \in \widehat{D}_{\hat{r}_1} \times \widehat{D}'_{\hat{r}_2} \; , \quad (A.10)$$

where $\omega(\hat{J}) := \nabla h(\hat{J})$. Assume also that

$$\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \leq \frac{\alpha\hat{r}_1}{2^8 K} \; . \qquad\qquad\qquad (A.11)$$

Then, there exist a real–analytic symplectic transformation

$$\Phi : (I, \varphi) \in \widehat{W}_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \to \Phi(I, \varphi) \in \widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}}$$

such that

$$H \circ \Phi(I, \varphi) = h(I) + g(I, \varphi_1, \varphi_2) + f_*(I, \varphi) \qquad\qquad (A.12)$$

with

$$\left\| g - \frac{1}{2\pi}\int_0^{2\pi} f_K(I, \varphi)d\varphi_3 \right\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \frac{2^6 K}{\alpha\hat{r}_1} \left(\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \right)^2 \leq \frac{1}{4}\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \; , \; (A.13)$$

$$\|f_*\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \; \exp(-K\hat{s}/6) \; ,$$

$$\|\Phi(I, \varphi) - (I, \varphi)\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \hat{c} \; \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \; ,$$

where $f_K(I, \varphi) := \sum_{|k| \leq K} \hat{f}_k(I) \exp(ik \cdot \varphi)$ and $\hat{c} > 0$ is a suitable constant.

---

[28]In fact, when $\varepsilon = 0$, $(d/dt)\hat{\psi}_1 = O(\varepsilon^\ell)$, $(d/dt)\hat{\psi}_2 = 0$ while $(d/dt)\hat{\psi}_3 = \omega$.
[29]The proof is given in [9] and the same comments in the Remark 6–(ii) apply word–by–word to the present situation.

Let $c = 10$, $0 < \ell < 1$ and ε small (in particular $\varepsilon^\ell \ll d$); let also $\bar{J}_2$ and L be so that (A.3) is (abundantly) verified. If we choose also

$$\hat{r}_1 := \varepsilon^\ell \ , \qquad \hat{r}_2 := \frac{d}{10} \ , \tag{A.14}$$

we see that the functions ν_i and κ_i (and hence the functions c_i, d_i, $\widehat{F}_i$) are analytic and bounded, for any $\hat{s} > 0$ and any $\widehat{R} > 0$, in the domain $(\widehat{D}_{\hat{r}_1} \times \widehat{D}'_{\hat{r}_2}) \times \mathbb{T}^3_{\hat{s}}$ where

$$\widehat{D} := [-10\varepsilon^\ell, 10\varepsilon^\ell] \ , \quad \widehat{D}' := \{\hat{J}_2 : \ |\hat{J}_2 - \bar{J}_2| \leq 2d\} \times \{|\hat{J}_3| \leq \widehat{R}\} \ .$$

We can now apply Lemma A.1 to the Hamiltonian $\widehat{H}_{\varepsilon,\mu}(\hat{J}, \hat{\psi}) = H(\hat{J}, \hat{\psi})$ with

$$h(\hat{J}) := \frac{\hat{J}_1^2}{2} + \omega \hat{J}_3 \ , \quad f(\hat{J}, \hat{\psi}) := \widehat{F}_0(\hat{J}_1, \hat{J}_2, \hat{\psi}) + \mu \widehat{F}_1(\hat{J}_1, \hat{J}_2, \hat{\psi}) \ .$$

Under the above position, for $0 \leq \mu \leq 1$, we have that

$$\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \leq \text{const } \varepsilon \ .$$

Thus, letting $\alpha = \omega/2$ and $K := \omega/(4\varepsilon^\ell)$, we see that (A.11) is satisfied for any $\ell < 1/2$. Now, observe that

$$\frac{1}{2\pi} \int_0^{2\pi} \widehat{F}_0(\hat{J}_1, \hat{J}_2, \hat{\psi}) d\hat{\psi}_3 = \hat{c}_0 + \hat{d}_1 \cos \hat{\psi}_1 \ , \tag{A.15}$$

where $\hat{c}_0 := \hat{c}_0(\hat{J}_1, \hat{J}_2)$ and $\hat{d}_1 := \hat{d}_1(\hat{J}_1, \hat{J}_2)$ are defined as

$$\hat{c}_0(\hat{J}_1, \hat{J}_2) := c_0(\hat{J}_1, 2\hat{J}_1 + \hat{J}_2) \ , \qquad \hat{d}_1(\hat{J}_1, \hat{J}_2) := d_1(\hat{J}_1, 2\hat{J}_1 + \hat{J}_2) \ . \tag{A.16}$$

Thus, by Lemma A.1 and (A.15), we find that $\widehat{H}_{\varepsilon,\mu} \circ \Phi(I, \varphi)$ has the form

$$\frac{I_1^2}{2} + \omega I_3 + \varepsilon \Big(\hat{c}_0(I_1, I_2) + \hat{d}_1(I_1, I_2) \cos \varphi_1 \Big)$$
$$+ \tilde{g}(I_1, I_2, \varphi_1, \varphi_2; \varepsilon, \mu) + f_*(I_1, I_2, \varphi; \varepsilon, \mu) \ , \tag{A.17}$$

where (if g is as in the Lemma) $\tilde{g} := g - \varepsilon[\hat{c}_0(I_1, I_2) + \hat{d}_1(I_1, I_2) \cos \varphi_1]$. The function f_* is exponentially small,

$$\|f_*\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \ \exp(-K\hat{s}/6) \leq \text{const } \varepsilon \exp\Big(-\frac{\omega \hat{s}}{24} \frac{1}{\varepsilon^\ell}\Big) \ , \tag{A.18}$$

and, in view of (A.13), the definition of $\tilde{g}$ and (A.15), the function $\tilde{g}$ satisfies the bound

$$\|\tilde{g}\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \text{const } (\varepsilon^{2(1-\ell)} + \varepsilon\mu) \ . \tag{A.19}$$

Thus, *assuming* $|\mu| \leq \varepsilon^c$ *with* $c > 1/2$ *and* $0 < \ell < 1/4$, in the above region of phase space, the D'Alembert Hamiltonian is described, *up to the exponentially small term in* (A.18), by the Hamiltonian[30]

$$H_D(I_1, I_2, \varphi_1, \varphi_2; \varepsilon, \mu) = \frac{I_1^2}{2} + \varepsilon \Big(\hat{c}_0(I_1, I_2) + \hat{d}_1(I_1, I_2) \cos \varphi_1 \Big)$$
$$+ \varepsilon^a G(I_1, I_2, \varphi_1, \varphi_2; \varepsilon, \mu) \ , \tag{A.20}$$

where

$$a := \min\{ 2(1 - \ell) \ , \ 1 + c \} > \frac{3}{2} \ , \quad G := \frac{\tilde{g}}{\varepsilon^a} \ , \quad \|G\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \text{const} \ . \tag{A.21}$$

---

[30]If we disregard f_* then ωI_3 becomes a constant, which we may drop.

REMARK 9. (i) The form of H_D has been, for us, the main motivation to discuss the dynamics of models described by (1.5)–(1.8).

(ii) The theory developed in this paper cannot be applied directly to H_D because of the following two reasons: (a) The "intermediate Hamiltonian" $\hat{c}_0(I_1, I_2) + \hat{d}_1(I_1, I_2) \cos \varphi_1$ has a more complicate dependence on the action variables than the one considered in this paper and one would need to extend Lemma 2.1. (b) The second (technical) reason is that the I_1–domain of analyticity of H_D is small with ε, while in our model we assumed H_1 analytic in a given (ε–independent) region.

(iii) While problem (a) needs further investigations, problem (b) may be easily overcome. More precisely, the proofs of §3 work also if one allows r_1 (i.e., the smaller action radius of analyticity of the perturbation H_1) to depend upon ε, say $r_1 = \text{const } \varepsilon^\ell$, provided $a > \ell + 3/2$. In fact, the crucial point where a dependence upon ε of r_1 comes in, is in the second step of the proof of Theorem 1.4 and, more precisely, in checking the condition (2.26) for applicability of the normal form lemma 2.2. If $r_1 = \text{const } \varepsilon^\ell$, for some $\ell > 0$, we see, by (3.52), that we must have

$$a > \ell + b + \lambda + q_0 + \frac{1}{2} = \ell + \frac{3}{2} + 3b + q_0 \ , \qquad \text{(A.22)}$$

which is equivalent to require (compare (1.9) and (1.13))

$$a > \ell + \frac{3}{2} \ , \qquad b < \frac{1}{3}\left(a - \frac{3}{2} - \ell\right) \ , \quad q < \left(a - \frac{3}{2} - \ell - 3b\right) \ . \qquad \text{(A.23)}$$

(iv) In the case of the D'Alembert reduced Hamiltonian H_D, we have $a = 2(1-\ell)$, so that (A.23) would be satisfied provided $0 < \ell < 1/6$. Of course, the argument in (iii) is applicable to H_D only if the *qualitative* dependence on ε of the analyticity domain in the analogous of Lemma 2.1 remains the same.

Appendix B. **Real–analytic action–angle variables.** Here, we shall give a complete proof of Lemma 2.1. For the purpose of this appendix we shall denote the variables (I_1, φ_1) with (the standard pendulum coordinate names) (p, q) and set $E(p, q) := H_{\mathrm{p}}(p, q; \varepsilon) := H_{\mathrm{p}}(I_1, \varphi_1; \varepsilon)$. We shall also denote the action–angle variables for $E(p, q)$ by (P, Q) (which therefore coincide with the variables $(\hat{I}_1, \hat{\varphi}_1)$ of Lemma 2.1).

We shall use the following notation: if A, B are two strictly positive functions we shall say $A \sim B$ if there exist positive constants c^+, c^- so that $c^- A \le B \le c^+ A$ pointwise. For example,

$$\sqrt{A + B} - \sqrt{A} = \frac{B}{\sqrt{A + B} + \sqrt{A}} \sim \frac{B}{\sqrt{A + B}}. \qquad \text{(B.1)}$$

Obviously, "$\sim$" is transitive. Also, if A, B and C are strictly positive, then $A \sim B$ implies $A + C \sim B + C$. Let $x = x_1 + ix_2$ be a complex number (with $x_i \in \mathbb{R}$). Since[31] $x_1^2 + x_2^2 \sim (|x_1| + x_2)^2$, from (B.1) it follows immediately the following trivial lemma

---

[31] $(|x_1| + x_2)^2/2 \le x_1^2 + x_2^2 \le (|x_1| + x_2)^2$.

LEMMA B.1. *If $x_1 \geq 0$, then $\sqrt{x_1 \pm ix_2} = w_1 \pm iw_2$; if $x_1 < 0$ then $\sqrt{x_1 \pm ix_2} = w_2 \pm iw_1$ where*

$$w_1 = \frac{1}{\sqrt{2}}\sqrt{x_1 + \sqrt{x_1^2 + x_2^2}} \sim (x_1^2 + x_2^2)^{1/4} \sim \sqrt{|x_1| + x_2} \qquad (B.2)$$

$$w_2 = \frac{1}{\sqrt{2}}\sqrt{-x_1 + \sqrt{x_1^2 + x_2^2}} \sim \frac{x_2}{(x_1^2 + x_2^2)^{1/4}} \sim \frac{x_2}{\sqrt{|x_1| + x_2}} \qquad (B.3)$$

If $x_1 \geq 0$ then $(x_1 \pm ix_2)^{-1/2} = y_1 \mp iy_2$; if $x_1 < 0$ then $(x_1 \pm ix_2)^{-1/2} = y_2 \mp iy_1$ where

$$y_1 = \frac{\sqrt{x_1 + \sqrt{x_1^2 + x_2^2}}}{\sqrt{2}\sqrt{x_1^2 + x_2^2}} \sim \frac{1}{\sqrt{|x_1| + x_2}} \;, y_2 = \frac{\sqrt{-x_1 + \sqrt{x_1^2 + x_2^2}}}{\sqrt{2}\sqrt{x_1^2 + x_2^2}} \sim \frac{x_2}{(|x_1| + x_2)^{3/2}}$$
(B.4)

If $x_1 \geq 0$ then $(x_1 \pm ix_2)^{-3/2} = z_1 \mp iz_2$; if $x_1 < 0$ then $(x_1 \pm ix_2)^{-3/2} = -z_2 \mp iz_1$ where $z_1 = y_1(y_1^2 - 3y_2^2)$, $z_2 = y_2(3y_1^2 - y_2^2)$ and y_1, y_2 are as above. Furthermore, if $y_1 \geq 2y_2$ then $y_1^2 - 3y_2^2 \sim y_1^2$, $3y_1^2 - y_2^2 \sim y_1^2$ and

$$z_1 \sim \frac{1}{(|x_1| + x_2)^{3/2}} \;, \qquad z_2 \sim \frac{x_2}{(|x_1| + x_2)^{5/2}} \qquad (B.5)$$

We divide the proof of Lemma 2.1 in several steps considering, in particular, separately positive and negative pendulum energy $E = E(p, q)$. In the following we shall consider energies $E = E_1 + iE_2 \in \mathbb{C}$ such that

$$|E_2| \leq \tilde{c}|E_1| \qquad (B.6)$$

for a suitable $0 < \tilde{c} < 1$.

First step: action variable (positive energy). In such a case, as well known, the action variable for $E(p, q)$ is given by

$$P^+(E) = \frac{\sqrt{2}}{\pi} \int_0^\pi \sqrt{E + \varepsilon(1 + \cos\psi)}\, d\psi \qquad (B.7)$$

and, denoting by with a dot the derivative with respect to E, we have

$$\dot{P}^+(E) = \frac{\sqrt{2}}{2\pi} \int_0^\pi \frac{1}{\sqrt{E + \varepsilon(1 + \cos\psi)}} d\psi \;, \qquad (B.8)$$

$$\ddot{P}^+(E) = -\frac{\sqrt{2}}{4\pi} \int_0^\pi \frac{1}{(E + \varepsilon(1 + \cos\psi))^{3/2}} d\psi$$

and, in general,

$$\frac{d^{n+1}}{dh^{n+1}} P^+(E) = (-1)^n \frac{\sqrt{2}}{2\pi} \frac{(2n-1)!!}{2^n} \int_0^\pi \frac{1}{(E + \varepsilon(1 + \cos\psi))^{n+1/2}} d\psi \;.$$

Notice that the above functions, viewed as functions of E_2 at E_1 fixed, have even real part and odd imaginary part. Thus in the following *we may consider only the case $E_2 \geq 0$*.

Setting $E + \varepsilon(1 + \cos\psi) := E_1 + \varepsilon(1 + \cos\psi) + iE_2 := x_1(\psi) + ix_2$ with $x_1(\psi) = E_1 + \varepsilon(1 + \cos\psi)$ and $x_2 = E_2$, we get (notice that $x_1(\psi) \geq E_1 > E_2 = x_2$) $x_1(\psi) + x_2 \sim x_1(\psi)$. Thus, $P^+(E) = P_1^+(E) + iP_2^+(E)$, $\dot{P}^+(E) = \dot{P}_1^+(E) + i\dot{P}_2^+(E)$, $\ddot{P}^+(E) = \ddot{P}_1^+(E) + i\ddot{P}_2^+(E)$.

From (B.2) it follows

$$P_1^+ \sim \int_0^\pi \sqrt{x_1(\psi)}\, d\psi \sim \int_0^{\pi/4} \sqrt{x_1(\psi)}\, d\psi \sim \int_0^{\pi/4} \sqrt{E_1 + \varepsilon}\, d\psi \sim \sqrt{E_1 + \varepsilon}\,.$$

Since, for $\psi \in (0, \pi)$, it is $x_1(\psi) = E_1 + \varepsilon(1 + \cos\psi) \sim E_1 + \varepsilon(\pi - \psi)$ (making in the integrals the change of variable $y = (\pi - \psi)\sqrt{\varepsilon/E_1}$), from (B.4) we get[32]

$$
\begin{aligned}
\dot{P}_1^+(E) &\sim \int_0^\pi \frac{d\psi}{\sqrt{x_1(\psi)}} \sim \int_0^\pi \frac{d\psi}{\sqrt{E_1 + \varepsilon(\pi - \psi)^2}} = \frac{1}{\sqrt{\varepsilon}} \int_0^{\pi\sqrt{\frac{\varepsilon}{E_1}}} \frac{dy}{\sqrt{1 + y^2}} \\
&= \frac{1}{\sqrt{\varepsilon}} \operatorname{arcsinh}\left(\pi\sqrt{\frac{\varepsilon}{E_1}}\right) \sim \frac{1}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{E_1}}\right)
\end{aligned}
$$

Similarly,

$$P_2^+(E) \sim \int_0^\pi \frac{d\psi}{\sqrt{x_1(\psi)}} \sim \frac{E_2}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{E_1}}\right).$$

From (B.4) it follows that

$$
\begin{aligned}
-\ddot{P}_1^+(E) &\sim \int_0^\pi \frac{d\psi}{(x_1(\psi))^{3/2}} \sim \int_0^\pi \frac{d\psi}{(E_1 + \varepsilon(\pi - \psi)^2)^{3/2}} \\
&= \frac{1}{E_1\sqrt{\varepsilon}} \int_0^{\pi\sqrt{\frac{\varepsilon}{E_1}}} \frac{dy}{(1 + y^2)^{3/2}} = \frac{\pi}{E_1^{3/2}\sqrt{1 + \pi^2\varepsilon/E_1}} \sim \frac{1}{E_1\sqrt{E_1 + \varepsilon}}
\end{aligned}
$$

and

$$-\dot{P}_2^+(E) \sim E_2 \int_0^\pi \frac{d\psi}{(x_1(\psi))^{3/2}} \sim \frac{E_2}{E_1\sqrt{E_1 + \varepsilon}}.$$

Thus, using (B.4), we find[33]

$$
\begin{aligned}
-\ddot{P}_2^+(E) &\sim E_2 \int_0^\pi \frac{d\psi}{(x_1(\psi))^{5/2}} \sim E_2 \int_0^\pi \frac{d\psi}{(E_1 + \varepsilon(\pi - \psi)^2)^{5/2}} \\
&= \frac{E_2}{E_1\sqrt{\varepsilon}} \int_0^{\pi\sqrt{\frac{\varepsilon}{E_1}}} \frac{dy}{(1 + y^2)^{5/2}} \sim \frac{E_2}{E_1^2\sqrt{E_1 + \varepsilon}}
\end{aligned}
$$

Summarizing the following estimates hold

$$P_1^+(E) \sim \sqrt{E_1 + \varepsilon}\,, \qquad P_2^+(E) \sim \frac{E_2}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{E_1}}\right), \qquad \text{(B.9)}$$

$$\dot{P}_1^+(E) \sim \frac{1}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{E_1}}\right), \qquad -\dot{P}_2^+(E) \sim \frac{E_2}{E_1\sqrt{E_1 + \varepsilon}}, \quad \text{(B.10)}$$

$$-\ddot{P}_1^+(E) \sim \frac{1}{E_1\sqrt{E_1 + \varepsilon}}, \qquad -\ddot{P}_2^+(E) \sim \frac{E_2}{E_1^2\sqrt{E_1 + \varepsilon}}\,. \qquad \text{(B.11)}$$

Second step: action variable (negative energy). In this case the action variable is given by

$$P^-(E) = \frac{2\sqrt{2}}{\pi} \int_0^{\psi_0(E)} \sqrt{E + \varepsilon(1 + \cos\psi)}\, d\psi$$

---

[32]Use $\operatorname{arcsinh}(t) = \ln(t + \sqrt{1 + t^2}) \sim \ln(1 + t)$.

[33]In the last estimate we considered separately $E_1 \leq \varepsilon$ (in which case $\sqrt{E_1 + \varepsilon} \sim \sqrt{\varepsilon}$ and $\int_0^{\pi\sqrt{\frac{\varepsilon}{E_1}}} \frac{dy}{(1 + y^2)^{5/2}} \sim 1$) and $E_1 > \varepsilon$ (in which case $\sqrt{E_1 + \varepsilon} \sim \sqrt{E_1}$ and $(1 + y^2)^{-5/2} \sim 1$.

where $\psi_0(E)$ is the first positive number such that $E + \varepsilon(1 + \cos\psi_0(E)) = 0$. Differentiating

$$\dot{P}^-(E) = \frac{\sqrt{2}}{\pi}\int_0^{\psi_0(E)}\frac{1}{\sqrt{E + \varepsilon(1 + \cos\psi)}}\,d\psi\ .$$

Making the change of variable $\psi = \arccos(1 - \tilde{E}/\varepsilon + \xi\tilde{E}/\varepsilon)$ where $\tilde{E} := E + 2\varepsilon$ we get

$$P^-(E) = \frac{2\sqrt{2}}{\pi}\int_0^1\frac{\tilde{E}\sqrt{\xi}}{\sqrt{1-\xi}\sqrt{\tilde{E}\xi - E}}\,d\xi\ ,$$

$$\dot{P}^-(E) = \frac{\sqrt{2}}{\pi}\int_0^1\frac{1}{\sqrt{\xi}\sqrt{1-\xi}\sqrt{\tilde{E}\xi - E}}\,d\xi\ . \tag{B.12}$$

Thus,

$$\ddot{P}^-(E) = \frac{\sqrt{2}}{2\pi}\int_0^1\frac{\sqrt{1-\xi}}{\sqrt{\xi}(\tilde{E}\xi - E)^{3/2}}\,d\xi \tag{B.13}$$

and, in general,

$$\frac{d^{n+1}}{dh^{n+1}}P^-(E) = \frac{\sqrt{2}}{\pi}\frac{(2n-1)!!}{2^n}\int_0^1\frac{(1-\xi)^n}{\sqrt{\xi}\sqrt{1-\xi}(\tilde{E}\xi - E)^{n+1/2}}\,d\xi.$$

As above, for symmetry reasons, we may consider only $E_2 \geq 0$. Observe that

$$\tilde{E}\xi - E = (-E_1(1-\xi)+2\varepsilon\xi) - iE_2(1-\xi) = (\tilde{E}_1\xi - E_1) - iE_2(1-\xi) = x_1(\xi) - ix_2(\xi)$$

with $x_1(\xi) := \tilde{E}_1\xi - E_1$ and $x_2(\xi) := E_2(1-\xi)$. When $0 \leq \xi \leq 1$ it is[34] $x_1(\xi) \geq x_2(\xi) \geq 0$. Let now $y_1(\xi), y_2(\xi)$ be as in (B.4). Then

$$y_1 \sim \frac{1}{\sqrt{x_1}}\ ,\qquad y_2 \sim \frac{x_2}{x_1^{3/2}} \tag{B.14}$$

so that $y_1 \geq y_2$ provided $\tilde{c}$ is small enough. If $z_1(\xi), z_2(\xi)$ are as in (B.5), then

$$z_1 \sim \frac{1}{x_1^{3/2}}\ ,\qquad z_2 \sim \frac{x_2}{x_1^{5/2}}. \tag{B.15}$$

Obviously:

$$\frac{1}{\sqrt{\tilde{E}\xi - E}} = y_1 + iy_2\ ,\qquad \frac{1}{(\tilde{E}\xi - E)^{3/2}} = z_1 + iz_2\ . \tag{B.16}$$

By (B.12) and (B.16) we see that[35]

$$P_1^-(E) = \frac{2\sqrt{2}}{\pi}\int_0^1(\tilde{E}_1 y_1 - E_2 y_2)\frac{\sqrt{\xi}}{\sqrt{1-\xi}}\,d\xi \sim \int_0^1\frac{\tilde{E}_1\sqrt{\xi}}{\sqrt{1-\xi}\sqrt{\tilde{E}_1\xi - E_1}}\,d\xi\ . \tag{B.17}$$

Similarly, from (B.12), (B.16) and (B.14) we get:

$$\begin{aligned}P_2^-(E) &= \frac{2\sqrt{2}}{\pi}\int_0^1(\tilde{E}_1 y_2 + E_2 y_1)\frac{\sqrt{\xi}}{\sqrt{1-\xi}}\,d\xi \\ &\sim \int_0^1\frac{\tilde{E}_1 E_2\sqrt{\xi}\sqrt{1-\xi}}{(\tilde{E}_1\xi - E_1)^{3/2}}\,d\xi + \int_0^1\frac{E_2\sqrt{\xi}}{\sqrt{1-\xi}\sqrt{\tilde{E}_1\xi - E_1}}\,d\xi\end{aligned} \tag{B.18}$$

---

[34]The function $x_1(\xi)-x_2(\xi)$ is increasing so that $x_1(\xi)-x_2(\xi) \geq x_1(0)-x_2(0) = -E_1-E_2 > 0$.
[35]Use $\tilde{E}_1 y_1 - E_2 y_2 \sim \tilde{E}_1 y_1$.

Finally, from (B.12), (B.13), (B.14), (B.15), (B.16), there follows

$$\dot{P}_1^-(E) \sim \int_0^1 \frac{1}{\sqrt{\xi}\sqrt{1-\xi}\sqrt{\tilde{E}_1\xi - E_1}}\, d\xi \ , \quad \dot{P}_2^-(E) \sim E_2 \int_0^1 \frac{\sqrt{1-\xi}}{\sqrt{\xi}(\tilde{E}_1\xi - E_1)^{3/2}}\, d\xi \ ,$$

$$\ddot{P}_2^-(E) \sim \int_0^1 \frac{\sqrt{1-\xi}}{\sqrt{\xi}(\tilde{E}_1\xi - E_1)^{3/2}}\, d\xi \ , \quad \ddot{P}_2^-(E) \sim \int_0^1 \frac{(1-\xi)^{3/2}}{\sqrt{\xi}(\tilde{E}_1\xi - E_1)^{5/2}}\, d\xi. \quad \text{(B.19)}$$

If $-2\varepsilon < E_1 < -\varepsilon$ (since, in such case, $\tilde{E}_1\xi - E_1 \sim \varepsilon$, $-E_1 \sim \varepsilon$) we have

$$P_1^-(E) \sim \frac{\tilde{E}_1}{\sqrt{\varepsilon}}, \qquad\qquad P_2^-(E) \sim \frac{E_2}{\sqrt{\varepsilon}},$$

$$\dot{P}_1^-(E) \sim \frac{1}{\sqrt{\varepsilon}}, \qquad\qquad \dot{P}_2^-(E) \sim \frac{E_2}{\varepsilon^{3/2}},$$

$$\ddot{P}_1^-(E) \sim \frac{1}{\varepsilon^{3/2}}, \qquad\qquad \ddot{P}_2^-(E) \sim \frac{E_2}{\varepsilon^{5/2}} \ . \qquad\qquad \text{(B.20)}$$

The case $-\varepsilon < E_1 < 0$ (i.e. $\tilde{E}_1 \sim \varepsilon$) is a bit more complicated and it is convenient to break up the integrals in (B.17) as $\int_0^1 = \int_0^{1/2} + \int_{1/2}^1$. The latter integrals are easier to handle since if $1/2 \le \xi \le 1$ then $\sqrt{\xi} \sim 1$ and $\tilde{E}_1\xi - E_1 \sim \varepsilon$ and therefore the estimates in (B.20) follow. As for the other integrals, since $0 \le \xi \le 1/2$, one has $1 - \xi \sim 1$. Substituting $t = \frac{\tilde{E}_1}{-E_1}\xi$ (so that $\tilde{E}_1\xi - E_1 = -E_1(t+1)$) and denoting $a = -\tilde{E}_1/2E_1$, in view of the estimates in (B.20) and of the estimates done in the integrals over $(1/2, 1)$, we obtain

$$P_1^-(E) \sim \frac{\tilde{E}_1}{\sqrt{\varepsilon}} + \frac{-E_1}{\sqrt{\tilde{E}_1}} \int_0^a \frac{\sqrt{t}}{\sqrt{t+1}}\, dt \sim \frac{\tilde{E}_1}{\sqrt{\varepsilon}} \ ,$$

$$P_2^-(E) \sim \frac{E_2}{\sqrt{\varepsilon}}\left(1 + \ln \frac{\varepsilon}{-E_1}\right) \ ,$$

$$\dot{P}_1^-(E) \sim \frac{1}{\sqrt{\varepsilon}} + \frac{1}{\sqrt{\varepsilon}} \int_0^a \frac{1}{\sqrt{t}\sqrt{t+1}}\, dt \sim \frac{1}{\sqrt{\varepsilon}}\left(1 + \ln \frac{\varepsilon}{-E_1}\right) \ ,$$

$$\dot{P}_2^-(E) \sim \frac{E_2}{\varepsilon^{3/2}} + \frac{E_2}{-E_1\sqrt{\varepsilon}} \int_0^a \frac{1}{\sqrt{t}(t+1)^{3/2}}\, dt \sim \frac{E_2}{-E_1\sqrt{\varepsilon}} \ ,$$

$$\ddot{P}_1^-(E) \sim \frac{1}{\varepsilon^{3/2}} + \frac{1}{-E_1\sqrt{\varepsilon}} \int_0^a \frac{1}{\sqrt{t}(t+1)^{3/2}}\, dt \sim \frac{1}{\varepsilon^{3/2}} + \frac{1}{-E_1\sqrt{\varepsilon}} \sim \frac{1}{-E_1\sqrt{\varepsilon}} \ ,$$

$$\ddot{P}_2^-(E) \sim \frac{E_2}{\varepsilon^{5/2}} + \frac{E_2}{(-E_1)^2\sqrt{\varepsilon}} \int_0^a \frac{1}{\sqrt{t}(t+1)^{5/2}}\, dt \sim \frac{E_2}{(-E_1)^2\sqrt{\varepsilon}} \ .$$

Summarizing we find[36]

$$P_1^-(E) \quad \sim \quad \frac{\tilde{E}_1}{\sqrt{\varepsilon}} \ , \qquad P_2^-(E) \sim \frac{E_2}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{|E_1|}}\right) \ ,$$

$$\dot{P}_1^-(E) \quad \sim \quad \frac{1}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{|E_1|}}\right) \ , \qquad \dot{P}_2^-(E) \sim \frac{E_2}{|E_1|\sqrt{\varepsilon}} \ ,$$

$$\ddot{P}_1^-(E) \quad \sim \quad \frac{1}{|E_1|\sqrt{\varepsilon}} \ , \qquad \ddot{P}_2^-(E) \sim \frac{E_2}{E_1^2\sqrt{\varepsilon}} \ . \qquad\qquad \text{(B.21)}$$

---

[36]Note that if $x \in [1/2, \infty)$ then $1 + \ln x \sim \ln(1 + \sqrt{x})$.

Third step: Estimates on the action–analyticity radius. Let us define the following energy sets:

$$
\mathcal{E}^+ \;\; := \;\; \Big\{\, E = E_1 + iE_2 \in \mathbb{C},\ \text{s.t.}\ \frac{\eta}{2} < E_1 < E_0,\ |E_2| < E_2^*(E_1) \,\Big\}
$$

$$
\mathcal{E}^- \;\; := \;\; \Big\{\, E = E_1 + iE_2 \in \mathbb{C},\ \text{s.t.}\ -2\varepsilon + \frac{\eta}{2} < E_1 < -\frac{\eta}{2},\ |E_2| < E_2^*(E_1) \,\Big\}
$$

where $E_0 := E(\rho, 0)$ and $E_2^*(E_1) := \tilde{c}\eta \ln^{-1}(1 + \sqrt{\varepsilon/|E_1|})$ with $0 < \tilde{c} < 1$ a suitable small constant to be fixed later. From these definitions and from the assumption $\eta < \varepsilon/32$ it follows (B.6) for any $E \in \mathcal{E}^\pm$. Let us now define

$$
D^+ := (P^+(\eta), P^+(E_0 - \varepsilon)) , \qquad D^- := (P^-(-2\varepsilon + \eta), P^-(-\eta)) .
$$

We claim that $P^\pm(\mathcal{E}^\pm) \supset D^\pm(\sigma)_r$. By symmetry reasons we keep considering only $E_2 \geq 0$. Also we consider only the positive energy case since the other case is completely analogous. By (B.9), there exists $d_1 > 0$ such that:

$$
P_2^+(E_1 + iE_2^*(E_1)) \geq d_1 \frac{E_2^*(E_1)}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{|E_1|}}\right) = d_1 \tilde{c} \frac{\eta}{\sqrt{\varepsilon}} \geq r
$$

provided $c \leq d_1\tilde{c}$. Notice that, by definition, $P_1^+(E_1 + iE_2)$ is an increasing function of E_2. Thus, for any $|E_2| \leq E_2^*(E_0)$, we have $P_1^+(E_0 + iE_2) \geq P_1^+(E_0)$. From (B.10) it follows that there exists $d_2 > 0$ such that

$$
P_1^+(E_0) - P_1^+(E_0 - \varepsilon) > d_2\sqrt{\varepsilon} \ln\left(1 + \sqrt{\frac{\varepsilon}{|E_0|}}\right) \geq r
$$

provided c is small enough. Similarly, for any $|E_2| \leq E_2^*(\eta/2)$ we have that[37]

$$
P_1^+(\eta/2 + iE_2) \leq P_1^+(\eta/2 + iE_2^*(\eta/2)) \leq P_1^+(\eta/2 + E_2^*(\eta/2)) \leq P_1^+(3\eta/4) .
$$

We have proved that

$$
P_1^+(\eta) - P_1^+(3\eta/4) > \frac{d_2\eta}{4\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{|E_1|}}\right) \geq r
$$

for c small enough.

Fourth step: angle variable (positive energy); estimate on analyticity radius.

Let $g(E, \psi) := E + \varepsilon(1 + \cos\psi)$ and consider the set $\mathcal{M}_p^+(E) := \{q \in \mathbb{C} \text{ s.t. } g(E, q) \notin (\infty, 0]\}$. For any $E \in \mathcal{E}^+$, the functions

$$
S(E, q) := \sqrt{2} \int_0^q \sqrt{g(E, \psi)}\, d\psi , \quad \frac{\partial S}{\partial E}(E, q) , \quad \chi^+(E, q) := \Big(\frac{\partial P^+}{\partial E}(E)\Big)^{-1} \frac{\partial S}{\partial E}(E, q) \tag{B.22}
$$

are analytic on $\mathcal{M}_p^+(E)$. Observe that $q \in \mathbb{C}/2\pi\mathbb{Z} \mapsto \chi^+(E, q)$ is, for any fixed E, one–to–one. Also $Q^+(p, q) := \chi^+(E(p, q), q)$ is analytic and 2π-periodic. Let

$$
\Phi^+(p, q) := (P^+(E(p, q)), Q^+(p, q)) .
$$

Then Φ^+ is a symplectic map from $\mathcal{M}_p^+ := \{(p, q) \text{ s.t. } p \in \mathbb{C}^+ \cup \{0\},\ E(p, q) \in \mathcal{E}^+\}$ into $\mathbb{C} \times \mathbb{C}/2\pi\mathbb{Z}$. Since Φ^+ is one–to–one, its inverse $\phi^+ := (\Phi^+)^{-1}$ is well defined and analytic. In fact, $P^+(E)$ is bijective and for any (P^*, Q^*) there is a unique E^* such that $P^+(E^*) = P^*$ and, since χ^+ is one–to–one, there is a unique $q^* \in \mathbb{C}/2\pi\mathbb{Z}$ such that $\chi^+(E^*, q^*) = Q^*$. Finally, there is a unique $p^* \in \mathbb{C}^+ \cup \{0\}$ such that $E^* = E(p^*, q^*)$. Fix E. Observe that $\chi^+(E, q)$ is 2π-periodic and that

---

[37]Use $\sqrt{x_1 + \sqrt{x_1^2 + x_2^2}} \leq \sqrt{2(x_1 + x_2)}$ if $x_1, x_2 \leq 0$, and that $E_2^*(\eta/2) \geq \eta/4$ if $\tilde{c}$ is small.

$\chi^+(E,0) = 0, \chi^+(E,\pm\pi) = \pm\pi$. Let us first consider the case $E = E_1 \in \mathbb{R}$. In such a case[38] $\chi^+(E_1,(-\pi,\pi)) = (-\pi,\pi)$, $\chi^+(E_1,(0,\pm i\infty)) = (0,\pm is^+(E_1))$, $\chi^+(E_1,(\pm\pi,\pm\pi\pm i\psi_0(E_1))) = (\pm\pi,\pm\pi\pm is^+(E_1))$, where

$$ s^+(E_1) := \sqrt{2}(\frac{dP^+}{dE})^{-1} \int_0^\infty \frac{d\psi}{\sqrt{E_1 + \varepsilon(1+\cosh\psi)}} \ . $$

In fact, it is $\chi^+(E_1, \mathcal{M}_p^+(E_1)) = \mathbb{T}_{s^+(E_1)}$. Let, now, $E = E_1 + iE_2 \in \mathcal{E}^+$. In this case,

$$ \chi^+(E, \mathbb{T}_\sigma \cap \mathcal{M}_p^+(E)) \supseteq \mathbb{T}_{s^\langle E,\sigma\rangle} $$

where

$$ s^+(E,\sigma) := \min_{t\in(-\pi,\pi)} \mathrm{Im}\chi^+(E, t+i\sigma) \ , $$

and, as in the case $E = E_1$,

$$ s^+(E_1,\sigma) = \mathrm{Im}\chi^+(E_1, i\sigma) := \sqrt{2}(\frac{dP^+}{dE})^{-1} \int_0^\sigma \frac{d\psi}{\sqrt{E_1 + \varepsilon(1+\cosh\psi)}}. $$

It is easy to see that

$$ \int_0^\sigma \frac{d\psi}{\sqrt{E_1 + \varepsilon(1+\cosh\psi)}} \sim \int_0^\sigma \frac{d\psi}{\sqrt{\tilde{E}_1 + \varepsilon\psi^2}} \sim \frac{1}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{\tilde{E}_1}}\sigma\right) \ . $$

Thus, by (B.10), we get

$$ s^+(E_1,\sigma) \sim \frac{\ln\left(1 + \sqrt{\frac{\varepsilon}{\tilde{E}_1}}\sigma\right)}{\ln\left(1 + \sqrt{\frac{\varepsilon}{E_1}}\right)} \ , $$

which implies that, for any $\eta/2 < E_1 < E_0$ and for a suitable constant $c > 0$, it is

$$ s^+(E_1,\sigma) \geq 2c\frac{\sigma}{\ln(\varepsilon/\eta)}. $$

In the general case, using the estimates on P^+ and its derivatives, one has that if $E = E_1 + iE_2 \in \mathcal{E}^+$ then $s^+(E,\sigma) \geq \frac{1}{2}s^+(E_1,\sigma) \geq c\frac{\sigma}{\ln(\varepsilon/\eta)} =: s$. This proves the lemma for positive energies.

Fifth step: angle variable (negative energy). Consider the complex sets

$$ \mathcal{M} := \{q \in \mathbb{C}\, \text{s.t.}\, |\,\mathrm{Re}\, q| < \pi\} \ , \qquad \mathcal{M}_p^-(E) := \{q \in \mathcal{M}\, \text{s.t.}\, g(E,q) \notin (\infty,0]\}. $$

For any $E \in \mathcal{E}^-$ is defined, as in (B.22), the analytic function $S(E,q)$. We want to extend S on the whole set $\mathcal{M}$. At this purpose, fix $E \in \mathcal{E}^-$ and $q \in \mathcal{M}\backslash\mathcal{M}_p^-(E)$. We shall then think the integral in (B.22) evaluated on a curve $\gamma : [0,1] \to \mathcal{M}$ satisfying the following conditions: if $t \in [0,1)$, then $\gamma(t) \in \mathcal{M}_p^-(E)$, $\gamma(0) = 0$, $\gamma(1) = q$; there exists t^* such that, for $t \in [t^*, 1)$, one has $\mathrm{Im} g(E, \gamma(t)) < 0$. Note that in this way γ, which exist always, is a continuous on $[0,1]$. We shall do the the same for the integral representing $\frac{\partial S}{\partial E}(E,q)$. Let $\chi^-(E,q) := (\frac{\partial P^-}{\partial E}(E))^{-1}\frac{\partial S}{\partial E}(E,q)$ and observe that $q \mapsto \chi^-(E,q)$ is one–to–one for any fixed E. Define, also,

$$ Q^-(p,q) := \begin{cases} \chi^-(E(p,q),q) & \text{if } p \in \mathbb{C}^+ \\ \\ \pi - \chi^-(E(p,q),q) & \text{if } p \in \mathbb{C}^- \end{cases} $$

---

[38] If $a, b \in \mathbb{C}$ we denote $(a,b) := \{z := a + t(b-a), \text{with}\, t \in (0,1)\}$ and, in particular, $(a, a+i\infty) := \{a + it, \text{with}\, t \in (0,\infty)\}$.

while, for $p = 0$, (in which case $q = \pm\psi_0(E)$), we define $Q^-(0, \pm\psi_0(E)) := \pm\pi/2$. Finally, define[39]

$$W_1(E) := \{\chi^-(E, q), \text{ for } q \in \mathcal{M}, q \neq \pm\psi_0(E)\}, \qquad W_2(E) := \pi - W_2(E).$$

Denoting, $\mathcal{M}_p^- := \{(p, q) \in \mathbb{C} \times \mathcal{M} \text{ s.t. } E(p, q) \in \mathcal{E}^-\}$, we see that the mapping $\Phi^-(p, q) := (P^-(E(p, q)), Q^-(p, q))$ defines a real–analytic symplectic map $\Phi^- : \mathcal{M}_p^- \to \mathbb{C} \times \mathbb{C}$. The map Φ^- is one–to–one[40] and has an inverse $\phi^- := (\Phi^-)^{-1}$. If $Q^* \neq \pm\pi/2$, then $Q^* \in W_1(E^*)$ (resp. $Q^* \in W_2(E^*)$) and hence $p^* \in \mathbb{C}^+$ (resp. $p^* \in \mathbb{C}^-$). Moreover, since $q \mapsto \chi^-(E^*, q)$ is one–to–one, there exists a unique q^* such that $\chi^-(E^*, q^*) = Q^*$ (resp. $\pi - \chi^-(E^*, q^*) = Q^*$). Also p^* is uniquely determined since it verifies $E^* = (p^*)^2 - \varepsilon(1 + \cos q^*)$. If, instead, $Q^* = \pm\pi/2$ then $(p^*, q^*) = (0, \pm\psi_0(E^*))$. Let us show that ϕ^{-1} is analytic. Consider first P_0, (i.e. E_0) fixed and let us prove analyticity in Q. In[41] $\overset{\circ}{W}_1(E_0) = \{\chi^-(E_0, q), \text{ with } q \in \mathcal{M}(E_0)\}$ and $\overset{\circ}{W}_2(E_0) = \{\pi - \chi^-(E_0, q), \text{ with } q \in \mathcal{M}(E_0)\}$ everything is analytic, Let us check what happens on[42] $\partial W_1(E_0) \cap \partial W_2(E_0)$. Let us suppose (to fix ideas) that $Q_0 \in W_1(E_0)$. Let $Q_n \to Q_0$. Then we have to check that $(p_n, q_n) := \phi^{-1}(P_0, Q_n) \to \phi^{-1}(P_0, Q_0) =: (p_0, q_0)$. If $Q_n \in W_1(E_0)$, then, continuity in q_0 comes from the definition of χ^-. Also, since $p_n \in \mathbb{C}^+$ and $\mathrm{Im} g(E_0, q_n) < 0$, we have $p_n := \sqrt{2}\sqrt{g(E_0, q_n)} \to \sqrt{2}\sqrt{g(E_0, q_0)} =: p_0$. Let us turn now to the more delicate case $Q_n \in W_2(E_0)$. We shall consider the real case, $E_0 \in \mathbb{R}$, since the complex case is analogous but more clumsy because of the loss of symmetries. In this case $\mathrm{Re}\, Q_0 = \pm\pi/2$ and $\mathrm{Im} q_0 = 0$. The points q_n are such that $Q_n = \pi - \chi^-(E_0, q_n)$ and, by the symmetry of χ^-, we have that[43] $\chi^-(E_0, \bar{q}_n) = \pi - \bar{Q}_n \in W_1(E_0)$. Since $\bar{Q}_n \to \bar{Q}_0 = \pi - Q_0$, one has that $\chi^-(E_0, \bar{q}_n) \to Q_0$. If $\phi^{-1}(E_0, \pi - \bar{Q}_0) = \phi^{-1}(E_0, Q_0)$, since $\pi - \bar{Q}_n \in W_1(E_0)$ we see that $\bar{q}_n \to q_0$ and hence $q_n \to q_0$ (as $\mathrm{Im} q_0 = 0$). Observe, now, that $g(E_0, q_n) = g(E_0, \bar{q}_n)$. Then, $p_0 := \sqrt{2}\sqrt{g(E_0, q_0)}$, while to determine the p_n we have to choose the other branch of the square root, namely[44]

$$\begin{aligned} p_n &:= -\sqrt{2}\sqrt{g(E_0, q_n)} = -\sqrt{2}\sqrt{\overline{g(E_0, \bar{q}_n)}} \\ &= -\sqrt{2}\overline{\sqrt{g(E_0, \bar{q}_n)}} \to -\sqrt{2}\overline{\sqrt{g(E_0, q_0)}} = -\bar{p}_0 = p_0. \end{aligned}$$

The continuity for fixed Q_0 and $P_n \to P_0$ is checked similarly.

Sixth step: angle variable (negative energy); estimate on analyticity radius.

Fix E and observe that $\chi^-(E, 0) = 0$, $\chi^-(E, \pm\psi_0(E)) = \pm\pi/2$. Consider the case $E = E_1 \in \mathbb{R}$. As in the positive energy case we find $\chi^-(E_1, (-\psi_0(E_1), \psi_0(E_1))) = (-\pi/2, \pi/2)$ and $\chi^-(E_1, (0, \pm i\infty)) = (0, \pm i s^-(E_1))$, where

$$s^-(E_1) := \sqrt{2}\left(\frac{dP^-}{dE}\right)^{-1} \int_0^\infty \frac{d\psi}{\sqrt{E_1 + \varepsilon(1 + \cosh\psi)}}.$$

---

[39] Note that $W_1(E) \cap W_2(E) = \emptyset$.

[40] $P^-(E)$ is one–to–one and to any fixed $(P^*, Q^*) = \Phi(p^*, q^+)$ there corresponds a unique E^* such that $P^-(E^*) = P^*$.

[41] $\overset{\circ}{W}$ denotes, as usual, the interior of W; ∂W denotes the boundary points of W.

[42] It is enough to check that, if $Q_0 \in \partial W_1(E_0) \cap \partial W_2(E_0)$, then ϕ^{-1} is continuous in (P_0, Q_0).

[43] $\bar{q}$ denotes, as usual, the complex conjugated of $q \in \mathbb{C}$.

[44] Recall that $\mathrm{Re}\, p_0 = 0$, since $g(E_0, q_0) \in (-\infty, 0)$.

In fact, $\chi^-(E_1, \mathcal{M}_p^-(E_1)) = \{\,|\operatorname{Re} q| < \pi/2\,\} = \frac{1}{2}\mathcal{M}$. Let now, in general, $E = E_1 + iE_2 \in \mathcal{E}^-$. Then,

$$\chi^-(E, \mathbb{T}_\sigma \cap \mathcal{M}_p^-(E)) \supseteq \mathbb{T}_{s^-(E,\sigma)} \cap \mathcal{M}_p^-(E) \ ,$$

where

$$s^-(E,\sigma) := \min_{t \in (-\pi,\pi)} \operatorname{Im}\chi^-(E, t + i\sigma).$$

As in the case $E = E_1$ we have, as in the positive energy case,

$$
\begin{aligned}
s^-(E_1, \sigma) &= \operatorname{Im}\chi^-(E_1, i\sigma) := \sqrt{2}(\frac{dP^-}{dE})^{-1} \int_0^\sigma \frac{d\psi}{\sqrt{E_1 + \varepsilon(1 + \cosh\psi)}} \\
&\sim \sqrt{2}(\frac{dP^-}{dE})^{-1}\frac{1}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{\tilde{\tilde{E}}_1}}\sigma\right) \ .
\end{aligned}
$$

Also, using (B.21), we find

$$s^-(E_1, \sigma) \sim \frac{\ln\left(1 + \sqrt{\frac{\varepsilon}{\tilde{\tilde{E}}_1}}\sigma\right)}{\ln\left(1 + \sqrt{\frac{\varepsilon}{|E_1|}}\right)} \ ,$$

which implies that there exists a constant c such that, for any E_1 with $2\varepsilon + \eta/2 < E_1 < -\eta/2$, one has

$$s^-(E_1, \sigma) \geq 2c\frac{\sigma}{\ln(\varepsilon/\eta)} \ .$$

In the general case, by the estimates on P^- and its derivatives, one finds that, if $E = E_1 + iE_2 \in \mathcal{E}^-$, then

$$s^-(E, \sigma) \geq \frac{1}{2}s^-(E_1, \sigma) \geq c\frac{\sigma}{\ln(\varepsilon/\eta)} =: s \ .$$

The lemma is proved also in the negative energy case. $\qquad\square$

REFERENCES

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

[2] Arnold V. I.: *Small divisor problems in classical and Celestial Mechanics*, Russian Math. Survey **18**, No. 6, 85–192 (1963)

[3] Benettin G., Fassò F., Guzzo M.: *Nekhoroshev-stability of L4 and L5 in the spatial restricted three-body problem*, Regul. Chaotic Dyn. **3**, no. 3, 56–72 (1998)

[4] 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)

[5] Delshams A., Gutierrez P.: *Effective stability and KAM theory* J. Differential Equations 128, no. 2, 415–490 (1996)

[6] Nejshtadt A.I.: *Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions* Prikl. Mat. Mekh. **45**, No. 6, 1016–1025 (1981) (Russian); English trans.: J. Appl. Math. Mech. **45**, No. 6, 766–772 (1982)

[7] Nekhoroshev N.N.: *An exponential estimate of the time of stability of nearly–integrable Hamiltonian systems.* Russ. Math. Surv. **32**, No. 6, 5–66 (1977)

[8] Poincaré H.: *Les Methodes Nouvelles de la Mechanique Celeste*, Gauthier Villars, Paris (1892)

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

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