Journal of Nonlinear Science (2023) 33:90
https://doi.org/10.1007/s00332-023-09948-4

Journal of Nonlinear Science

Quantitative KAM Theory, with an Application to the Three-Body Problem

Gabriella Pinzari[1] · **Xiang Liu**[1]

Received: 24 February 2023 / Accepted: 18 July 2023

Abstract
Based on quantitative "KAM theory", we state and prove two theorems about the continuation of maximal and whiskered quasi-periodic motions to slightly perturbed systems exhibiting proper degeneracy. Next, we apply such results to prove that, in the three-body problem, there is a small set in phase space where it is possible to detect both such families of tori. We also estimate the density of such motions in proper ambient spaces. Up to our knowledge, this is the first proof of co-existence of stable and whiskered tori in a physical system.

Keywords Properly degenerate Hamiltonian · Symplectic coordinates · Symmetry reductions

Mathematics Subject Classification 37J40 · 37J11 · 37J06

Contents

Communicated by Alain Goriely.

✉ Gabriella Pinzari
pinzari@math.unipd.it

1 Department of Mathematics, University of Padova, Padua, Italy

Published online: 02 August 2023

Springer

1 Overview

1.1 Two KAM Theorems for Properly Degenerate Hamiltonian Systems We deal with Hamiltonians which meet the demand of being close-to-be-integrable [see, e.g., Gallavotti (1986)], but, in addition, with the number of degrees of freedom of perturbing term being possibly larger than the one of the unperturbed part. Such kind of Hamiltonians often arise in problems of celestial mechanics and are referred to as "properly degenerate", after (Arnold 1963). We denote them as

$$H(I, \varphi, p, q; \mu) = H_0(I) + \mu\, P(I, \varphi, p, q; \mu),$$

where the coordinates $(I, \varphi) = (I_1, \ldots, I_n, \varphi_1, \ldots, \varphi_n)$ are of "action-angle" kind (after a possible application of the Liouville–Arnold theorem to the unperturbed term), while (for our needs) the $(p, q) = (p_1, \ldots, p_m, q_1, \ldots, q_m)$ are "rectangular", namely, take value in a small ball (say, of radius (ε_0)) about some point (say, the origin). The symplectic form is standard:

$$\Omega = dI \wedge d\varphi + dp \wedge dq = \sum_{i=1}^{n} dI_i \wedge d\varphi_i + \sum_{i=1}^{m} dp_i \wedge dq_i .$$

We work in the real-analytic framework, which means that we assume that H admits a holomorphic extension on a complex neighborhood of the real "phase space" (namely, the domain)

$$\mathcal{P}_{\varepsilon_0} := V \times \mathbb{T}^n \times B^{2m}_{\varepsilon_0},$$

where $V \subset \mathbb{R}^n$ is bounded, open and connected, $(\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z}))$ is the "flat torus", B^{2m}_{ε} is the $2m$-dimensional ball around 0 of radius ε, relatively to some norm in $\mathbb{R}^{2m}$. In this framework, we present[1] two "KAM theorems" which deal with different situations. A basic assumption, common to both statements, and often referred to as "Kolmogorov condition", is:

(A_1) the map $I \to \partial_I H_0(I)$ is a diffeomorphism of V.

However, due to the proper degeneracy mentioned above, such assumption is to be reinforced with some statement concerning the perturbing term, or, more precisely, its Lagrange average

$$P_{\mathrm{av}}(I, p, q; \mu) := \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} P(I, \varphi, p, q; \mu) d^n \varphi$$

with respect to the φ-coordinates. Such extra-assumption will be different in the two statements; therefore, we quote them below.

[1] We refer to specialized literature for historical notices and constructive approaches to KAM theory: see, e.g., Gallavotti (1994), Gentile and Gallavotti (1995), Bonetto et al. (1998), Chierchia and Procesi (2019) and references therein.

The first result is a revisitation of the so-called Fundamental Theorem by V.I. Arnold, Arnold (1963). Such theorem has been already studied, generalized and extended in previous works (Chierchia and Pinzari 2010; Pinzari 2018). Here, we deal with the situation (not considered in the aforementioned papers) where $P_{\rm av}$ admits a "Birkhoff Normal Form" (BNF hereafter) about $(p, q) = (0, 0)$ of high[2] order; say s. As expected, a higher order of BNF allows to improve the measure of the "Kolmogorov set", namely the set given by the union of all KAM tori. We shall prove[3] the following

Theorem 1.1 *Assume* (A_1) *above and the following conditions:*

(A$_2$) $P_{\rm av}(I, p, q) = \sum_{j=1}^{s} \mathcal{P}_j(r; I) + {\rm O}_{2s+1}(p, q; I)$, *with* $r_i := \frac{p_i^2+q_i^2}{2}$ *and* $\mathcal{P}_j(r; I)$ *being a polynomial of degree* j *in* $r = (r_1, \cdots, r_m)$, *for some* $2 \le s \in \mathbb{N}$.
(A$_3$) *the* $m \times m$ *matrix* $\beta(I)$ *of the coefficients of the second-order term* $\mathcal{P}_2(r; I) = \frac{1}{2}\sum_{i,j=1}^{m} \beta_{ij}(I) r_i r_j$ *is non-degenerate:* $|\det \beta(I)| \ge$ const > 0 *for all* $I \in V$.

Then, there exist positive numbers $\varepsilon_* < \varepsilon_0$, C_* *and* c_* *such that, for*

$$0 < \varepsilon < \varepsilon_* , \qquad 0 < \mu < \frac{\varepsilon^{2s+2}}{C_*(\log \varepsilon^{-1})^{c_*}} . \tag{1}$$

one can find a set $\mathcal{K} \subset \mathcal{P}_\varepsilon$ *formed by the union of* H*-invariant* n*-dimensional Lagrangian tori, on which the* H*-motion is analytically conjugated to linear Diophantine quasi-periodic motions with frequencies* $(\omega_1, \omega_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ *with* $\omega_1 = O(1)$ *and* $\omega_2 = O(\mu)$. *The set* $\mathcal{K}$ *has positive Liouville–Lebesgue measure and satisfies*

$$\mathrm{meas}\mathcal{P}_\varepsilon > \mathrm{meas}\mathcal{K} > \left(1 - C_*\varepsilon^{s-\frac{3}{2}}\right) \mathrm{meas}\mathcal{P}_\varepsilon . \tag{2}$$

The second result deals with lower-dimensional quasi-periodic motions, the so-called whiskered tori. These are n-dimensional quasi-periodic motions (in a phase space of dimension $2n + 2m$), approached or reached at an exponential rate. For simplicity, in view of our application, we focus on the case $m = 1$. In addition, we allow a further degeneracy in the Hamiltonian: the unperturbed term H_0 may possibly depend not on all the I's, but only on a part of them.

Theorem 1.2 *Let* $m = 1$, *and let* H_0 *depend on the components* $I_1 = (I_{11}, \ldots, I_{1n_1})$ *of the* $I = (I_1, I_2)$*'s, with* $1 \le n_1 \le n := n_1 + n_2$. *Assume* (A_1) *with* I_1 *replacing* I *and, in addition, that*

(A$_2'$) $P_{\rm av}(I, p, q; \mu) = P_0(I, pq; \mu) + P_1(I, \varphi, p, q; \mu)$ *with* $\|P_1\| \le a\|P_0\|$;
(A$_3'$) $|\partial_{pq} P_0| \ge$ const > 0 *and* $|\det \partial^2_{I_2} P_0| \ge$ const > 0 *if* $n_2 \ne 0$.

[2] Arnold (1963), Chierchia and Pinzari (2010), Pinzari (2018) deal with the "minimal" case $s = 2$. The case $s = 2$ is called here "minimal" as we work in the framework of generalizations of the Kolmogorov condition (A_1) above. In Rüssmann (2001), Féjoz (2004), using different techniques, the case $s = 1$ has been considered.

[3] For simplicity of notations, we do not write μ among the arguments of the functions in Theorem 1.1 and 1.2.

Fix $\eta > 0$. *Then, there exist positive numbers* a_*, $\varepsilon_* < \varepsilon_0$, C_* *and* c_* *such that, if*

$$0 < \varepsilon < \varepsilon_*, \quad 0 < a < a_* \varepsilon^4 \,, \quad 0 < \mu < \frac{C_*(a\|P_0\|)^{1+\eta}}{(\log a^{-1})^{c_*}} \tag{3}$$

one can find a set $\mathcal{K}$ *formed by the union of* H*-invariant* n*-dimensional Lagrangian tori, on which the* H*-motion is analytically conjugated to linear Diophantine quasi-periodic motions with frequencies* $(\omega_1, \omega_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ *with* $\omega_1 = O(1)$ *and* $\omega_2 = O(\mu)$. *The projection* $\mathcal{K}_0$ *of set* $\mathcal{K}$ *on* $\mathcal{P}_0 := V \times \mathbb{T}^n$ *has positive Liouville–Lebesgue measure and satisfies*

$$\text{meas}\mathcal{P}_0 > \text{meas}\mathcal{K}_0 > \left(1 - C_*\sqrt{a}\right)\text{meas}\mathcal{P}_0 \,. \tag{4}$$

Furthermore, for any $\mathcal{T} \in \mathcal{K}$ *there exist two* $(n+1)$*-dimensional invariant manifolds* $\mathcal{W}_{\rm u}$, $\mathcal{W}_{\rm s} \subset \mathcal{P}_{\varepsilon_*}$ *such that* $\mathcal{T} = \mathcal{W}_{\rm u} \cap \mathcal{W}_{\rm s}$ *and the motions on* $\mathcal{W}_{\rm u}$, $\mathcal{W}_{\rm s}$ *leave, approach* $\mathcal{T}$ *at an exponential rate.*

Before we go on with describing how we aim to use the theorems above, we premise some comment.

(i) The conditions involving μ in (1) and (3) are not optimal. With a procedure similar to the one shown in Chierchia and Pinzari (2010, proof of Theorem 1.2, steps 1–4), one can show that they can be relaxed to, respectively

$$\mu < \frac{1}{C_*(\log \varepsilon^{-1})^{2b}}, \qquad \mu < \frac{1}{C_*(\log(a\|P_0\|)^{-1})^{2b}}$$

with some $C_*, b > 0$.

(ii) The careful bounds on the measure of the invariant sets provided in (2) and (4) are needed in view of our application. Indeed, we shall apply both the theorems above in order to prove that, in the three-body problem, closely to the *co-planar, co-circular, outer retrograde configuration* (see below for the exact definition), full-dimensional and "whiskered" quasi-periodic tori co-exist [the result was conjectured in Pinzari (2018)]. In the application, ε will correspond to the maximum eccentricity or inclination; a the semi-major axes ratio, and the use of a high-order BNF in Theorem 1.1 will be necessary because the size of the set goes to 0 with some power of ε ($s = 4$ will be enough for our application).

(iii) Following Chierchia and Gallavotti (1994), Theorem 1.2 might be extended to prove the existence of "diffusion paths" and "whisker ladders". We shall not do, as proving Arnold instability [in the sense of Arnold (1964)] for the system (5) below is not the purpose of this paper. We, however, remark that such kind of instability has been found for the *four-body problem* in a very similar framework (Clarke et al. 2022). We remark that proofs of chaos or Arnold instability in celestial mechanics are quite recent (Féjoz et al. 2014; Delshams et al. 2019), by the difficulty of overcoming the so-called problem of large gaps. See Guzzo et al. (2020) and references therein.

(iv) Another important aspect in view of the application described above is a rather standard consequence of the proof of Theorem 1.2: If P (namely, P_1) has an equilibrium at $(p, q) = 0$, then, along the motions of $\mathcal{K}$, the coordinates (p, q) remain fixed at (0, 0) (rather than varying closely to it), namely

$$\mathcal{K} \subset V \times \mathbb{T}^n \times \{(0,0)\}.$$

More generally, the stable and unstable invariant manifolds do not shift from the unperturbed ones:

$$\mathcal{W}_{\rm s} \subset \mathcal{P}_\varepsilon \cap \{q = 0\}, \qquad \mathcal{W}_{\rm u} \subset \mathcal{P}_\varepsilon \cap \{p = 0\}.$$

1.2 Application to the three-body problem We apply the results above to prove that, in a region of the phase space of the three-body problem, and under conditions that will be specified later, full dimensional and whiskered tori co-exist. We underline that the co-existence of such different kind of motions is not a mere consequence of the non-integrability of the system (as in such case the result would be somewhat expected) as it persists in two suitable integrable approximations of the system, close one to the other. Indeed, such motions will be found in a very small zone in the phase space of the three-body problem which simultaneously is in the neighborhood of an elliptic equilibrium of one of such approximations and in a hyperbolic one of the other. Such an occurrence is intimately related to the use of two different systems of coordinates, which are singular one with respect to the other, in the region of interest. The authors are not aware of the appearance of such phenomenon, previously.
After the "heliocentric reduction" of translational invariance, the three-body problem Hamiltonian with gravitational masses equal to m_0, μm_1 and μm_2 and Newton constant $\mathcal{G} \equiv 1$, takes the form of the two-particle system [see, e.g., Féjoz (2004), Laskar and Robutel (1995) for a derivation]:

$$\mathrm{H}_{3\mathrm{b}} = \sum_{i=1}^{2} \left(\frac{|y^{(\mathrm{i})}|^2}{2\mathrm{m}_i} - \frac{\mathrm{m}_i \mathrm{M}_i}{|x^{(\mathrm{i})}|} \right) + \mu \left(-\frac{m_1 m_2}{|x^{(1)} - x^{(2)}|} + \frac{y^{(1)} \cdot y^{(2)}}{m_0} \right) \tag{5}$$

with suitable values of $\mathrm{m}_i = m_i + \mathrm{O}(\mu)$, $\mathrm{M}_i = m_0 + \mathrm{O}(\mu)$. We consider the system in the Euclidean space, namely we take, in (5), $y^{(\mathrm{i})}, x^{(\mathrm{i})} \in \mathbb{R}^3$, with $x^{(1)} \neq x^{(2)}$.

We call *Kepler maps* the class of symplectic[4] coordinate systems $\mathcal{C} = (\Lambda_1, \Lambda_2, \ell_1, \ell_2, \mathrm{y}, \mathrm{x})$ for the Hamiltonian (5), where $\mathrm{y} = (\mathrm{y}_1, \ldots, \mathrm{y}_4)$, $\mathrm{x} = (\mathrm{x}_1, \ldots, \mathrm{x}_4)$, such that:

[4] Namely verifying

$$\Omega = d\Lambda \wedge d\ell + d\mathrm{y} \wedge d\mathrm{x} = \sum_{i=1}^{2} d\Lambda_i \wedge d\ell_i + \sum_{i=1}^{4} d\mathrm{y}_i \wedge d\mathrm{x}_i.$$

- $\Lambda_i = \mathrm{m}_i\sqrt{\mathrm{M}_i a_i}$, where a_i denotes the semi-major axis of the i^{th} instantaneous[5] ellipse;
- $\ell_1, \ell_2 \in \mathbb{T}$ are conjugated to Λ_1, Λ_2. Such angles are defined in a different way according to the choice of $\mathcal{C}$. In all known examples, they are related to the area spanned by the planet along the instantaneous ellipse.

Using a Kepler map, the Hamiltonian (5) takes the form

$$\mathrm{H}_{\mathcal{C}} = -\frac{\mathrm{m}_1^3\mathrm{M}_1^2}{2\Lambda_1^2} - \frac{\mathrm{m}_2^3\mathrm{M}_2^2}{2\Lambda_2^2} + \mu f_{\mathcal{C}}(\Lambda_1, \Lambda_2, \ell_1, \ell_2, \hat{\mathrm{y}}, \hat{\mathrm{x}}) \tag{6}$$

where $\hat{\mathrm{y}}$, $\hat{\mathrm{x}}$ include the couples $(\mathrm{y}_i, \mathrm{x}_i)$ such[6] that nor y_i nor x_i is negligible. $\hat{\mathrm{y}}$, $\hat{\mathrm{x}}$ are often called *degenerate coordinates*, because they do not appear in (6) when μ is set to zero. In other words, $\mathrm{H}_{\mathcal{C}}$ is a properly degenerate close-to-be-integrable system, in the sense of the previous paragraph.

We call *co-planar, co-circular, outer retrograde configuration* the configuration of two planets in circular and co-planar motions, with the angular momentum of the outer planet having opposite verse to the resulting one. In Pinzari (2018) it has been pointed out that, under a careful choice of $\mathcal{C}$ such configuration plays the rôle of an equilibrium for the (ℓ_1, ℓ_2)-averaged perturbing function

$$\overline{f}_{\mathcal{C}}(\Lambda_1, \Lambda_2, \hat{\mathrm{y}}, \hat{\mathrm{x}}) = \frac{1}{(2\pi)^2}\int_{[0,2\pi]^2} f_{\mathcal{C}}(\Lambda_1, \Lambda_2, \ell_1, \ell_2, \hat{\mathrm{y}}, \hat{\mathrm{x}})d\ell_1 d\ell_2.$$

But what matters more is that, closely to such equilibrium, there exist two such $\mathcal{C}_i$'s such that the Hamiltonian $\mathrm{H}_{\mathcal{C}_1}$ is suited to Theorem 1.1, while $\mathrm{H}_{\mathcal{C}_2}$ is suited to Theorem 1.2. This leads to the following result, which states co-existence of stable and whiskered quasi-periodic motions in the three-body problem. It will be made more precise (see Theorem 2.1) and proved along the paper.

Theorem A *In the vicinity of the co-planar, co-circular, outer retrograde configuration, and provided that the masses of the planets and the semi-axes ratio are small, there exists a positive measure set $\mathcal{K}_1$ made of* 5*-dimensional quasi-periodic motions $\mathcal{T}_1$'s "surrounding"* (*in a sense which will be specified*) 3*-dimensional quasi-periodic motions $\mathcal{T}_2$'s, each equipped with two invariant manifolds, called, respectively, unstable, stable manifold, where the motions are respectively asymptotic to the $\mathcal{T}_2$'s in the past, in the future.*

We conclude with saying how this paper is organized.

- In Sects. 2.1 and 2.2 we recall the main arguments of the discussion in Pinzari (2018), which lead to put the system (5) to a form suited to apply Theorems 1.1 and 1.2.

[5] With reference to the three-body Hamiltonian (5), the i^{th} *instantaneous ellipse* is the orbit generated by $\mathrm{h}_i := \frac{|y^{(\mathrm{i})}|^2}{2\mathrm{m}_i} - \frac{\mathrm{m}_i\mathrm{M}_i}{|x^{(\mathrm{i})}|}$ in a region of phase space where h_i is negative.

[6] The reason of this is that the Hamiltonian (5) has first integrals, as recalled in the next section.

- In Sects. 2.3 and 2.4 we check that the two domains where Theorems 1.1 and 1.2 apply have a non-empty intersection, and such intersection includes both families of tori. This check is subtle, because of the difference of the frameworks used.
- In Sect. 3, we prove Theorems 1.1 and 1.2 via a carefully quantified KAM theory.

2 Ellipticity and Hyperbolicity Closely to Co-planar, Co-circular, Outer Retrograde Configuration

Putting the system in a form suited to Theorem 1.1 requires identifying an elliptic equilibrium, while Theorem 1.2 calls for a hyperbolic one.

Denoting as $(\mathrm{C}^{(j)} := x^{(j)} \times y^{(j)}$ the angular momenta of the planets, we proceed to study motions evolving from initial data close to the manifold

$$\mathcal{M}_\pi := \Big\{(y, x) : \ \mathrm{C}^{(1)} \parallel (-\mathrm{C}^{(2)}) \parallel \mathrm{C}, \ \text{and } x^{(1)}, \ x^{(2)} \text{ describe circular motions.}\Big\}. \tag{7}$$

The sub-fix "π" recalls that $\mathrm{C}^{(1)}$ and $\mathrm{C}^{(2)}$ are opposite. In the two next sections, we recall material from Pinzari (2018), which highlights a sort of "double (elliptic, hyperbolic) nature" of $\mathcal{M}_\pi$.

2.1 Ellipticity (with BNF)

Basically[7], the construction of the elliptic equilibrium—and of its associated BNF—proceeds as in Chierchia and Pinzari (2011). We briefly resume the procedure here. We fix a domain $\mathcal{D}_c \subset \mathbb{R}^{12}$ for impulse-position "Cartesian" coordinates

$$c = (y, x) := (y^{(1)}, y^{(2)}, x^{(1)}, x^{(2)})$$

of two point masses relatively to a prefixed orthonormal frame $(k^{(1)}, k^{(2)}, k^{(3)})$ in $\mathbb{R}^3$. As a first step, we switch to a set of coordinates, well known in the literature, which we name JRD, after C. G. J. Jacobi, R. Radau and A. Deprit (Jacobi 1842; Radau 1868; Deprit 1983), who, at different stages, contributed to their construction.
We fix a region of phase space where the orbits $t \to (x^{(j)}(t), y^{(j)}(t))$ generated by the unperturbed "Kepler" Hamiltonians

$$\mathrm{h}_\mathrm{k}^{(j)} := \frac{|y^{(j)}|^2}{2\mathrm{m}_j} - \frac{\mathrm{m}_j\mathrm{M}_j}{|x^{(j)}|}$$

in (5) are ellipses with non-vanishing eccentricity. Then, we denote as $\mathrm{P}^{(j)}$ the unit vectors pointing in the directions of the perihelia; as a_j the semi-major axes; as ℓ_j the

[7] As pointed out in Pinzari (2018), the only note-worthing difference with the case studied in Chierchia and Pinzari (2011) (which deals with prograde motions of the planets, namely, revolving all in the same verse) is that here the elliptic character of the equilibrium does not follow for free from the symmetry of the Hamiltonian, but is checked manually.

"mean anomaly" of $x^{(j)}$(which, we recall, is defined as area of the elliptic sector from $\mathrm{P}^{(j)}$ to $x^{(j)}$ "normalized at 2π"); as $\mathrm{C}^{(j)} = x^{(j)} \times y^{(j)}$, $j = 1, 2$, the angular momenta of the two planets and $\mathrm{C} := \mathrm{C}^{(1)} + \mathrm{C}^{(2)}$ the total angular momentum integral. We assume that the "nodes"

$$\nu_1 := k^{(3)} \times \mathrm{C}\,, \quad \nu := \mathrm{C} \times \mathrm{C}^{(1)} = \mathrm{C}^{(2)} \times \mathrm{C}^{(1)} \tag{8}$$

do not vanish, anytime. Such condition is equivalent to ask that the planes determined by the instantaneous ellipses and the $(k^{(1)}, k^{(2)})$ plane never pairwise coincide. As in previous works, we use the following notations. For three vectors u, v, w with u, $v \perp w$, we denote as $\alpha_w(u, v)$ the angle formed by u to v relatively to the positive (counterclockwise) orientation established by w. Then, the JRD coordinates are here denoted with the symbols

$$jrd := \left(\widehat{jrd} := (\Lambda_1, \Lambda_2, \mathrm{G}_1, \mathrm{G}_2, \ell_1, \ell_2, \gamma_1, \gamma_2), (\mathrm{G}, \mathrm{Z}, \gamma, \zeta)\right) \in \mathbb{R}^4 \times \mathbb{T}^4 \times \mathbb{R}^2 \times \mathbb{T}^2 \tag{9}$$

and defined via the formulae

$$\begin{cases} \mathrm{Z} := \mathrm{C} \cdot k^{(3)} \\ \mathrm{G} := \|\mathrm{C}\| \\ \mathrm{G}_1 := \|\mathrm{C}^{(1)}\| \\ \mathrm{G}_2 := \|\mathrm{C}^{(2)}\| \\ \Lambda_j := \mathrm{M}_j \sqrt{\mathrm{m}_j a_j} \end{cases} \qquad \begin{cases} \zeta := \alpha_{k^{(3)}}(k^{(1)}, \nu_1) \\ \gamma := \alpha_{\mathrm{C}}(\nu_1, \nu) \\ \gamma_1 := \alpha_{\mathrm{C}^{(1)}}(\nu, \mathrm{P}^{(1)}) \\ \gamma_2 := \alpha_{\mathrm{C}^{(2)}}(\nu, \mathrm{P}^{(2)}) \\ \ell_j := \text{mean anomaly of } x^{(j)} \end{cases} \tag{10}$$

The main point of JRD is that Z, ζ and γ are ignorable coordinates and G is constant along the motions of SO(3)-invariant systems. Therefore, most of motions of SO(3)-invariant systems are effectively described by the "reduced" coordinates $\widehat{jrd}$. This strong property cannot be exploited in the case study of the paper, as the manifold $\mathcal{M}_\pi$ in (7) is a singularity of the change (10). More generally, any co-planar or circular[8] configuration is so. Pretty similarly as in Chierchia and Pinzari (2011), we bypass such difficulty switching to new coordinates denoted as

$$rps_\pi := \left(\widehat{rps}_\pi := (\Lambda_1, \Lambda_2, \lambda_1, \lambda_2, \eta_1, \eta_2, \xi_1, \xi_2, p, q), (Z, \zeta)\right)$$

[8] Circular configurations correspond to $\mathrm{G}_i = \Lambda_i$; co-planar configurations correspond to $\mathrm{G} = \sigma_1\mathrm{G}_1 + \sigma_2\mathrm{G}_2$, with $(\sigma_1, \sigma_2) \in \{\pm 1\}^2 \setminus \{(-1, -1)\}$.

where the Λ_j's, Z and ζ are the same[9] as in (9), while

$$\begin{cases} \lambda_1 = \ell_1 + \gamma_1 + \gamma \\ \lambda_2 = \ell_2 + \gamma_2 - \gamma \\ \eta_1 + \mathrm{i}\xi_1 = \sqrt{2(\Lambda_1 - \mathrm{G}_1)}e^{-\mathrm{i}(\gamma_1+\gamma)} \\ \eta_2 + \mathrm{i}\xi_2 = -\sqrt{2(\Lambda_2 - \mathrm{G}_2)}e^{\mathrm{i}(-\gamma_2+\gamma)} \\ p + \mathrm{i}q = -\sqrt{2(\mathrm{G} + \mathrm{G}_2 - \mathrm{G}_1)}e^{\mathrm{i}\gamma} \end{cases} \tag{11}$$

As in JRD, (Z, ζ) is a cyclic couple in SO(3)-invariant Hamiltonians but now no more cyclic coordinates but it appears. This leaves the system with 5 degrees of freedom and an extra-integral: the action G written using rps_π:

$$\mathrm{G}_{rps_\pi} := \Lambda_1 - \Lambda_2 - \frac{\eta_1^2 + \xi_1^2}{2} + \frac{\eta_2^2 + \xi_2^2}{2} + \frac{p^2 + q^2}{2}. \tag{12}$$

We denote as

$$\begin{aligned} \mathrm{H}_{rps_\pi} &:= -\frac{\mathrm{m}_1^3\mathrm{M}_1^2}{2\Lambda_1^2} - \frac{\mathrm{m}_2^3\mathrm{M}_2^2}{2\Lambda_2^2} + \mu\left(- \frac{m_1 m_2}{|x^{(1)}_{rps_\pi} - x^{(2)}_{rps_\pi}|} + \frac{y^{(1)}_{rps_\pi} \cdot y^{(2)}_{rps_\pi}}{m_0}\right) \\ &=: \mathrm{h_k}(\Lambda) + \mu f_{rps_\pi}(\Lambda, \lambda, z) \qquad z := (\eta, \xi, p, q) \end{aligned} \tag{13}$$

the Hamiltonian (5) written in RPS_π coordinates, and worry about it.
We note that the manifold $\mathcal{M}_\pi$ in (7) is now given by

$$\mathcal{M}_\pi = \left\{rps_\pi \,:\, z = 0\right\}.$$

Then we consider a neighborhood of $\mathcal{M}_\pi$ of the form

$$\mathcal{M}_{rps_\pi,\varepsilon_0} := \mathcal{L} \times \mathbb{T}^2 \times B^6_{\varepsilon_0}(0)\ ,$$

where $B^6_{\varepsilon_0}$ is the 6-ball centered at $0 \in \mathbb{R}^6$ with radius ε_0; $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$ and $\mathcal{L}$ is defined as

$$\mathcal{L} := \Big\{\Lambda = (\Lambda_1, \Lambda_2) : \ \Lambda_- < \Lambda_2 < \Lambda_+\ , \quad k_-\Lambda_2 < \Lambda_1 < k_+\Lambda_2\Big\}. \tag{14}$$

[9] There is an inessential difference between the Definition (11) and the one in Pinzari (2018, Eqs. (25), (26))). Denoting as $\overline{rps}_\pi := \Big((\Lambda_1, \Lambda_2, \overline{\lambda}_1, \overline{\lambda}_2, \overline{\eta}_1, \overline{\eta}_2, \overline{\xi}_1, \overline{\xi}_2, \overline{p}, \overline{q}), (P, Q)\Big)$ the coordinates defined in Pinzari (2018), we have the following relations:

$$\begin{cases} \overline{\lambda}_1 = \lambda_1 + \zeta \\ \overline{\lambda}_2 = \lambda_2 - \zeta \\ \overline{\eta}_1 + \mathrm{i}\overline{\xi}_1 = (\eta_1 + \mathrm{i}\xi_1)e^{-\mathrm{i}\zeta} \\ \overline{\eta}_2 + \mathrm{i}\overline{\xi}_2 = (\eta_2 + \mathrm{i}\xi_2)e^{\mathrm{i}\zeta} \\ \overline{p} + \mathrm{i}\overline{q} = (p + \mathrm{i}q)e^{\mathrm{i}\zeta} \\ P + \mathrm{i}Q = \sqrt{2(\mathrm{G} - \mathrm{Z})}\, e^{-\mathrm{i}\zeta} \end{cases}$$

But as (Z, ζ), (P, Q) and ζ do not appear in the Hamiltonian, its expression does not change.

Here, $0 < \Lambda_- < \Lambda_+$ are arbitrarily taken (more conditions on such numbers will be specified in the course of the paper) and, for fixed positive[10] numbers $0 < \alpha_- < \alpha_+ < 1$, $k_\pm$ are constants depending on $\alpha_\pm$ and the masses via

$$k_\pm := \frac{m_1}{m_2}\sqrt{\frac{m_0+\mu m_2}{m_0+\mu m_1}\alpha_\pm}\,. \tag{15}$$

We now take $0 < \delta < 1$ and[11] and assume

$$0 < \frac{m_2}{m_1} < \min\left\{\sqrt{(1-\delta)\alpha_-}\ ,\ 1-\delta\right\},\quad 0 < \mu < \mu_0(\delta) := \frac{\delta m_0}{m_1(1-\delta)-m_2} \tag{16}$$

Then we[12] have

[10] Observe that α_- and α_+ have the meaning of lower and upper bound for the semi-major axes ratio $\alpha = a_1/a_2$, namely,

$$\alpha_- \le \alpha \le \alpha_+ \qquad \forall\,(\Lambda_1, \Lambda_2) \in \mathcal{L}\,.$$

Indeed, from the formula

$$\frac{\Lambda_1}{\Lambda_2} = \frac{m_1}{m_2}\sqrt{\frac{m_0+\mu m_2}{m_0+\mu m_1}\alpha}$$

we find

$$\alpha \le \left(\frac{m_2}{m_1}\right)^2 \frac{m_0+\mu m_1}{m_0+\mu m_2} k_+^2 = \alpha_+$$

and, similarly, $\alpha \ge \alpha_-$.

[11] The reader might ask the reason of inequalities in (16). This is related to the fact that we want to investigate a region of phase space where the inner planet, labeled as "1", has a larger angular momentum, namely, $\mathrm{G}_1 > \mathrm{G}_2$, and, simultaneously, the masses of the planets, as well as their eccentricities and mutual inclination are small. As, when eccentricities and mutual inclination go to zero, the $\mathrm{G_i}$ reduce to Λ_i, by (14), the number k_- in (15) needs to be strictly larger than 1. Conditions (16) are apt to ensure this, as in fact they immediately imply

$$1-\delta \le \frac{m_0+\mu m_2}{m_0+\mu m_1} \le 1+\delta$$

hence, by (15),

$$k_- \ge \frac{m_1}{m_2}\sqrt{(1-\delta)\alpha_-} > 1.$$

[12] The proof in Pinzari (2018, Appendix A) is given with $\delta = 1 - \frac{1}{4\chi^2} \ge \frac{3}{4}$, but works well also for any $\delta \in (0, 1)$. Indeed, for $(\Lambda_1, \Lambda_2) \in \mathcal{L}$,

$$\Lambda_1 - \Lambda_2 = \Lambda_2\left(\frac{\Lambda_1}{\Lambda_2} - 1\right) \ge \Lambda_-(k_-1) \ge \Lambda_-\left(\frac{m_1}{m_2}\sqrt{(1-\delta)\alpha_-} - 1\right).$$

Proposition 2.1 [Pinzari (2018, Section III and Appendix A)] *One can find* $\varepsilon_0 > 0$*, depending only on* Λ_-*,* δ*,* α_-*,* m_1*,* m_2 *such that the function* H_{rps_π} *in* (13) *is real-analytic*[13] *for* $(\Lambda, \lambda, \eta, \xi, p, q) \in \mathcal{M}_{rps_\pi, \varepsilon_0}$*. In addition, for any* $s \in \mathbb{N}$*, there exists a positive number* $\alpha^\#$ *such that, if* $\alpha_+ < \alpha^\#$*, there exists a positive number* $\varepsilon_1 < \varepsilon_0$ *and a real-analytic canonical transformation*

$$\phi_{bnf}: \quad (\Lambda, \overline{\lambda}, \overline{\eta}, \overline{\xi}, \overline{p}, \overline{q}) \in \mathcal{M}_{rps_\pi, \varepsilon_1} \to (\Lambda, \lambda, \eta, \xi, p, q) \in \mathcal{M}_{rps_\pi, \varepsilon_0}$$

which carries $(\overline{\eta}, \overline{\xi}, \overline{p}, \overline{q}) = 0$ *to* $(\eta, \xi, p, q) = 0$ *for all* $(\overline{\Lambda}, \overline{\lambda}) \in \mathcal{L} \times \mathbb{T}^2$*, such that, if*

$$\mathrm{H}_{bnf} := \mathrm{H}_{rps_\pi} \circ \phi_{bnf} = \mathrm{h_k}(\Lambda) + \mu f_{bnf}(\Lambda, \overline{\lambda}, \overline{\eta}, \overline{\xi}, \overline{p}, \overline{q}) \tag{17}$$

then the averaged perturbing function

$$f_{bnf}^{\mathrm{av}}(\Lambda, \overline{\eta}, \overline{\xi}, \overline{p}, \overline{q}) := \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} f(\Lambda, \overline{\lambda}, \overline{\eta}, \overline{\xi}, \overline{p}, \overline{q}) d\overline{\lambda}_1 d\overline{\lambda}_2$$

"is in Birkhoff Normal Form of order s", namely:

$$f_{bnf}^{\mathrm{av}} = C_0(\Lambda) + \Omega \cdot \overline{\tau} + \frac{1}{2}\overline{\tau} \cdot \mathrm{T}(\Lambda)\overline{\tau} + \not\Vdash_{s \geq 3} \sum_{j=3}^{s} \mathcal{P}_j(\overline{\tau}; \Lambda) + \mathrm{O}_{2s+1}(\overline{\eta}, \overline{\xi}, \overline{p}, \overline{q}; \Lambda)$$

where $\Omega(\Lambda) = (\Omega_1(\Lambda), \Omega_2(\Lambda), \Omega_3(\Lambda))$*;* $\mathcal{P}_j(\overline{\tau}; \Lambda)$ *are homogeneous polynomials of degree* j *in* $\overline{\tau} := \left(\frac{\overline{\eta}_1^2 + \overline{\xi}_1^2}{2}, \frac{\overline{\eta}_2^2 + \overline{\xi}_2^2}{2}, \frac{\overline{p}^2 + \overline{q}^2}{2}\right)$ *and the determinant of the* 3×3 *matrix* $\mathrm{T}(\Lambda)$ *does not identically vanish. Moreover,* ϕ_{bnf} *leaves* G_{rps_π} *unvaried, meaning that the function*

$$\overline{\mathrm{G}} := \Lambda_1 - \Lambda_2 - \frac{\overline{\eta}_1^2 + \overline{\xi}_1^2}{2} + \frac{\overline{\eta}_2^2 + \overline{\xi}_2^2}{2} + \frac{\overline{p}^2 + \overline{q}^2}{2}$$

is still a first integral to $\overline{\mathrm{H}}$.

Therefore, for (Λ_1, Λ_2) on a complex neighborhood of $\mathcal{L}$ depending on Λ_-, m_1, m_2, α_- and δ we shall have $|\Lambda_1 - \Lambda_2| \geq \frac{\Lambda_-}{2}\left(\frac{m_1}{m_2}\sqrt{(1-\delta)\alpha_-} - 1\right)$ and, as in the proof of Pinzari (Pinzari (2018), Proposition III.2), one can take $\varepsilon_0 < \frac{\Lambda_-}{2}\left(\frac{m_1}{m_2}\sqrt{(1-\delta)\alpha_-} - 1\right)$ in order that the denominators of the functions c_1^*, c_2, c_2^* in (Pinzari 2018, Appendix A) do not vanish, and so small that collisions are excluded.

[13] Namely, analytic on a complex neighborhood of $\mathcal{M}_{rps_\pi, \varepsilon_0}$ and real-valued on $\mathcal{M}_{rps_\pi, \varepsilon_0}$.

2.2 Hyperbolicity

The hyperbolic character appears using a set of canonical coordinates, named *perihelia reduction* (P-*coordinates*). This is a further set of canonical coordinates

$$\mathrm{P} := \left(\widehat{\mathrm{P}}, (\mathrm{Z}, \mathrm{G}, \zeta, \mathrm{g})\right) \in \mathbb{R}^{3n-2} \times \mathbb{T}^{3n-2} \times \mathbb{R}^2 \times \mathbb{T}^2 \tag{18}$$

performing full reduction of SO(3) invariance for a n-particle system, which, in addition keeps regular for co-planar motions. The P-coordinates have been firstly introduced in Pinzari (2018), to which we refer for the proof of their canonical character. We remark that in (18), G, Z and ζ are the same as in JRD in (10). The coordinate g, conjugated to G, is not the same as in (10), but of course $(\mathrm{Z}, \zeta, \mathrm{g})$ are again ignorable and G is constant in SO(3) invariant systems. For the 3-body problem, namely, $n = 2$, the 8-plet $\widehat{\mathrm{P}}$ is given by

$$\widehat{\mathrm{P}} := (\Lambda_1, \Lambda_2, \mathrm{G}_2, \Theta, \ell_1, \ell_2, \mathrm{g}_2, \vartheta)$$

with Λ_j, ℓ_j, G_2 as in (10). To define Θ, g, ϑ and g_2, we assume that

$$\nu_1 := k^{(3)} \times \mathrm{C}, \qquad \mathrm{n}_1 := \mathrm{C} \times \mathrm{P}^{(1)}, \qquad \nu_2 := \mathrm{P}^{(1)} \times \mathrm{C}^{(2)}, \quad \mathrm{n}_2 = \mathrm{C}^{(2)} \times \mathrm{P}^{(2)} \tag{19}$$

do not vanish. Note that ν_1 in (19) is the same as in (8). We[14] let (under the same notations as in the previous section)

$$\Theta := \mathrm{C} \cdot \mathrm{P}^{(1)} = \mathrm{C}^{(2)} \cdot \mathrm{P}^{(1)} \qquad \begin{cases} \vartheta := \alpha_{\mathrm{P}^{(1)}}(\mathrm{n}_1, \nu_2) \\ \mathrm{g} := \alpha_{\mathrm{C}}(\nu_1, \mathrm{n}_1) \\ \mathrm{g}_2 := \alpha_{\mathrm{C}^{(2)}}(\nu_2, \mathrm{n}_2) \end{cases} \tag{20}$$

We now describe the rôle of the P-coordinates in the Hamiltonian (5). We denote as

$$\mathrm{H}_{\mathrm{P}} = \mathrm{h}_{\mathrm{k}}(\Lambda_1, \Lambda_2) + \mu f_p(\Lambda_1, \Lambda_2, \mathrm{G}_2, \Theta; \ell_1, \ell_2, \mathrm{g}_2, \vartheta; \mathrm{G})$$

where

$$\mathrm{h}_{\mathrm{k}}(\Lambda_1, \Lambda_2) = -\frac{\mathrm{m}_1^3 \mathrm{M}_1^2}{2\Lambda_1^2} - \frac{\mathrm{m}_2^3 \mathrm{M}_2^2}{2\Lambda_2^2}, \qquad f_p = -\frac{m_1 m_2}{|x_p^{(1)} - x_p^{(2)}|} + \frac{y_p^{(1)} \cdot y_p^{(2)}}{m_0}.$$

the Hamiltonian (5) expressed in terms of p, and

$$f_p^{\mathrm{av}} := \frac{1}{(2\pi)^2} \int_{[0,2\pi]^2} f_p d\ell_1 d\ell_2$$

[14] The second equality in the first equation in (20) is implied by $\mathrm{C} = \mathrm{C}^{(1)} + \mathrm{C}^{(2)}$ and $\mathrm{C}^{(1)} \cdot \mathrm{P}^{(1)} = 0$.

the doubly averaged perturbing function. We look at the expansion

$$f_p^{\rm av} = -\frac{m_1 m_2}{a_2}\Big(1+\alpha^2 {\rm P} + {\rm O}(\alpha^3)\Big)$$

where $\alpha := \frac{a_1}{a_2}$ is the semi-major axes ratio. We focus on the function P. Let $\mathcal{L}$ as in (14); $c \in (0, 1)$, and put

$$\begin{aligned}\mathcal{L}_p({\rm G}) \ :=& \Big\{\Lambda = (\Lambda_1, \Lambda_2) \in \mathcal{L}: \quad \Lambda_1 > {\rm G} + \frac{2}{c}\sqrt{\alpha_+}\Lambda_2\\ & 5\Lambda_1^2{\rm G} - ({\rm G} + \frac{2}{c}\sqrt{\alpha_+}\Lambda_1)^2(4{\rm G} + \frac{2}{c}\sqrt{\alpha_+}\Lambda_1) > 0,\\ & 5\Lambda_1^2{\rm G} - ({\rm G}+\Lambda_2)(4{\rm G}+\Lambda_2) > 0\\ & \Lambda_2 > {\rm G}\,, \quad \Lambda_1 > 2{\rm G}\Big\} \qquad (21)\\ \mathcal{G}_p(\Lambda_1, \Lambda_2, {\rm G}) \ :=& \Big\{{\rm G}_2 : \ \max\{\frac{2}{c}\sqrt{\alpha_+}\Lambda_2, {\rm G}\} < {\rm G}_2 < \Lambda_2\Big\}\\ \mathcal{B}_p({\rm G}) \ :=& \Big\{(\Theta, \vartheta) : \ |\Theta| < \frac{\rm G}{2}, \ |\vartheta| < \frac{\pi}{2}\Big\} \qquad (22)\end{aligned}$$

and finally

$$\mathcal{A}_p({\rm G}) \ := \ \Big\{(\Lambda_1, \Lambda_2, {\rm G}_2) : \ (\Lambda_1, \Lambda_2) \in \mathcal{L}_p({\rm G}), \ {\rm G}_2 \in \mathcal{G}_p(\Lambda_1, \Lambda_2)\Big\}$$

Moreover, we let

$$\mathcal{N}({\rm G}) := \mathcal{A}_p({\rm G}) \times \mathbb{T}^3 \times \mathcal{B}_p({\rm G}), \quad \mathcal{N}_0({\rm G}) := \mathcal{A}_p({\rm G}) \times \mathbb{T}^3 \times \big\{0, 0\big\}. \qquad (23)$$

Note that phase points in $\mathcal{N}_0$ has the geometrical meaning of co-planar motions with the outer planet in retrograde motion.

Proposition 2.2 (Pinzari 2018, Section IV) *The 4 degrees of freedom Hamiltonian* H_p *is real-analytic in* $\mathcal{N}$. *It has an equilibrium on* $\mathcal{N}_0$. *Such equilibrium turns to be hyperbolic*[15] *for* P.

2.3 Existence and Co-Existence of two Families of Tori

Theorems 1.1 and 1.2 can now be used to prove the existence of both full-dimensional and whiskered, co-dimension 2 tori in the three-body problem. Indeed,

---

[15] In Pinzari (2018) a slightly more general result is proved: the equilibrium is hyperbolic when $\mathcal{L}_p$ in (21) is defined without the inequality

$$5\Lambda_1^2{\rm G} - ({\rm G}+\Lambda_2)(4{\rm G}+\Lambda_2) > 0 \qquad (24)$$

and $\mathcal{G}_p$ in (22) is taken to be $\Big\{{\rm G}_2 : \ \max\{\frac{2}{c}\sqrt{\alpha_+}\Lambda_2, {\rm G}\} < {\rm G}_2 < \min\{\Lambda_2, {\rm G}^\star\}\Big\}$, with ${\rm G}^\star$ the unique root of the polynomial ${\rm G}_2 \to 5\Lambda_1^2{\rm G} - ({\rm G}+{\rm G}_2)(4{\rm G}+{\rm G}_2)$. But as (24) ensures $\Lambda_2 < {\rm G}^\star$, under such restriction, $\mathcal{G}_p$ can be taken as in (22).

– Under conditions (1), by Theorem 1.1, an invariant[16] set $\mathcal{F} \subset \mathcal{M}_\varepsilon$ for the Hamiltonian H_{rps_π} with 5-dimensional frequencies is found, whose measure satisfies

$$\mathrm{meas}\mathcal{M}_\varepsilon > \mathrm{meas}\mathcal{F} > \left(1 - C_* \varepsilon^{\frac{1}{2}+\overline{s}}\right) \mathrm{meas}\mathcal{M}_\varepsilon \tag{25}$$

where $\overline{s} = s - 2$.
– Under conditions (3) with $a = \alpha_+$, by Theorem 1.2, for any $\mathrm{G} \in \mathbb{R}_+$, one finds an invariant set $\mathcal{H}(\mathrm{G}) \subset \mathcal{N}_0(\mathrm{G})$ with 3-dimensional frequencies for H_p and equipped with 4-dimensional stable and unstable manifolds[17], whose measure satisfies

$$\mathrm{meas}\mathcal{N}_0(\mathrm{G}) > \mathrm{meas}\mathcal{H}(\mathrm{G}) > \left(1 - C_* \sqrt{\alpha_+}\right) \mathrm{meas}\mathcal{N}_0(\mathrm{G}) . \tag{26}$$

In the next, we show that the invariant sets $\mathcal{F}$ and $\mathcal{H}(\mathrm{G})$ constructed above "have a common domain of existence". We have to make this assertion more precise, mainly because $\mathcal{F}$ and $\mathcal{H}(\mathrm{G})$ have been constructed with different formalisms.
Let

$$\phi^p_{rps_\pi} : rps_\pi \to p \tag{27}$$

the canonical change of coordinates between rps_π and p, well defined in a full measure set.
Let $\mathbb{G}_*$, $\mathbb{G}_0$ the respective images under the function (12):

$$\mathbb{G}_0 := \mathrm{G}_{rps_\pi}(\mathcal{M}_\varepsilon), \qquad \mathbb{G}_* := \mathrm{G}_{rps_\pi}(\mathcal{F})$$

of the sets $\mathcal{M}_\varepsilon$, $\mathcal{F}$. As $\mathcal{F} \subset \mathcal{M}_\varepsilon$, then $\mathbb{G}_* \subset \mathbb{G}_0$. For any $\mathrm{G}_0 \in \mathbb{G}_0$, $\mathrm{G}_* \in \mathbb{G}_*$, let

$$\mathcal{M}_\varepsilon(\mathrm{G}_0) := \mathcal{M}_\varepsilon \cap \{\mathrm{G}_{rps_\pi} = \mathrm{G}_0\}, \qquad \mathcal{F}(\mathrm{G}_*) := \mathcal{F} \cap \{\mathrm{G}_{rps_\pi} = \mathrm{G}_*\}$$

$\mathcal{M}_\varepsilon(\mathrm{G}_0)$ and $\mathcal{F}(\mathrm{G}_*)$ are invariant sets because G_{rps_π} is conserved along the motions of H_{rps}.
Define:

$$\mathcal{M}'_\varepsilon(\mathrm{G}_0) := \phi^p_{rps_\pi}(\mathcal{M}_\varepsilon(\mathrm{G}_0)), \qquad \mathcal{F}'(\mathrm{G}_*) := \phi^p_{rps_\pi}(\mathcal{F}(\mathrm{G}_*)) .$$

---

[16] More precisely, Theorem 1.1 is applied to the Hamiltonian H_{bnf} in (17), hence with

$$n_1 = 2,\ n_2 = 3,\ V = \mathcal{L},\ \varepsilon = \varepsilon_1,\ H_0 = \mathrm{h_k},\ P = f_{bnf}$$

corresponding to the image under $\overline{\phi}$ of the invariant set obtained through the thesis of Theorem 1.1.
[17] Theorem 1.2 is applied to the Hamiltonian H_p of Proposition 2.2, hence with

$$n_1 = 2,\ n_2 = 1,\ V = \mathcal{A}_p(\mathrm{G})$$
$$H_0 = \mathrm{h_k} - \frac{m_1 m_2}{a_2},\ P_0 = -\frac{m_1 m_2}{a_2}\alpha^2 \mathrm{P},\ P_1 = -\frac{m_1 m_2}{a_2}\mathrm{O}(\alpha^3),\ a = \alpha_+$$

At the cost of eliminating zero-measure sets from $\mathbb{G}_0$, $\mathbb{G}_*$, the sets $\mathcal{F}'(\mathrm{G}_*)$, $\mathcal{M}'_\varepsilon(\mathrm{G}_0)$ are well-defined, for all $\mathrm{G}_0 \in \mathbb{G}_0$, $\mathrm{G}_* \in \mathbb{G}_*$. Then split

$$\mathcal{M}'_\varepsilon(\mathrm{G}_0) = \widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_0) \times \{\mathrm{G} = \mathrm{G}_0,\ \mathrm{g} \in \mathbb{T}\} \qquad \mathcal{F}'(\mathrm{G}_*) = \widehat{\mathcal{F}}'(\mathrm{G}_*) \times \{\mathrm{G} = \mathrm{G}_*,\ \mathrm{g} \in \mathbb{T}\}$$

The volume-preserving property of $\phi^p_{rps_\pi}$ in (27), the monotonicity of the Lebesgue integral and the bounds in (25) guarantee that

$$\mathrm{meas}\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*) > \ \mathrm{meas}\widehat{\mathcal{F}}'(\mathrm{G}_*) > \left(1 - C_1\varepsilon^{\frac{1}{2}+\bar{s}}\right)\mathrm{meas}\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*) \qquad \forall\ \mathrm{G}_* \in \mathbb{G}_* \tag{28}$$

with some $C_1 > 0$.

Recall now the definition of $\mathcal{N}(\mathrm{G})$, $\mathcal{N}_0(\mathrm{G})$ in (23) and $\mathcal{H}(\mathrm{G})$ in (26). The main result of the paper is the following

Theorem 2.1 *Let $\sigma > 0$ half-integer. There exist ε_*, $c_0 \in (0, 1)$ such that, if $\varepsilon < \varepsilon_*$, $\mathrm{G}_* \in \mathbb{G}_*$, $\mathrm{G}_* > c_0^{-1}\varepsilon^2$, $\alpha_+ \le c_0\varepsilon^{12}$ and μ verifies (1), (3) with $a = \alpha_+$ and $s = \sigma + \frac{7}{2}$, then there exists a non-empty set $\mathcal{A}_\star(\mathrm{G}_*)$ such that, letting*

$$\mathcal{Q}(\mathrm{G}_*) := \mathcal{A}_\star(\mathrm{G}_*) \times \mathbb{T}^3 \times \mathcal{B}_1(\varepsilon, \mathrm{G}_\star), \qquad \mathcal{Q}_0(\mathrm{G}_*) := \mathcal{A}_\star(\mathrm{G}_*) \times \mathbb{T}^3 \times \{(0, 0)\}$$

and denoting $\widehat{\mathcal{F}}'_*(\mathrm{G}_*)$, $\widehat{\mathcal{H}}_*(\mathrm{G}_*)$ the respective intersections of $\widehat{\mathcal{F}}'(\mathrm{G}_*)$, $\widehat{\mathcal{H}}(\mathrm{G}_*)$ with $\mathcal{Q}(\mathrm{G}_*)$, $\mathcal{Q}_0(\mathrm{G}_*)$ then $\widehat{\mathcal{F}}'_*(\mathrm{G}_*)$, $\widehat{\mathcal{H}}_*(\mathrm{G}_*)$ are non-empty and in fact verify

$$\mathrm{meas}\mathcal{Q}(\mathrm{G}_*) \ge \ \mathrm{meas}\widehat{\mathcal{F}}'_*(\mathrm{G}_*) \ge \left(1 - \frac{\varepsilon^\sigma}{\varepsilon_*^\sigma}\right)\mathrm{meas}\mathcal{Q}(\mathrm{G}_*) \tag{29}$$

$$\mathrm{meas}\mathcal{Q}_0(\mathrm{G}_*) \ge \ \mathrm{meas}\widehat{\mathcal{H}}_*(\mathrm{G}_*) \ge \left(1 - \frac{\alpha_+}{c_0\varepsilon^{12}}\right)\mathrm{meas}\mathcal{Q}_0(\mathrm{G}_*). \tag{30}$$

The proof of Theorem 2.1 relies on some technical result (Propositions 2.3, 2.4 and 2.5) which we now state and prove later.

Proposition 2.3 *Let, for a suitable pure number $\underline{k} \in (1, 2)$, $\Lambda_- < \mathrm{G}$, $k_- \le \underline{k}\, k_+ \ge 2$, $\alpha_+ \le \frac{c^2}{16}$. Choose Λ_+ as the unique value of $\Lambda_2 > \mathrm{G}$ such that $\mathcal{C}$ and the straight line $\Lambda_1 = 2\Lambda_2$ meet at $(\Lambda_1, \Lambda_2) = (2\Lambda_+, \Lambda_+)$. Let*

$$\mathcal{L}_0(\mathrm{G}) := \left\{(\Lambda_1, \Lambda_2) :\quad \mathrm{G} \le \Lambda_2 \le \Lambda_+\,,\quad (\mathrm{G} + \Lambda_2)\sqrt{\frac{4\mathrm{G} + \Lambda_2}{5\mathrm{G}}} < \Lambda_1 < \min\{k_+\,\Lambda_2,\ 2\Lambda_+\}\right\}$$

$$\mathcal{A}_0(\mathrm{G}) := \left\{(\Lambda_1, \Lambda_2, \mathrm{G}_2) :\ (\Lambda_1, \Lambda_2) \in \mathcal{L}_0(\mathrm{G}),\ \mathrm{G}_2 \in \mathcal{G}_p(\Lambda_1, \Lambda_2)\right\}$$

Then, the set

$$\mathcal{N}_0(\mathrm{G}) := \mathcal{A}_0(\mathrm{G}) \times \mathbb{T}^3 \times \mathcal{B}_p(\mathrm{G})$$

is a subset of $\mathcal{N}(\mathrm{G})$.

Proposition 2.4 *There exists* $c_1 \in (0,1)$ *depending only on* Λ_+/G, Λ_-/G *such that, letting, for any* $\gamma < c_1^2\varepsilon^2$,

$$\mathcal{L}_1(\mathrm{G}) := \Big\{(\Lambda_1, \Lambda_2) \in \mathcal{L}\,,\ |\Lambda_1 - \Lambda_2 - \mathrm{G}| < c_1^2\varepsilon^2\Big\}$$

$$\mathcal{G}_1(\Lambda_2) := \Big\{\mathrm{G}_2 :\ \Lambda_2 - c_1^2\varepsilon^2 < \mathrm{G}_2 < \Lambda_2 - \gamma\Big\}$$

$$\mathcal{A}_1(\mathrm{G}) := \Big\{(\Lambda_1, \Lambda_2, \mathrm{G}_2) :\quad (\Lambda_1, \Lambda_2) \in \mathcal{L}_1\,,\quad \mathrm{G}_2 \in \mathcal{G}_1(\Lambda_2)\Big\}$$

$$\mathcal{B}_1(\mathrm{G}, \varepsilon) := \Big\{(\Theta, \vartheta) :\quad \Theta^2 < c_1^2\mathrm{G}\varepsilon^2,\ \vartheta^2 < c_1^2\frac{\varepsilon^2}{\mathrm{G}}\Big\}$$

then the set

$$\mathcal{N}_1(\mathrm{G}, \varepsilon) := \mathcal{A}_1(\mathrm{G}) \times \mathbb{T}^3 \times \mathcal{B}_1(\mathrm{G}, \varepsilon)$$

is a subset of $\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G})$.

Proposition 2.5 *Assume* $\mathrm{G} \geq 10c_1^2\varepsilon^2$ *and* $\alpha_+ < \frac{c^2}{16}$. *Then,* $\mathcal{A}_0(\mathrm{G})$ *and* $\mathcal{A}_1(\mathrm{G})$ *have a non-empty intersection* $\mathcal{A}_\star(\mathrm{G})$, *verifying*

$$\mathrm{meas}(\mathcal{A}_\star(\mathrm{G})) \geq \frac{9}{10}(c_1^2\varepsilon^2 - \gamma)c_1^4\varepsilon^4$$

We prove how Theorem 2.1 follows from the above propositions. $\mathcal{Q}(\mathrm{G}_*)$ is a subset of $\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*)$ and $\mathcal{N}_\varepsilon(\mathrm{G}_*)$, and

$$\mathrm{meas}\,\mathcal{Q}(\mathrm{G}_*) = C_1\varepsilon^8 = C_2\varepsilon^2\,\mathrm{meas}\widehat{\mathcal{M}}'_\varepsilon.$$

The bound in (28) guarantees that

$$\mathrm{meas}\left(\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*) \setminus \widehat{\mathcal{F}}'(\mathrm{G}_*)\right) < C_3\varepsilon^{\frac{1}{2}+\bar{s}}\,\mathrm{meas}\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*) \qquad \forall\, \mathrm{G}_* \in \mathbb{G}_*.$$

On the other hand, if $\widehat{\mathcal{F}}'_\varepsilon(\mathrm{G}_*) \cap \mathcal{Q}(\mathrm{G}_*)$ was empty, we would have

$$\mathrm{meas}\left(\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*) \setminus \widehat{\mathcal{F}}'(\mathrm{G}_*)\right) \geq \mathrm{meas}\mathcal{Q}(\mathrm{G}_*) = C_2\varepsilon^2\,\mathrm{meas}\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*)$$

which contradicts the previous inequality if $\bar{s} > \frac{3}{2}$ and ε is small. Finally, if $\mathrm{G}_* \in \mathbb{G}_*$,

$$\mathrm{meas}\left(\mathcal{Q}_\varepsilon(\mathrm{G}_*) \setminus \widehat{\mathcal{F}}'(\mathrm{G}_*)\right) \leq \mathrm{meas}\left(\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*) \setminus \widehat{\mathcal{F}}'(\mathrm{G}_*)\right) < C_3\varepsilon^{\frac{1}{2}+\bar{s}}\,\mathrm{meas}\widehat{\mathcal{M}}'_\varepsilon(\mathrm{G}_*)$$
$$= C_4\varepsilon^{\bar{s}-\frac{3}{2}}\,\mathrm{meas}\mathcal{Q}_\varepsilon(\mathrm{G}_*)$$

and we have (29) with $\sigma = \bar{s} - \frac{3}{2} = s - \frac{7}{2}$, with $s \geq 4$. The proof of (30) is similar.

2.4 Proof of Propositions 2.3, 2.4 and 2.5

Proof of Proposition 2.3 We only need to prove that $\mathcal{L}_0(\mathrm{G}) \subset \mathcal{L}_p(\mathrm{G})$. We switch to the coordinates

$$y := \frac{\Lambda_1}{\mathrm{G}}, \qquad x := \frac{\Lambda_2}{\mathrm{G}}.$$

We denote as $\mathcal{X}_p := \mathrm{G}^{-1}\mathcal{L}_p$ the domain of (y, x), and as

$$x_- := \frac{\Lambda_-}{\mathrm{G}}, \qquad x_+ := \frac{\Lambda_+}{\mathrm{G}}$$

$\mathcal{X}_p$ can be written as the intersection of the three sets:

$$\begin{aligned}
\mathcal{X}_1 &:= \left\{(y, x) : \quad 1 \le x \le x_+ \,, \quad y > 2, \ \max\{k_- x, (1+x)\sqrt{\frac{4+x}{5}}\} < y < k_+ x\right\} \\
\mathcal{X}_2 &:= \left\{(y, x) : \quad 1 \le x \le x_+ \,, \quad y > 1 + \frac{2}{c}\sqrt{\alpha_+ x}\right\} \\
\mathcal{X}_3 &:= \left\{(y, x) : \quad 1 \le x \le x_+, \quad y > 2\,, \quad 5y^2 - (1 + \frac{2}{c}\sqrt{\alpha_+ y})^2(4 + \frac{2}{c}\sqrt{\alpha_+ y}) > 0\right\}
\end{aligned}$$

We prove $\mathcal{X}_0 := \mathrm{G}^{-1}\mathcal{L}_0$ is a subset of all of them. The curve

$$\mathcal{C} : \qquad y = (1+x)\sqrt{\frac{4+x}{5}} \qquad x \ge 1$$

passes through $P_0 = (1, 2)$. We denote as $\underline{k}$ the slope of the straight line $y = kx$ which is tangent at $\mathcal{C}$ at P_0. The slope of the straight line $y = kx$ through P_0 is obviously $\overline{k} = 2$. We assume that

$$k_- \le \underline{k}, \qquad k_+ \ge \overline{k}$$

and choose (x_+, y_+) as the only (x, y) with $x > 1$ such that $\mathcal{C}$ meets $y = 2x$ at (x, y). Under such assumptions, we have:

$$\mathcal{X}_1 = \left\{(y, x) : \quad 1 \le x \le x_+ \,, \quad (1+x)\sqrt{\frac{4+x}{5}} < y < k_+ x\right\} \supset \mathcal{X}_0$$

The straight line which is tangent at $\mathcal{C}$ at $P_0 = (1, 2)$ has equation

$$y = \frac{6}{5}x + \frac{4}{5}$$

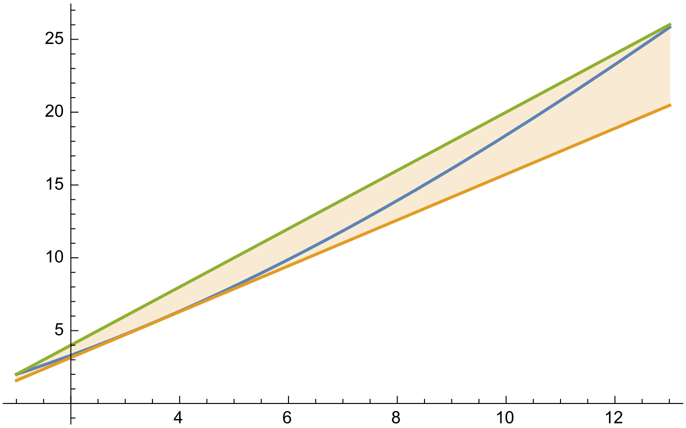

Fig. 1 The blue curve is $\mathcal{C}$; the orange line has slope k_-, the green one has slope k_+ (MATHEMATICA)

Since we $\alpha_+ < \frac{c^2}{4}$, $x > 1$ and $\mathcal{C}$ is convex, we have

$$1 + \frac{2}{c}\sqrt{\alpha_+}x \leq 1 + x \leq \frac{6}{5}x + \frac{4}{5} \leq (1+x)\sqrt{\frac{4+x}{5}}$$

This shows that $\mathcal{X}_2 \supset \mathcal{X}_0$. As for $\mathcal{X}_3$, we note that for

$$\alpha_+ \leq \frac{c^2}{16}$$

it is

$$5y^2 - (1 + \frac{2}{c}\sqrt{\alpha_+}y)^2(4 + \frac{2}{c}\sqrt{\alpha_+}y) \geq 5y^2 - \left(1 + \frac{y}{2}\right)^2\left(4 + \frac{y}{2}\right)$$
$$= \frac{1}{4}(y-2)(y-y_-)(y_+-y).$$

with

$$y_\pm = 13 \pm \sqrt{185}.$$

As $y_- < 0$ and $(y-2)(y_+-y) \geq 0$ on $\mathcal{X}_0$, we have that $\mathcal{X}_3 \supset \mathcal{X}_0$ (Figs. 1, 2). □

Remark 2.1 The numbers $\underline{k}$, Λ_+ of Proposition 2.3 can be chosen as

$$\Lambda_+ = \frac{G}{2}\left(13 + \sqrt{185}\right), \quad \underline{k} = \frac{1}{4}\sqrt{\frac{3}{10}(69 + 11\sqrt{33})} \sim 1.57$$

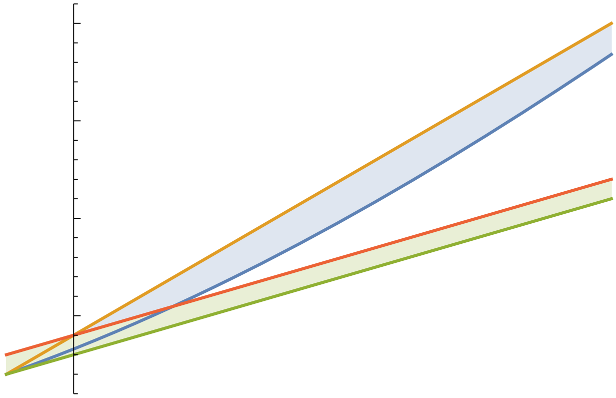

Fig. 2 $\mathcal{L}_0(\mathrm{G})$ (blue) and $\mathcal{L}_1(\mathrm{G})$ (green)

Λ_+ is related to the number x_+ computed along the proof via $\Lambda_+ = x_+\mathrm{G}$. $\underline{k}$ is defined as the slope of the straight line $y = kx$ which is tangent at $\mathcal{C}$. We can compute it eliminating y between the two equations; we obtain the cubic equation

$$x^3 + (6 - 5k^2)x^2 + 9x + 4 = 0. \tag{31}$$

The tangency condition is imposed identifying this equation with

$$(x - a)^2(x - b) = 0 \tag{32}$$

where a is the abscissa of the tangency point. Equating the respective coefficients of (31) and (32), we obtain

$$\begin{cases} -(b + 2a) = 6 - 5k^2 \\ 2ab + a^2 = 9 \\ -a^2 b = 4 \end{cases} \tag{33}$$

Eliminating b through the second and the third equations, we obtain

$$a^3 - 9a - 8 = 0$$

which has the following three roots:

$$a_0 = -1, \qquad a_\pm = \frac{1 \pm \sqrt{33}}{2}.$$

The only admissible value is then

$$a = a_+ = \frac{1 + \sqrt{33}}{2}\ .$$

In correspondence of this value for a, solving the system in (33), we find

$$b = \frac{-17+\sqrt{33}}{32}, \qquad k = \frac{1}{4}\sqrt{\frac{3}{10}(69+11\sqrt{33})} = \underline{k}.$$

□

Proof of Proposition 2.4 From (11), we get

$$|z|^2 = \eta_1^2+\xi_1^2+\eta_2^2+\xi_2^2+p^2+q^2 = 2(\mathrm{G}+\mathrm{G}_2-\mathrm{G}_1)+2(\Lambda_1-\mathrm{G}_1)+2(\Lambda_2-\mathrm{G}_2).$$

From the equality

$$\begin{aligned}\mathrm{G}_1 &= \sqrt{\mathrm{G}^2+\mathrm{G}_2^2-2\Theta^2+2\sqrt{\mathrm{G}^2-\Theta^2}\sqrt{\mathrm{G}_2^2-\Theta^2}\cos\vartheta}\\ &= \mathrm{G}+\mathrm{G}_2+\mathrm{O}\left(\frac{\Theta^2}{\mathrm{G}+\mathrm{G}_2}\right)+\mathrm{O}\left(\frac{\Theta^2\mathrm{G}_2}{\mathrm{G}(\mathrm{G}+\mathrm{G}_2)}\right)+\mathrm{O}\left(\frac{\Theta^2\mathrm{G}}{\mathrm{G}_2(\mathrm{G}+\mathrm{G}_2)}\right)+\mathrm{O}\left(\frac{\vartheta^2\mathrm{G}\mathrm{G}_2}{\mathrm{G}+\mathrm{G}_2}\right)\end{aligned}$$

and the definition of $\mathcal{N}_1(\mathrm{G})$, the assertion trivially follows. □

Proof of Proposition 2.5 Let $\Lambda_2^\star$ be the abscissa, in the plane (Λ_2, Λ_1), of the intersection point between the curves

$$\Lambda_1 = (\mathrm{G}+\Lambda_2)\sqrt{\frac{4\mathrm{G}+\Lambda_2}{5\mathrm{G}}}, \qquad \text{and} \qquad \Lambda_1 = \Lambda_2+\mathrm{G}+c_1^2\varepsilon^2 .$$

Using the coordinate $x := \frac{\Lambda_2}{\mathrm{G}}$. With $x^\star := \frac{\Lambda_2^\star}{\mathrm{G}}$, $\theta := \frac{c_1^2\varepsilon^2}{\mathrm{G}}$, $\zeta := \frac{\gamma}{\mathrm{G}}$, where $\zeta < \theta$, the set $\mathcal{A}_\star(\mathrm{G}) := \mathcal{A}_0(\mathrm{G}) \cap \mathcal{A}_1(\mathrm{G})$ has measure

$$\mathrm{meas}(\mathcal{A}_\star(\mathrm{G})) = \mathrm{G}^3\int_{1+\zeta}^{x^\star} F_1(x)F_2'(x)dx$$

where

$$\begin{aligned}F_1(x) &= \min\left\{2x,\ x+1+\theta\right\}-(1+x)\sqrt{\frac{4+x}{5}}\\ F_2(x) &= \min\left\{\theta-\zeta,\ x-1-\zeta,\ mx-\zeta\right\}\end{aligned}$$

and where, for short, we have let $m := 1-\frac{2}{c}\sqrt{\alpha_+}$. Then,

$$\mathrm{meas}(\mathcal{A}_\star(\mathrm{G})) \geq \mathrm{G}^3\int_{1+\zeta}^{x^\star} F_1(x)F_2(x)dx . \tag{34}$$

To go further, we need a quantitative bound on $x^\star$. Indeed, we have

Claim 2.1 *If* $0 < \theta < \frac{1}{10}$, *then* $1+4\theta < x^\star < 1+6\theta$.

The proof of the claim is postponed below, in order not to interrupt the main proof. Since we have assumed $\mathrm{G} \geq 10c_1^2\varepsilon^2$ and $\alpha_+ \leq \frac{c^2}{16}$, then $\mathrm{G} \geq \frac{\frac{12}{c}\sqrt{\alpha_+}}{1-\frac{2}{c}\sqrt{\alpha_+}}c_1^2\varepsilon^2$. In the new variables, this is $\theta \leq \frac{m}{6(1-m)}$. But then

$$x^* < 1 + 6\theta \leq \frac{1}{1-m} \qquad \Longrightarrow \quad x - 1 - \zeta \leq mx - \zeta \quad \forall\, x < x^*$$

whence

$$F_2(x) = \begin{cases} x - 1 - \zeta \text{ if } 1 + \zeta \leq x \leq 1 + \theta \\ \theta - \zeta \qquad \text{if } 1 + \theta < x \leq x^\star \end{cases}$$

Observe that the second inequality is well put, because $x^\star > 1 + 4\theta$, as said. The function $F_1(x)$ splits in the same intervals:

$$F_1(x) = \begin{cases} 2x - (1+x)\sqrt{\frac{4+x}{5}} & \text{if } 1 + \zeta \leq x \leq 1 + \theta \\ x + 1 + \theta - (1+x)\sqrt{\frac{4+x}{5}} & \text{if } 1 + \theta < x \leq x^\star \end{cases} \tag{35}$$

Since $\zeta < \theta$, a lower bound to the integral in (34) is given by

$$\int_{1+\zeta}^{x^\star} F_1(x)F_2(x)dx \geq \int_{1+\theta}^{x^\star} F_1(x)F_2(x)dx = (\theta - \zeta)\int_{1+\theta}^{x^\star} F(x)dx$$

with

$$F(x) := x + 1 + \theta - (1+x)\sqrt{\frac{x+4}{5}} \tag{36}$$

the function in the second line in (35). Since F is the difference of a linear function and a convex one, it is concave. Then, we have

$$F(x) \geq F(1) + \frac{F(x^\star) - F(1)}{x^\star - 1}(x-1) \qquad \forall\, 1 \leq x \leq x^\star$$

since $F(x^\star) = 0$ and $F(1) = \theta$, this inequality becomes

$$F(x) \geq \frac{x^\star - x}{x^\star - 1}\theta \qquad \forall\, 1 \leq x \leq x^\star$$

hence

$$\int_{1+\theta}^{x^\star} F(x)dx \geq \frac{\theta}{x^\star - 1}\int_{1+\theta}^{x^\star}(x^\star - x)dx = \frac{\theta}{2}\frac{(x^\star - 1 - \theta)^2}{x^\star - 1} \geq \frac{9}{10}\theta^2$$

having used $1 + 4\theta < x^\star < 1 + 6\theta$.

It remains to prove Claim 2.1. $x^\star$ is defined as the zero of the function F in (36) in the range $(1, +\infty)$. Multiplying the left hand side of Equation

$$x + 1 + \theta - (1 + x)\sqrt{\frac{x + 4}{5}} = 0$$

by $x + 1 + \theta + (1 + x)\sqrt{\frac{x+4}{5}}$, we obtain the algebraic equation of degree three

$$x^3 + x^2 + (1 + 10\theta)x - 1 - 10\theta - 5\theta^2 = 0$$

which, for $x \geq -1$ is completely equivalent to the initial equation. We aim to apply a bisection argument to the function at left hand side, which we denote as $G(x)$. We have

$$G(1 + 4\theta) = \theta(64\theta^2 + 19\theta - 4)\ , \qquad G(1 + 6\theta) = \theta(216\theta^2 + 79\theta + 4)$$

and it is immediate to check that

$$G(1 + 4\theta) < 0 \qquad G(1 + 6\theta) > 0 \qquad \forall\, 0 < \theta < \frac{-19 + \sqrt{1385}}{128} = 0.142\ldots$$

To prove uniqueness, just observe that the function $x \in (0, +\infty) \to G(x)$ is increasing for all $\theta > 0$. This completes the proof. □

3 Quantitative KAM Theory

3.1 Proof of Theorem 1.1

The proof of Theorem 1.1 is based on an application of Chierchia and Pinzari (2010, Proposition 3). The method is completely analogous to the one used in the proof of Chierchia and Pinzari (2010, Theorem 1.3), so we shall only say what to change in the proof of Chierchia and Pinzari (2010, Theorem 1.3) in order to obtain the proof of Theorem 1.1. The polynomial $N(I, r)$ in the first non-numbered formula in Chierchia and Pinzari (2010, Section 4) is to be changed as

$$N(I, r) = P_0(I) + \sum_{i=1}^{m} \Omega_i(I) r_i + \frac{1}{2}\sum_{i,j=1}^{m} \beta_{ij}(I) r_i r_j + \Bbbk_{s\geq 3} \sum_{j=3}^{s} \mathcal{P}_j(r; I). \quad (37)$$

Equations (60) and (61) in Chierchia and Pinzari (2010) can be modified, respectively, as

$$\sup_{B_\varepsilon^{2m} \times V_{\rho_0}} |\tilde{P}_{\rm av}| \leq C\varepsilon^{2s+1} \qquad \forall\, 0 < \varepsilon < \varepsilon_0$$

$$\mu < \frac{\varepsilon^{2s+2}}{(\log \varepsilon^{-1})^{2\tau+1}} \qquad \overline{\gamma} > \left(\frac{6(2s+1)}{s_0}\right)^{\tau+\frac{1}{2}} \frac{\sqrt{\mu}(\log \varepsilon^{-1})^{\tau+\frac{1}{2}}}{\varepsilon^{s+\frac{1}{2}}} . \tag{38}$$

Analogously to Chierchia and Pinzari (2010), one next applies Lemma A.1 in Chierchia and Pinzari (2010), but modifying the choice of K as

$$K = \frac{6(2s+1)}{s_0} \log \varepsilon^{-1} \tag{39}$$

and leaving the other quantities unvaried. A bound as in Equation (62) in Chierchia and Pinzari (2010) is so obtained, with H_0 as in Chierchia and Pinzari (2010), $N(\overline{I}, \overline{r})$ as in (37), $\mu \widetilde{P}_{\rm av}(\overline{p}, \overline{q}, \overline{I}) = f_{bnf}(I, \overline{p}, \overline{q}) - N(\overline{I}, \overline{r})$ uniformly bounded by $C\mu\varepsilon^{2s+1}$, by (A_2). Due to the choice of K in (39) and the one for $\overline{\gamma}$ in (38), a bound similar to the one in Equation (63) in Chierchia and Pinzari (2010) holds, with the right hand side replaced by $\overline{C}\mu\varepsilon^{2s+1}$. At this point, one follows the indications in Step 2 of the proof of Theorem 1.3 in Chierchia and Pinzari (2010). Namely, one has to repeat the procedure in Steps 5 and 6 of the proof Theorem 1.4 [previously proved in altchierchiaPi10], with the following modification. The annulus $\mathcal{A}(\varepsilon)$ in Equation (47) in Chierchia and Pinzari (2010) is to be taken as

$$\mathcal{A}(\varepsilon) = \left\{ J \in \mathbb{R}^m : \ \check{c}_1 \varepsilon^{s+\frac{1}{2}} < J_i < \check{c}_2 \varepsilon^2 \, , \quad 1 < i < m \right\}$$

and the number $\check{\rho}$ in

Equation (48) in Chierchia and Pinzari (2010) is to be replaced with $\check{\rho} := \min \{\check{c}_1 \varepsilon^{s+\frac{1}{2}}/2, \ \overline{\rho}/48\}$. The other quantities remain unvaried. In the remaining Steps 5 and 6 of the proof of Theorem 1.4 in Chierchia and Pinzari (2010) replace the number "5" appearing in all the formulae with $(2s+1)$ and $\varepsilon^{n_2/2}$ in Equation (56) (and the formulae below) in Chierchia and Pinzari (2010) with $\varepsilon^{m(s-\frac{3}{2})}$. □

3.2 Proof of Theorem 1.2

The proof of Theorem 1.2 proceeds along the same lines as the proof of Theorem 1.1, apart for being based on a generalization (Theorem 3.1 below) of Chierchia and Pinzari (2010, Proposition 3) which now we state.

As in Chierchia and Pinzari (2010) $\mathcal{D}_{\gamma_1,\gamma_2,\tau} \subset \mathbb{R}^n$ denotes the set of vectors $\omega = (\omega_1, \omega_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ satisfying for any $k = (k_1, k_2) \in \mathbb{Z}^{n_1} \times \mathbb{Z}^{n_2} \setminus \{0\}$, inequality

$$|\omega_1 \cdot k_1 + \omega_2 \cdot k_2| \geq \begin{cases} \dfrac{\gamma_1}{|k|^\tau} & \text{if} \quad k_1 \neq 0 \, ; \\ \\ \dfrac{\gamma_2}{|k_2|^\tau} & \text{if} \quad k_1 = 0 \, , \quad k_2 \neq 0 \, . \end{cases} \tag{40}$$

Theorem 3.1 *Let* $n_1, n_2 \in \mathbb{N}$, $n := n_1 + n_2$, $\tau > n$, $\gamma_1 \geq \gamma_2 > 0$, $0 < s \leq \frac{\varepsilon}{\overline{\varepsilon}+\varepsilon}$, $\rho > 0$, $A := D_\rho \times B^2_{\overline{\varepsilon}+\varepsilon}$, *and let*

$$\mathrm{H}(I, \psi, p, q) = \mathrm{h}(I, pq) + \mathrm{f}(I, \psi, p, q)$$

be real-analytic on $A \times \mathbb{T}^n_{\overline{s}+s}$. *Let*

$$I = (I_1, I_1),\ \ \varpi(I, pq) := \partial_{(I,pq)}\mathrm{h}(I, pq) = (\omega_1(I_1, I_2, pq), \omega_2(I_1, I_2, pq), \nu(I_1, I_2, pq))$$

with $\omega_k(I_1, I_2, pq) := \partial_{I_k}\mathrm{h}(I_1, I_2, pq)$, *and assume that the map* $I \in D_\rho \to \omega(I, J)$ *is a diffeomorphism of* D_ρ *for all* $J = pq$, *with* $(p, q) \in B^2_\varepsilon$, *with non-singular Hessian matrix* $U(I, J) := \partial^2_I \mathrm{h}(I, J)$. *Let*[18]

$$M \geq \|\partial\omega\|_A\ ,\ \ \widehat{M} \geq \|\partial\omega_1\|_A,\ \ \overline{M} \geq \|U^{-1}\|_A\ ,\ \ E \geq \|\mathrm{f}\|_{\rho,\overline{s}+s},\quad \lambda \leq \inf|\mathrm{Re}\,\nu|_A\ .$$

Assume, for[19] *simplicity,*

$$2\frac{s^\tau\gamma_2}{6^\tau\lambda} \leq 1. \tag{41}$$

Define

$$\begin{aligned}
&\widehat{c} := 2^7(n+1)(24)^\tau\ ,\quad \widetilde{c} := 2^6\\
&K := \frac{32}{s}\ \log_+\left(\frac{EM^2L}{\gamma_1^2}\right)^{-1} \quad \text{where} \quad \log_+ a := \max\{1, \log a\}\\
&\widehat{\rho} := \min\left\{\frac{\gamma_1}{2MK^{\tau+1}}\ ,\ \frac{\gamma_2}{2\widehat{M}K^{\tau+1}},\ \rho\right\},\quad \widetilde{\rho} := \min\left\{\widehat{\rho},\ \frac{\varepsilon^2}{s}\right\}\\
&L := \max\left\{\overline{M},\ M^{-1},\ \widehat{M}^{-1}\right\}\\
&\widehat{E} := \frac{EL}{\widehat{\rho}\widetilde{\rho}}\ ,\qquad \widetilde{E} := \frac{E}{\lambda\varepsilon^2}.
\end{aligned}$$

Finally, let $\overline{M}_1$, $\overline{M}_2$ *upper bounds on the norms of the sub-matrices* $n_1 \times n$, $n_2 \times n$ *of* U^{-1} *of the first* n_1, *last* n_2 *rows*[20]. *Assume the perturbation* f *so small that the following "KAM conditions" hold*

$$\widehat{c}\widehat{E} < 1\ ,\quad \widetilde{c}\widetilde{E} < 1 \tag{42}$$

[18] The norms will be specified in the next Sect. 3.3.

[19] (41) is a simplifying assumption. It may be relaxed.

[20] *That is*, $\overline{M}_i \geq \sup_{D_\rho}\|T_i\|\ ,\quad i = 1,\ 2\ ,\quad \text{if}\quad U^{-1} = \begin{pmatrix} T_1\\ T_2\end{pmatrix}$.

Then, for any $(\pi, \kappa) \in B^2_{\bar\varepsilon}$ *and any* $\omega_* \in \Omega_*(\pi\kappa) := \omega(D, \pi\kappa) \cap \mathcal{D}_{\gamma_1,\gamma_2,\tau}$, *one can find a unique real-analytic embedding*

$$\begin{aligned}\phi_{\omega_*} : \ \mathbb{T}^n \times \{(\pi,\kappa)\} \to \ & \mathrm{Re}\,(D_r) \times \mathbb{T}^n \times B^2_{\bar\varepsilon + r'} \\ (\vartheta, \pi, \kappa) \to \ & \Big(v(\vartheta,\pi,\kappa;\omega_*), \vartheta + u(\vartheta,\pi,\kappa;\omega_*), \pi + w(\vartheta,\pi,\kappa;\omega_*), \kappa + y(\vartheta,\pi,\kappa;\omega_*)\Big)\end{aligned} \tag{43}$$

such that $\mathcal{M}_{\omega_*} := \phi_{\omega_*}(\mathbb{T}^n \times B^2_{\bar\varepsilon})$ *is a real-analytic* $(n+2)$*-dimensional manifold, on which the* H*-flow is analytically conjugated to*

$$(\vartheta, \pi, \kappa) \in \mathbb{T}^n \times B^2_{\bar\varepsilon} \to (\vartheta + \omega_* t,\ \pi \to \pi e^{-\nu_*(\omega_*, \pi\kappa)t},\ \kappa \to \kappa e^{\nu_*(\omega_*, \pi\kappa)t}). \tag{44}$$

In particular, the manifolds

$$\mathrm{T}_{\omega_*} := \phi_{\omega_*}\left(\mathbb{T}^n \times \{(0,0)\}\right)$$

are real-analytic n-dimensional H*-invariant tori embedded in* $\mathrm{Re}\,(D_r) \times \mathbb{T}^n \times B^2_{\bar\varepsilon}$, *equipped with* $(n+1)$*-dimensional manifolds*

$$\mathcal{M}_{\mathrm{u}} := \phi_{\omega_*}\left(\mathbb{T}^n \times \{0\} \times B^1_{\bar\varepsilon}\right), \qquad \mathcal{M}_{\mathrm{s}} := \phi_{\omega_*}\left(\mathbb{T}^n \times B^1_{\bar\varepsilon} \times \{0\}\right)$$

on which the motions leave, approach T_{ω_*} *at an exponential rate. Let* $\mathrm{T}_{\omega_*,0}$ *denote the projection of* T_{ω_*} *on the* (I, φ)*-variables, and* $\mathrm{K}_0 := \bigcup_{\omega_* \in \Omega_*} \mathrm{T}_{\omega_*,0}$. *Then* K_0 *satisfies the following measure*[21] *estimate:*

$$\mathrm{meas}_{2n}(\mathrm{Re}\,(D_r) \times \mathbb{T}^n \setminus \mathrm{K}_0) \le c_n\Big(\mathrm{meas}(D \setminus D_{\gamma_1,\gamma_2,\tau} \times \mathbb{T}^n) + \mathrm{meas}(\mathrm{Re}\,(D_r) \setminus D) \times \mathbb{T}^n\Big), \tag{45}$$

where $D_{\gamma_1,\gamma_2,\tau}$ *denotes the* $\omega_0(\cdot, 0)$*-preimage of* $\mathcal{D}_{\gamma_1,\gamma_2,\tau}$ *and* c_n *can be taken to be* $c_n = (1 + (1 + 2^8 n E)^{2n})^2$.

Finally, the following uniform estimates hold for the embedding ϕ_{ω_*}:

$$\begin{aligned}&|v_1(\vartheta,\pi,\kappa;\omega_*) - I^0_1(\pi\kappa;\omega_*)| \le 6n\left(\frac{\overline{M}_1}{\overline{\overline{M}}} + \frac{\widehat{M}}{M}\right)\widehat{E}\,\widetilde{\rho}\\ &|v_2(\vartheta,\pi,\kappa;\omega_*) - I^0_2(\pi\kappa;\omega_*)| \le 6n\left(\frac{\overline{M}_2}{\overline{\overline{M}}} + \frac{\widehat{M}}{M}\right)\widehat{E}\,\widetilde{\rho}\,,\\ &|u(\vartheta,\pi,\kappa;\omega_*)| \le 2\,\widehat{E}\,s, \quad |w(\vartheta,\pi,\kappa;\omega_*)| \le 2\,\widehat{E}\,\varepsilon\\ &|y(\vartheta,\pi,\kappa;\omega_*)| \le 2\,\widehat{E}\,\varepsilon\end{aligned} \tag{46}$$

[21] meas_n denotes the n-dimensional Lebesgue measure.

where $v(\vartheta, \pi, \kappa; \omega_*) = (v_1(\vartheta, \pi, \kappa; \omega_*), v_2(\vartheta, \pi, \kappa; \omega_*))$ *and* $I^0(\pi\kappa; \omega_*) = (I_1^0(\pi\kappa; \omega_*), I_2^0(\pi\kappa; \omega_*)) \in D$ *is the* $\omega(\cdot, \pi\kappa)$*—pre-image of* $\omega_* \in \Omega_*(\pi\kappa)$. *where* $r := 8n\widehat{E}\widetilde{\rho}$, $r' = 2\widehat{E}\varepsilon$

The proof of Theorem 3.1 is deferred to the next Sect. 3.3. Here, we prove how Theorem 1.2 follows from it.
As said, we follow the same ideas of the proof of Theorem 3.1, which in turn follows (Chierchia and Pinzari 2010, Theorem 1.3). By (A_2'),

$$P_{\rm av}(I, p, q) = P_0(I, pq) + P_1(I, p, q) \quad \text{where} \quad |P_1| \le a\|P_0\| =: \epsilon. \tag{47}$$

At this point, proceeding as in Chierchia and Pinzari (2010, Proof of Theorem 1.3, Step 1) but with ϵ^5 replaced by ϵ, under condition

$$\mu < \frac{\epsilon^{1+\eta}}{(\log(\epsilon^{-1}))^{2\tau+1}}\,, \qquad \overline{\gamma} \ge C\Big(\frac{6}{s_0}\Big)^{\tau+\frac12} \frac{\sqrt{\mu}(\log\epsilon^{-1})^{\tau+\frac12}}{\sqrt{\epsilon}}\,,$$

by an application of Chierchia and Pinzari (2010, Lemma A.1), with $\overline{K} = \frac{6}{s_0}\log\epsilon^{-1}$, $r_p = r_q = \epsilon_0$, $r = 4\rho = \overline{\rho} := \min\left\{\frac{\overline{\gamma}}{2\overline{M}\,\overline{K}^{\tau+1}},\ \rho_0\right\}$ (with $\overline{M} := \sup|\partial^2_{I_1} H_0|$), $\rho_p = \rho_q = \epsilon_0/4$, $\sigma = s_0/4$, $\ell_1 = n_1$, $\ell_2 = 0$, $m = n_2$ $h = H_0$, $g \equiv 0$, $f = \mu P$, $A = \overline{D} := \omega_0^{-1}\mathcal{D}_{\gamma,\tau}$ (where ω_0 is as in A_1 and $\mathcal{D}_{\gamma,\tau}$ is the usual Diophantine set in $\mathbb{R}^n$, namely the set (40) with $\gamma_1 = \gamma_2$), $B = B' = \{0\}$, $s = s_0$, $\alpha_1 = \alpha_2 = \overline{\alpha} = \frac{\overline{\gamma}}{2\overline{K}^{\tau}}$, and $\Lambda = \{0\}$, on the domain $W_{\overline{v},\overline{s}}$ where $\overline{v} = (\overline{\rho}/2, \epsilon_0/2)$ and $\overline{s} = s_0/2$, one finds a real-analytic and symplectic transformation $\overline{\phi}$ which carries H to

$$\begin{aligned}\overline{H}(\overline{I}, \overline{\varphi}, \overline{p}, \overline{q}) &:= H \circ \overline{\phi}(\overline{I}, \overline{\varphi}, \overline{p}, \overline{q})\\ &= H_0(\overline{I}) + \mu P_0(\overline{I}, \overline{pq}) + \mu P_1(\overline{I}, \overline{\varphi}, \overline{p}, \overline{q}) + \widetilde{P}(\overline{I}, \overline{\varphi}, \overline{p}, \overline{q})\\ &= H_0(\overline{I}) + \mu P_0(\overline{I}, \overline{pq}) + \mu\overline{P}(\overline{I}, \overline{\varphi}, \overline{p}, \overline{q})\end{aligned}$$

where

$$\|\widetilde{P}\|_{\overline{v},\overline{s}} \le \overline{C}\mu\max\{\frac{\mu\overline{K}^{2\tau+1}}{\overline{\gamma}^2}, \frac{\mu\overline{K}^{\tau}}{\overline{\gamma}}\; e^{-\overline{K}s_0/2}\} \le \overline{C}\mu\epsilon = \overline{C}\mu a\|P_0\|\,,$$

whence (by (47)) also $\overline{P} = \mu\widetilde{P}_{\rm av} + \widetilde{P}$ is bounded by $C\mu a\|P_0\|$ on $W_{\overline{v},\overline{s}}$.
The next step is to apply Theorem 3.1 to the Hamiltonian $\overline{H}$. Since we can take

$$\begin{aligned}&M = C\,, \quad \widehat{M} = C\mu\|P_0\|\,, \quad \overline{M} = C(\mu\|P_0\|)^{-1}\,, \quad E = C\mu a\|P_0\|\\ &\bar{M}_1 = C\,, \quad \bar{M}_2 = C(\mu\|P_0\|)^{-1}, \quad \lambda = C^{-1}\mu\|P_0\|\end{aligned}$$

the numbers L, K, $\widehat{\rho}$ and $\widetilde{\rho}$ can be bounded, respectively, as

$$L \le C(\mu\|P_0\|)^{-1}\,, \qquad K \le C\log\,(a/\gamma_1{}^2)^{-1}$$

and

$$\widehat{\rho} \geq c\ \min\left\{\frac{\gamma_1}{(\log\,(a/{\gamma_1}^2)^{-1})^{\tau+1}}\ ,\ \frac{\overline{\gamma}_2}{(\log\,(a/{\gamma_1}^2)^{-1})^{\tau+1}}\ ,\ \frac{\bar{\gamma}}{(\log\epsilon^{-1})^{\bar{\tau}+1}}\ ,\ \rho_0\right\}$$
$$\widetilde{\rho} \geq c\ \min\left\{\frac{\gamma_1}{(\log\,(a/{\gamma_1}^2)^{-1})^{\tau+1}}\ ,\ \frac{\overline{\gamma}_2}{(\log\,(a/{\gamma_1}^2)^{-1})^{\tau+1}}\ ,\ \frac{\bar{\gamma}}{(\log\epsilon^{-1})^{\bar{\tau}+1}}\ ,\ \rho_0,\ \varepsilon^2\right\}$$

having let $\gamma_2 := \mu\|P_0\|\overline{\gamma}_2$. Condition (41) is trivially satisfied for any $\overline{\gamma} < 1$, $s \leq 6$, while, from the bounds

$$\widehat{c}\widehat{E} \leq Ca\max\left\{\frac{(\log\,(a/{\gamma_1}^2)^{-1})^{2(\tau+1)}}{\gamma_1^2}v\ \frac{(\log\,(a/{\gamma_1}^2)^{-1})^{2(\tau+1)}}{\overline{\gamma}_2^2},\ \frac{(\log\epsilon)^{-1})^{2(\tau+1)}}{\overline{\gamma}^2},\ \frac{1}{\rho_0^2},\ \frac{1}{\varepsilon^4}\right\}\ ,$$
$$\widetilde{c}\widetilde{E} \leq C\frac{a}{\varepsilon^2}$$

one sees that conditions (42) hold taking

$$\overline{\gamma} = \gamma_1 = \overline{\gamma}_2 = \widehat{C}\sqrt{a},\quad a < \widehat{C}^{-1}\varepsilon^4$$

with a suitable $\widehat{C} > 1$. By the thesis of Theorem 3.1, we can find a set of n-dimensional invariant tori $\mathcal{K} \subset \mathcal{P}$ whose projection $\mathcal{K}_0$ on $\mathcal{P}_0$ satisfies the measure estimate

$$\text{meas}\mathcal{P}_0 \geq\ \text{meas}\mathcal{K}_0 \geq (1 - C'(\overline{\gamma} + \gamma_1 + \gamma_2))\,\text{meas}\mathcal{P}_0 \geq (1 - C\sqrt{a})\,\text{meas}\mathcal{P}_0\ .$$

□

3.3 Proof of Theorem 3.1

We fix the following notations.

- in $\mathbb{R}^n$ we fix the 1-norm: $|I| := |I|_1 := \sum_{1\leq i\leq n_1}|I_i|$;
- in $\mathbb{T}^n$ we fix the "sup-metric": $|\varphi| := |\varphi|_\infty := \max_{1\leq i\leq n}|\varphi_i|$ (mod 2π);
- in $\mathbb{R}$ we fix the sup norm: $|(p,q)| := |(p,q)|_\infty := \max\{|p|,|q|\}$;
- for matrices we use the "sup-norm": $|\beta| := |\beta|_\infty := \max_{i,j}|\beta_{ij}|$;
- we denote as $B^n_\varepsilon(z_0)$ the complex ball having radius ε centered at $z_0 \in \mathbb{C}^n$. If $z_0 = 0$, we simply write B^n_ε.
- if $A \subset \mathbb{R}^n$, and $r > 0$, we denote by $A_r := \bigcup_{x_0\in A} B^n_r(x_0)$ the complex r-neighborhood of A (according to the prefixed norms/metrics above);
- given $A \subset \mathbb{R}^n$ and positive numbers r, ε, s, we let

$$v := (r,\varepsilon)\ ,\quad U_v := A_r \times B^2_\varepsilon\ ,\quad W_{v,s} := U_v \times \mathbb{T}^n_s$$

- if f is real-analytic on a complex domain of the form W_{v_0,s_0}, with $v_0 = (r_0, \varepsilon_0)$, $r_0 > r$, $\varepsilon_0 > \varepsilon$, $s_0 > s$, we denote by $\|f\|_{v,s}$ its "sup-Taylor–Fourier norm":

$$\|f\|_{v,s} := \sum_{k,\alpha,\beta}\sup_{U_v}|f_{\alpha,\beta,k}|e^{|k|s}\varepsilon^{|(\alpha,\beta)|} \tag{48}$$

with $|k|:=|k|_1$, $|(\alpha,\beta)|:=|\alpha|_1+|\beta|_1$, where $f_{k,\alpha,\beta}(I)$ denotes the coefficients in the expansion

$$f=\sum_{\substack{(k,\alpha,\beta)\in\mathbf{Z}^n\times\mathbb{N}^\ell\times\mathbb{N}^\ell\\ \alpha_i\neq\beta_i\forall i}} f_{k,\alpha,\beta}(I)e^{ik\cdot\varphi}p^\alpha q^\beta;$$

- if f is as in the previous item, $K>0$ and $\mathbb{L}$ is a sub-lattice of $\mathbf{Z}^n$, $T_K f$ and $\P_{\mathbb{L}} f$ denote, respectively, the K-truncation and the $\mathbb{L}$-projection of f:

$$T_K f := \sum_{\substack{(k,\alpha,\beta)\in\mathbf{Z}^n\times\mathbb{N}^\ell\times\mathbb{N}^\ell\\ \alpha_i\neq\beta_i\forall i\,|k|_1\leq K}} f_{k,\alpha,\beta}(I)e^{ik\cdot\varphi}p^\alpha q^\beta\,,\quad \P_{\mathbb{L}} f := \sum_{\substack{(k,\alpha,\beta)\in\mathbf{Z}^n\times\mathbb{N}^\ell\times\mathbb{N}^\ell\\ \alpha_i\neq\beta_i\forall i,k\in\mathbb{L}}} f_{k,\alpha,\beta}(I)e^{ik\cdot\varphi}p^\alpha q^\beta$$

with $f_{k,\alpha,\beta}(I):=f_{k,\alpha,\beta}(I,0,0)$. We say that f *is* $(K,\mathbb{L})$ *in normal form* if $f=\P_{\mathbb{L}}T_K f$. If $\mathbb{L}$ is strictly larger than $\{0\}$, we say that f *is resonant normal form.*

Proposition 3.1 (Partially hyperbolic averaging theory) *Let* $H=h(I_1,I_2,pq)+f(I,\varphi,p,q)$ *be a real-analytic function on* W_{v_0,s_0}*, with* $v_0=(r_0,\varepsilon_0)$*. Let* K*,* r*,* s*,* ε*,* $\hat r$*,* $\hat s$*, positive numbers, with* $\hat r<r/4$*,* $\hat s<s/4$ *and* $\hat\varepsilon<\varepsilon/4$*. Put* $\hat\sigma:=\min\left\{\hat s,\frac{\hat\varepsilon}{\varepsilon}\right\}$*. Assume there exist positive numbers* α_1*,* $\alpha_2>0$*, with* $\alpha_1\geq\alpha_2$*, such that, for all* $k=(k_1,k_2,k_3)\in\mathbb{Z}^{n+1}$*,* $0<|k|\leq K$ *and for all* $(I,p,q)\in U_{r,\varepsilon}$*,*

$$|\omega_1\cdot k_1+\omega_2\cdot k_2-\mathrm{i}k_3\nu|\geq\begin{cases}\alpha_1 \text{ if } k_1\neq 0\\ \alpha_2 \text{ if } k_1=0,\ (k_2,k_3)\neq(0,0)\end{cases}\tag{49}$$

and

$$K\widehat\sigma\geq 8\log 2,\qquad \frac{2^3c_1K\widehat\sigma}{\alpha_2\delta}\|f\|_{r,s,\varepsilon}<1,\qquad \delta:=\min\{\hat r\hat s,\ \hat\varepsilon^2\}\tag{50}$$

with a suitable number c_1*. Then, one can find a real-analytic and symplectic transformation*

$$\Phi_*:\quad W_{r_*,s_*,\varepsilon_*}\to W_{r,s,\varepsilon}$$

with $r_*=r-4\hat r$*,* $s_*=s-4\hat s$*,* $\varepsilon_*=\varepsilon-4\hat\varepsilon$*, which conjugates* H *to*

$$H_*(I,\varphi,p,q):=H\circ\Phi_*=\mathrm{h}(I,pq)+g(I,\varphi,p,q)+f_*(I,\varphi,p,q),$$

where g *is* $(K,\{0\})$ *in normal form, and* g*,* f *verify*

$$\|g-\P_0T_K f\|_{r_*,s_*,\varepsilon_*}\leq\frac{8c_1\,\|f\|^2_{r,s,\varepsilon}}{\alpha_2\delta}$$

$$\|f_*\|_{r_*,s_*,\varepsilon_*} \le e^{-K\hat{\sigma}/4}\|f\|_{r_*,s_*,\varepsilon_*} \tag{51}$$

Finally, Φ_* verifies

$$\max\left\{\alpha_1\hat{s}|I_1 - I_1'|,\ \alpha_2\hat{s}|I_2 - I_2'|,\ \alpha_2\hat{r}\,|\varphi - \varphi'|,\ \alpha_2\hat{\varepsilon}\,|p - p'|,\ \alpha_2\hat{\varepsilon},\ |q - q'|\right\} \le 2c_1E. \tag{52}$$

Proposition 3.1 is an extension of the Normal Form Lemma by Pöschel (1993). The extension pertains at introducing the (p, q) coordinates in the integrable part and leaving the amounts of analyticity $\hat{r}$, $\hat{s}$ and $\hat{\varepsilon}$ as independent. This is needed in order to construct the motions (44), where the coordinates (π, κ) are not set to $(0, 0)$, but take value in a small neighborhood of it. A more complete statement implying Proposition 3.1 is quoted and proved in Sect. 3.4.
Below, we let $B := B^2_{\bar{\varepsilon}}(0)$; therefore, B_ε will stand for $B^2_{\bar{\varepsilon}+\varepsilon}(0)$.

Lemma 3.1 *(KAM Step Lemma) Under the same assumptions and notations as in Theorem 3.1, there exists a sequence of numbers ρ_j, ε_j, s_j; of domains*

$$(W_j)_{\rho_j,\varepsilon_j,s_j} = (A_j)_{\rho_j,\varepsilon_j} \times \mathbb{T}^n_{\bar{s}+s_j}, \quad \text{with} \quad (A_j)_{\rho_j,\varepsilon_j} := \bigcup_{(p_j,q_j)\in B_{\varepsilon_j}} \big(D_j(p_jq_j)\big)_{\rho_j} \times \{(p_j, q_j)\}$$

and a real-analytic and symplectic transformations

$$\Psi_{j+1} :\ (I_{j+1}, \varphi_{j+1}, p_{j+1}, q_{j+1}) \in (W_{j+1})_{\rho_{j+1},\varepsilon_{j+1},s_{j+1}} \ \to\ (I_j, \varphi_j, p_j, q_j) \in (W_j)_{\rho_j,\varepsilon_j,s_j} \tag{53}$$

such that

$$\begin{aligned}\mathrm{H}_{j+1}(I_{j+1}, \varphi_{j+1}, p_{j+1}, q_{j+1}) &= \mathrm{H}_j \circ \Psi_{j+1}(I_{j+1}, \varphi_{j+1}, p_{j+1}, q_{j+1})\\ &= \mathrm{h}_{j+1}(I_{j+1}, p_{j+1}q_{j+1})\\ &\quad + \mathrm{f}_{j+1}(I_{j+1}, \varphi_{j+1}, p_{j+1}, q_{j+1})\end{aligned}$$

and such that the following holds. Letting $E_0 := E$, $(M_0, \overline{M}_0, \widehat{M}_0, L_0) = (M, \overline{M}, \widehat{M}, L)$, $s_0 := s$, $\rho_0 := \rho$, $\varepsilon_0 := \varepsilon_0$, $\lambda_0 := \lambda$ and, given, for $0 \le j \in \mathbb{Z}$, E_j, $(M_j, \overline{M}_j, \widehat{M}_j, L_j)$, s_j, ρ_j, ε_j, λ_j, define

$$K_j := \frac{32}{s_j}\log_+\left(\frac{E_jL_jM_j^2}{{\gamma_1}^2}\right)^{-1} \tag{54}$$

$$\widehat{\rho}_j := \min\left\{\frac{\gamma_1}{2M_jK_j^{\tau+1}},\ \frac{\gamma_2}{2\widehat{M}_jK_j^{\tau+1}},\ \frac{\lambda_j}{2M_jK_j},\ \frac{\lambda_j}{2\widehat{M}_jK_j},\ \rho_j\right\}, \tag{55}$$

$$\widetilde{\rho}_j := \min\left\{\widehat{\rho}_j,\ \frac{\varepsilon_j^2}{s_j}\right\}, \quad \widehat{E}_j := \frac{E_jL_j}{\widehat{\rho}_j\widetilde{\rho}_j}$$

$$
\begin{aligned}
E_{j+1} &:= \frac{E_j L_j M_j^2}{{\gamma_1}^2}, \quad (M_{j+1}, \overline{M}_{j+1}, \widehat{M}_{j+1}, L_{j+1}) = 2(M_j, \overline{M}_j, \widehat{M}_j, L_j) \\
\rho_{j+1} &:= \frac{\widehat{\rho}_j}{4}, \ \varepsilon_{j+1} := \frac{\varepsilon_j}{4}, \ , \lambda_{j+1} := \lambda_j - 2^8 \frac{E_j}{\varepsilon_j^2}, \quad s_{j+1} := \frac{s_j}{4}.
\end{aligned} \tag{56}
$$

Then, for all $(p_{j+1}, q_{j+1}) \in B_{\varepsilon_{j+1}}$,

(i) $D_{j+1}(p_{j+1}q_{j+1}) \subseteq (D_j(p_{j+1}q_{j+1}))_{\widehat{\rho}_j/4}$. *Letting*

$$
\begin{aligned}
\varpi_{j+1} &:= \partial_{(I_{j+1}, p_{j+1}q_{j+1})} \mathrm{h}_{j+1}(I_{j+1}, p_{j+1}q_{j+1})) \\
&= (\omega_{j+1}(I_{j+1}, p_{j+1}q_{j+1}), \nu_{j+1}(I_{j+1}, p_{j+1}q_{j+1}))
\end{aligned}
$$

the map $I_{j+1} \to \omega_{j+1}(I_{j+1}, p_{j+1}q_{j+1})$ *is a diffeomorphism of* $\big(D_{j+1}(p_{j+1}q_{j+1})\big)_{\rho_j}$ *verifying*

$$
\omega_{j+1}(D_{j+1}(p_{j+1}q_{j+1})), p_{j+1}q_{j+1}) = \omega_j(D_j(p_{j+1}q_{j+1})), p_{j+1}q_{j+1}).
$$

The map

$$
\begin{aligned}
\widehat{\iota}_{j+1}(p_{j+1}q_{j+1}) &= (\widehat{\iota}_{j+1,1}(p_{j+1}q_{j+1}), \widehat{\iota}_{j+1,2}(p_{j+1}q_{j+1})) : \\
D_j(p_{j+1}q_{j+1}) &\to D_{j+1}(p_{j+1}q_{j+1}) \\
I_j(p_{j+1}q_{j+1}) &\to I_{j+1}(p_{j+1}q_{j+1}) := \omega_{j+1}^{-1}\big(\omega_j(I_j, p_{j+1}q_{j+1}), p_{j+1}q_{j+1}\big)
\end{aligned}
$$

verifies

$$
\sup_{D_j} |\widehat{\iota}_{j+1,1}(p_{j+1}q_{j+1}) - \mathrm{id}\,| \le 3n \frac{\overline{M}_1}{\overline{\overline{M}}} \widehat{E}_j \widetilde{\rho}_j \le 3n \widehat{E}_j \widetilde{\rho}_j \ ,
$$

$$
\sup_{D_j} |\widehat{\iota}_{j+1,2}(p_{j+1}q_{j+1}) - \mathrm{id}\,| \le 3n \frac{\overline{M}_2}{\overline{\overline{M}}} \widehat{E}_j \widetilde{\rho}_j \le 3n \widehat{E}_j \widetilde{\rho}_j \tag{57}
$$

$$
\mathcal{L}(\widehat{\iota}_{j+1}(p_{j+1}q_{j+1}) - \mathrm{id}\,) \le 2^9 n \widehat{E}_j \tag{58}
$$

(ii) the perturbation f_j *has sup-Fourier norm*

$$
\| f_j \|_{(W_j)_{\rho_j, \varepsilon_j, s_j}} \le E_j
$$

(iii) the real-analytic symplectomorphisms Ψ_{j+1} *in* (53) *verify*

$$
\sup_{(W_{j+1})_{\rho_{j+1}, \varepsilon_{j+1}, s_{j+1}}} |I_{j,1}(I_{j+1}, \varphi_{j+1}, p_{j+1}, q_{j+1}) - I_{j+1,1}| \le \frac{3}{4} \frac{\widehat{M}_j}{M_j} \widehat{E}_j \widetilde{\rho}_j
$$

$$
\sup_{(W_{j+1})_{\rho_{j+1}, \varepsilon_{j+1}, s_{j+1}}} |I_{j,2}(I_{j+1}, \varphi_{j+1}, p_{j+1}, q_{j+1}) - I_{j+1,2}| \le \frac{3}{4} \widehat{E}_j \widetilde{\rho}_j
$$

$$\sup_{(W_{j+1})_{\rho_{j+1},\varepsilon_{j+1},s_{j+1}}} |\varphi_j(I_{j+1},\varphi_{j+1},p_{j+1},q_{j+1})-\varphi_{j+1}| \le \frac{3}{4}\widehat{E}_j s_j$$
$$\sup_{(W_{j+1})_{\rho_{j+1},\varepsilon_{j+1},s_{j+1}}} |p_j(I_{j+1},\varphi_{j+1},p_{j+1},q_{j+1})-p_{j+1}| \le \frac{3}{4}\widehat{E}_j \varepsilon_j$$
$$\sup_{(W_{j+1})_{\rho_{j+1},\varepsilon_{j+1},s_{j+1}}} |q_j(I_{j+1},\varphi_{j+1},p_{j+1},q_{j+1})-q_{j+1}| \le \frac{3}{4}\widehat{E}_j \varepsilon_j. \tag{59}$$

The rescaled dimensionless map $\check{\Phi}_{j+1} := \mathrm{id} + 1_{\widehat{\rho}_0^{-1},s_0^{-1},\varepsilon_0^{-1}}\big(\Phi_{j+1}-\mathrm{id}\big)\circ 1_{\widehat{\rho}_0,s_0,\varepsilon_0}$ *has Lipschitz constant on* $(W_{j+1})_{\rho_{j+1}/\widehat{\rho}_0,\varepsilon_{j+1}/\varepsilon_0,s_{j+1}/s_0}$

$$\mathcal{L}(\check{\Phi}_{j+1}-\mathrm{id}) \le 6(n+1)\big(12\cdot(24)^\tau\big)^j\widehat{E}_j\,; \tag{60}$$

(iv) for any $j\ge 0$, $\widehat{E}_{j+1} < \widehat{E}_j^2$, $\lambda_j \ge \frac{\lambda_0}{2}$.

Proof The proof of this proposition is obtained generalizing (Chierchia and Pinzari 2010, Lemma B.1). We shall limit ourselves to describe only the different points, leaving to the interested reader the easy work of completing details.
We construct the transformations (53) by recursion, based on Proposition 3.1. For simplicity of notations, we shall systematically eliminate the sub-fix "j" and replace "$j+1$" with a "+". As an example, instead of (53), we shall write

$$\Psi_+ : W_+ \to W.$$

When needed, the base step will be labeled as "0" (e.g., (76) below). Let us assume (inductively) that

$$\omega(D,pq)\subset\mathcal{D}_{\gamma_1,\gamma_2,\tau}\quad\forall\,(p,q)\in B_\varepsilon \tag{61}$$
$$\widehat{c}\widehat{E} < 1 \tag{62}$$
$$\lambda \ge \max\left\{\frac{\gamma_2}{K^\tau},\ \frac{\lambda_0}{2}\right\}. \tag{63}$$

Condition (61) is verified at the base step provided one takes $D_0 = \omega_0^{-1}(\mathcal{D}_{\gamma_1,\gamma_2,\tau}, p_0q_0)$; (62) is so by assumption, while (63) follows from (41):

$$\lambda_0 \ge \frac{\lambda_0}{2} \ge \frac{s_0^\tau\gamma_2}{6^\tau} \ge \frac{\gamma_2}{K_0^\tau}. \tag{64}$$

We aim to apply Proposition 3.1 with ε, s of Proposition 3.1 corresponding now to $\overline{\varepsilon}+\varepsilon$, $\overline{s}+s$, and//

$$r=\widehat{\rho},\quad \hat{r}=\frac{\hat{\rho}}{8},\quad \hat{s}:=\frac{s}{8},\quad \hat{\varepsilon}:=\frac{\varepsilon}{8},\quad \mathbb{L}=\{0\}.$$

We check that that (61) and (62) imply conditions (49) and (50). We start with (49). If $(I, p, q) = (I_1, I_2, p, q) \in A_{\widehat{\rho},\varepsilon}$ and $k \in \mathbf{Z}^3 \setminus \{0\}$, with $|k|_1 \le K$, then there exists some $I_0(pq) = (I_{01}(pq), I_{02}(pq))$ such that $|I - I_0(pq)| < \widehat{\rho}$ and $\omega(I_0(pq), pq) = (\omega_{01}, \omega_{02}) \in \mathcal{D}_{\gamma_1,\gamma_2,\tau}$. We have

$$
\begin{aligned}
|\varpi(I, pq)\cdot k| &= \Big|\omega_{01}\cdot k_1 + \omega_{02}\cdot k_2 + (\omega_1(I, pq) - \omega_1(I(pq), pq))\cdot k_1 \\
&\quad + (\omega_2(I, pq) - \omega_2(I(pq), pq)))\cdot k_2 - \mathrm{i}\nu(I, pq)k_3\Big| \\
&\ge \begin{cases} \min\left\{\dfrac{\gamma_1}{2K^\tau}, \dfrac{\lambda}{2}\right\} & \text{if } k_1 \ne 0 \\ \min\left\{\dfrac{\gamma_2}{2K^\tau}, \dfrac{\lambda}{2}\right\} & \text{if } k_1 = 0,\ k_2 \ne 0 \\ \lambda & \text{if } k_1 = k_2 = 0,\ k_3 \ne 0 \end{cases} \\
&\ge \begin{cases} \alpha_1 := \dfrac{\gamma_1}{2K^\tau} & \text{if } k_1 \ne 0 \\ \alpha_2 := \dfrac{\gamma_2}{2K^\tau} & \text{if } k_1 = 0,\ (k_2, k_3) \ne (0, 0) \end{cases}
\end{aligned}
$$

having used (63). The bounds above have been obtained considering separately the cases $k_3 \ne 0$ and $k_3 = 0$, and:
–if $k_3 \ne 0$, taking the infimum of the modulus of the imaginary part of the expression between the |'s; observing that $\overline{\omega}_0 = (\overline{\omega}_{01}, \overline{\omega}_{02})$ are real and bounding the differences $|\operatorname{Im}\big(\omega_i(I, pq) - \omega_i(I(pq), pq)\big)|$ with $MK\widehat{\rho}$ (when $i = 1$), $\widehat{M}K\widehat{\rho}$ (when $i = 2$) and using the definition of $\widehat{\rho}$ in (55).
–if $k_3 = 0$, using the Diophantine inequality and again bounding the differences $|\operatorname{Im}\big(\omega_i(I, pq) - \omega_i(I(pq), pq)\big)|$ as in the previous case and using the definition $\widehat{\rho}$.
We now check condition (50). The inequality $Ks > 8\log 2$ is trivial by definition of K (see (54)), and also, the smallness condition (50) is easily met, since $\hat{\sigma} = \min\{\frac{1}{8}\frac{\varepsilon}{\bar{\varepsilon}+\varepsilon},\ \frac{s}{8}\} = \frac{s}{8}$, $\delta = 2^{-6}\min\{\hat{\rho}s,\ \varepsilon^2\} = 2^{-6}\widetilde{\rho}s$ (by the definition of $\widetilde{\rho}$ in (56)), whence

$$
2^3 c_1 \frac{K\frac{s}{8}}{\alpha_2\delta}\|f\|_{W_{\widehat{\rho},\varepsilon,s}} \le 2^6 c_1 \frac{EL}{\widehat{\rho}\widetilde{\rho}} \le \widehat{c}\widehat{E} < 1
$$

having used $L \ge \widehat{M}^{-1}, M^{-1}$, so $\alpha_2 \ge KL^{-1}\widehat{\rho}$, $2^6 c_1 < \widehat{c}$, and (62). Thus, by Proposition 3.1, H may be conjugated to

$$
\mathrm{H}_+ := \mathrm{H}\circ\Psi_+ = \mathrm{h}_+(I_+, p_+q_+) + f_+(I_+, \varphi_+, p_+, q_+)
$$

where

$$
\mathrm{h}_+(I_+, p_+q_+) = \mathrm{h}(I_+, p_+q_+) + \mathrm{g}(I_+, p_+q_+)
$$

while, by (51) and the choice of K,

$$\| f_+ \|_{\widehat{\rho}/2,\varepsilon/2,s/2} \leq e^{-Ks/32} E \leq \frac{ELM^2}{{\gamma_1}^2} E = E_+ \ . \tag{65}$$

The conjugation is realized by an analytic transformation

$$\Psi_+ : \quad (I_+, \varphi_+, p_+, q_+,) \in W_{\widehat{\rho}/2,\varepsilon/2,s/2} \to (I, \varphi, p, q) \in W_{\widehat{\rho},\varepsilon,s} \ .$$

Using (52), $\widetilde{\rho} \leq \varepsilon^2/s$, $\alpha_1 \geq MK\widehat{\rho}$, $\alpha_2 = \frac{\gamma_2}{2K^\tau} \geq L^{-1}K\widehat{\rho}$, $Ks \geq 6$ and the definition of $\widehat{E}$, we obtain the bound (59) with, at the left hand side, the set $W_{\widehat{\rho}/2,\varepsilon/2,s/2}$. Below we shall prove that $W_{+\rho_+,\varepsilon_+,s_+} \subset W_{\widehat{\rho}/2,\varepsilon/2,s/2}$, so we shall have (59).
We now evaluate the generalized frequency

$$\varpi_+(I_+, p_+q_+) := \partial_{I_+,p_+q_+} \mathrm{h}_+(I_+, p_+q_+) = \big(\omega_+(I_+, p_+q_+),\ \nu_+(I_+, p_+q_+)\big).$$

with

$$\omega_+(I_+, p_+q_+) := \partial_{I_+} \mathrm{h}_+(I_+, p_+q_+) = \partial_{I_+} \mathrm{h}(I_+, p_+q_+) + \partial_{I_+} \mathrm{g}(I_+, p_+q_+) \tag{66}$$

(the "new frequency map") and

$$\nu_+(I_+, p_+q_+) := \partial_{p_+q_+} \mathrm{h}_+(I_+, p_+q_+) = \nu(I_+, p_+q_+) + \partial_{p_+q_+} \mathrm{g}(I_+, p_+q_+) \tag{67}$$

(the "new Lyapunov exponent").

Lemma 3.2 *Let $(p_+, q_+) \in B_{\varepsilon/2}$. The new frequency map ω_+ is injective on $D(p_+q_+)_{\widehat{\rho}/2}$ and maps $D(p_+q_+)_{\widehat{\rho}/4}$ over $\omega(D, p - +q_+)$. The map $\widehat{\iota}_+(p_+q_+) = (\widehat{\iota}_{+1}(p_+q_+), \widehat{\iota}_{+2}(p_+q_+)) := \omega_+^{-1} \circ \omega|_{D(p_+q_+)}$ which assigns to a point $I_0 \in D(p_+q_+)$ the $\omega_+(\cdot, p_+q_+)$-preimage of $\omega(I_0, p_+q_+)$ in $D(p_+q_+)_{\widehat{\rho}/4}$ satisfies*

$$\sup_{(A_+)_{\rho_+,\varepsilon_+}} |\widehat{\iota}_{+1}(p_+q_+) - \mathrm{id}\,| \leq 3n \frac{\overline{M}_1 E}{\widehat{\rho}} \leq 3n \frac{\overline{M} E}{\widehat{\rho}} \ ,$$
$$\sup_{(A_+)_{\rho_+,\varepsilon_+}} |\widehat{\iota}_{+2}(p_+q_+) - \mathrm{id}\,| \leq 3n \frac{\overline{M}_2 E}{\widehat{\rho}} \leq 3n \frac{\overline{M} E}{\widehat{\rho}} \ ,$$
$$\mathcal{L}(\widehat{\iota}_+(p_+q_+) - \mathrm{id}\,) \leq 2^9 n \frac{\overline{M} E}{\widehat{\rho}^2} \ . \tag{68}$$

The Jacobian matrix $U_+ := \partial^2_{I_+} \mathrm{h}_+(I_+, p_+q_+)$ is non-singular on $D_{\widehat{\rho}/4} \times B^2_{\varepsilon/2}$ and the following bounds hold

$$M_+ := 2M \geq \sup_{(A_+)_{\rho_+,\varepsilon_+}} \|U_+\| \ , \quad \widehat{M}_+ := 2\widehat{M} \geq \sup_{(A_+)_{\rho_+,\varepsilon_+}} \|\widehat{U}_+\| \ ,$$

$$\overline{M}_{+} := 2\overline{M} \geq \sup_{(A_+)_{\rho_+,\varepsilon_+}} \|U_+^{-1}\| \,, \quad \overline{M}_{i+} := 2\overline{M}_i \geq \sup_{(A_+)_{\rho_+,\varepsilon_+}} \|T_{i+}\| \,, \quad i = 1,\ 2. \tag{69}$$

where $U_+^{-1} =: \begin{pmatrix} T_{+1} \\ T_{+2} \end{pmatrix}$. *Finally, the new Lyapunov exponent* $\nu_+(I_+, p_+q_+)$ *satisfies*

$$\lambda_+ := \lambda - 2^4 \frac{E}{\varepsilon^2} \leq \inf_{(A_+)_{\rho_+,\varepsilon_+}} |\operatorname{Re} \nu_+|. \tag{70}$$

Postponing for the moment the proof of this lemma, we let $\rho_+ := \widehat{\rho}/2$, $s_+ := s/2$, $\varepsilon_+ = \varepsilon/2$ and $D_+(p_+q_+) := \widehat{\iota}_+(p_+q_+)(D(p_+q_+))$. By Lemma 3.2, D_+ is a subset of $D_{\widehat{\rho}/4}$ and hence

$$(D_+)_{\rho_+} \subset D_{\widehat{\rho}/2}\ . \tag{71}$$

We prove that $\widehat{E}_+ = \frac{E_+L_+}{\widehat{\rho}_+^2} \leq \widehat{E}^2$. Since

$$s_+ = \frac{s}{4} \quad \text{and} \quad x_+ := \Big(\frac{E_+L_+M_+^2}{{\gamma_1}^2}\Big)^{-1} = \frac{x^2}{8} \quad \text{where} \quad x := \Big(\frac{ELM^2}{{\gamma_1}^2}\Big)^{-1} \tag{72}$$

we have

$$K_+ = \frac{2^5}{s_+} \log x_+ = \frac{2^7}{s} \log \frac{x^2}{8} = \frac{2^8}{s} \log_+ x - \frac{3 \cdot 2^7}{s} \log_+ 2 < 8K. \tag{73}$$

Finally, (42), (70) and the definition of $\tilde{E}$ imply $\lambda_+ \geq \frac{\lambda}{2}$. Collecting all bounds, we get

$$\begin{aligned} \widehat{\rho}_+ &= \min\left\{\frac{\gamma_1}{2M_+K_+^{\tau+1}},\, ,\ \frac{\gamma_2}{2\widehat{M}_+K_+^{\tau+1}},\ \frac{\lambda_+}{2M_+K_+},\ \frac{\lambda_+}{2\widehat{M}_+K_+},\ \rho_+ = \frac{\widehat{\rho}}{2}\right\} \geq \frac{\widehat{\rho}}{2 \cdot 8^{\tau+1}} \\ \widetilde{\rho}_+ &= \min\left\{\widehat{\rho}_+,\ \frac{\varepsilon_+^2}{s_+}\right\} \geq \frac{\widetilde{\rho}}{2 \cdot 8^{\tau+1}} \end{aligned} \tag{74}$$

and

$$\widehat{E}_+ = \frac{E_+L_+}{\widehat{\rho}_+\widetilde{\rho}_+} \leq \frac{E^2LM^2}{{\gamma_1}^2} \frac{2L}{\widehat{\rho}\widetilde{\rho}} 4 \cdot 8^{2(\tau+1)} = 8 \cdot 8^{2(\tau+1)} \frac{E\,LM^2}{{\gamma_1}^2} \widehat{E}$$

Now, using, in the last inequality, the bound

$$\frac{E\,LM^2}{{\gamma_1}^2} \leq \frac{1}{4} \left(\frac{s}{6}\right)^{2(\tau+1)} \frac{EL}{\widehat{\rho}^2} \leq \frac{1}{4} \left(\frac{s}{6}\right)^{2(\tau+1)} \widehat{E}$$

(since $\rho\hat{}\le\frac{\gamma_1}{2MK^{\tau+1}}$ and $K\ge 6/s$) we find

$$\widehat{E}_+ \le 2(\frac{4}{3}s)^{\tau+1}\widehat{E}^2 < \widehat{E}^2 \tag{75}$$

(having used $s \le 1/2$). We now prove that $\lambda_+ \ge \frac{\lambda_0}{2}$. Iterating (70) and using $\widehat{\rho}_k \le \widehat{\rho}_{k-1}/4$, $\widetilde{\rho}_k \le \widetilde{\rho}_{k-1}/4$, $\varepsilon_k = \varepsilon_{k-1}/4$, $L_k = 2L_{k-1}$, (75) and the second condition in (42) with $\tilde{c} = 2^6$, we get

$$\begin{aligned}\lambda_+ &= \lambda_{j+1} = \lambda_0 - 2^4\sum_{k=1}^{j}\frac{E_k}{\varepsilon_k^2} \ge \lambda_0 - 2^4\sum_{k=1}^{j}\widehat{E}_k\frac{\widehat{\rho}_k\widetilde{\rho}_k}{\varepsilon_k^2 L_k} \ge \lambda_0 - 2^4\frac{\widehat{\rho}_0\widetilde{\rho}_0}{\varepsilon_0^2 L_0}\sum_{k=1}^{j}\widehat{E}_k\\ &\ge \lambda_0 - 2^5\frac{\widehat{\rho}_0\widetilde{\rho}_0}{\varepsilon_0^2 L_0}\widehat{E}_0\\ &= \lambda_0 - 2^5\frac{E_0}{\varepsilon_0^2} \ge \frac{\lambda_0}{2}.\end{aligned} \tag{76}$$

This allows to check (63) at the next step: using (64) and (73), we have

$$\lambda_+ \ge \frac{\lambda_0}{2} \ge \frac{\gamma_2}{K_0^\tau} \ge \frac{\gamma_2}{K_+^\tau}.$$

Finally, (57) and (58) follow from (68), while the estimate in (60) is a consequence of (59), (71), (72), (74), inequality $LM \ge 1$ and Cauchy estimates:

$$\begin{aligned}\mathcal{L}(\check{\Phi}_{j+1} - \mathrm{id}) &\le 2(n+1)\sup_{(\check{W}_{j+1})_{\rho_{j+1},\varepsilon_{j+1},s_{j+1}}}\|D(\check{\Phi}_{j+1} - \mathrm{id})\|_\infty\\ &\le 2(n+1)\frac{\frac{3}{4}\widehat{E}_j\max\{\widehat{\rho}_j/\rho_0, s_j/s_0, \varepsilon_j/\varepsilon_0\}}{\min\{\widehat{\rho}_j/(4\widehat{\rho}_0), s_j/(4s_0), \varepsilon_j/(4\varepsilon_0)\}}\\ &\le 2(n+1)\frac{3/4(1/4)^j}{1/4\left(\frac{1}{2(24)^{\tau+1}}\right)^j}\widehat{E}_j = 6(n+1)\big(12\cdot(24)^\tau\big)^j\widehat{E}_j\ .\end{aligned}$$

Proof of Lemma 3.2 The proof of this proposition is obtained generalizing (Chierchia and Pinzari 2010, Lemma B.2). As above, we limit to discuss only the different parts. By (51),

$$\sup_{D_{\widehat{\rho}/2}\times B^2_{\varepsilon/2}}|\mathrm{g}| \le \sup_{D_{\widehat{\rho}/2}\times B^2_{\varepsilon/2}}|\mathrm{g}-\overline{f}| + \sup_{D_{\widehat{\rho}/2}\times B^2_{\varepsilon/2}}|\overline{f}| \le \frac{3}{2}E\ ,$$

(where $\overline{f}$ denotes the average of f). Therefore we may bound

$$\sup_{D_{\widehat{\rho}/4}\times B^2_{\varepsilon/2}}\|(\partial^2_{I_+}\mathrm{h})^{-1}\partial^2_{I_+}\mathrm{g}\| \le 2\overline{M}\frac{\frac{3}{2}E}{(\widehat{\rho}/4)^2} \le 2^6\frac{\overline{ME}}{\widehat{\rho}^2} \le 2^6\frac{\overline{ME}}{\widehat{\rho}^2} < \frac{1}{2}$$

This shows that the function (66) has a Jacobian matrix

$$\partial_{I_+}\omega_+(I_+, p_+q_+) = \partial^2_{I_+}\mathrm{h}_+(I_+, p_+q_+) = \partial^2_{I_+}\mathrm{h}(I_+, p_+q_+) + \partial^2_{I_+}\mathrm{g}(I_+, p_+q_+)$$

which is invertible for all $(p_+, q_+) \in B^2_{\varepsilon/2}$ and satisfies

$$\overline{M}_+ := \sup_{D_{\widehat{\rho}/4}\times B^2_{\varepsilon/2}} \left\| \left(\partial_{I_+}\omega_+(I_+, p_+q_+)\right)^{-1} \right\| \le 2\overline{M}$$

In a similar way one proves (69). Next, for any fixed $(p_+, q_+) \in B^2_{\varepsilon/2}$ and $\overline{\omega} = \omega(I(p_+q_+), p_+q_+) \in \omega(D, p_+q_+)$ with $I(p_+q_+) \in D$, we want to find $I_+ = I_+(p_+q_+) \in D_+$ such that

$$\omega_+(I_+(p_+q_+), p_+q_+) = \overline{\omega} = \omega(I(p_+q_+), p_+q_+) \tag{77}$$

To this end, we consider the function

$$I_+ \in D_{\widehat{\rho}/2} \to F(I_+, p_+q_+) := \omega_+(I_+, p_+q_+) - \overline{\omega} \quad (p_+, q_+) \in B^2_{\varepsilon/2}$$

As F differs from ω_+ by a constant, we have

$$m := \sup_{D_{\widehat{\rho}/4}\times B^2_{\varepsilon/2}} \left\| \left(\partial_{I_+} F(I_+, p_+q_+)\right)^{-1} \right\| = \sup_{D_{\widehat{\rho}/4}\times B^2_{\varepsilon/2}} \left\| \left(\partial_{I_+}\omega_+(I_+, p_+q_+)\right)^{-1} \right\| \le 2M.$$

Similarly, we bound the quantities

$$Q := |\partial^2_{I_+} F(I)| = |\partial^3_{I_+}\mathrm{g}(I_+, p_+q_+)| \le 6\frac{\frac{3}{2}E}{(\widehat{\rho}/4)^3} < 2^{10}\frac{E}{\widehat{\rho}^3}.$$

and

$$P := |F(I(p_+q_+))| = |\partial_{I_+}\mathrm{g}(I(p_+q_+), p_+q_+)| \le \frac{\frac{3}{2}E}{(\widehat{\rho}/4)} \le 2^3\frac{E}{\widehat{\rho}}.$$

Putting everything together, we get

$$4m^2PQ \le 2^{16}\frac{M^2E^2}{\widehat{\rho}^4} \le \widehat{c}^2\widehat{E}^2 < 1$$

By the implicit function theorem (e.g., (Celletti and Chierchia 1998, Theorem 1 and Remark 1)), Equation (77) has a unique solution

$$(p_+, q_+) \in B_{\varepsilon/2} \to I_+(p_+q_+) \in B_r(I(p_+q_+)),$$

with

$$r = 2mP \le 2^5 \frac{ME}{\widehat{\rho}} \le \frac{\widehat{\rho}}{4}$$

so we can take

$$D_+(p_+q_+) = \bigcup_{\overline{\omega} \in \omega(D, p_+q_+)} \{I_+(p_+q_+)\}$$

This ensures that (61) holds also for D_+.

Finally, the real part of the function (67) satisfies the lower bound

$$\inf_{D_{\widehat{\rho}/2} \times B^2_{\varepsilon/4}} |\operatorname{Re} \nu_+| \ge \lambda - \frac{E}{(\varepsilon/4)^2} = \lambda_+.$$

The proof of (68) proceeds as in Chierchia and Pinzari (2010, proof of Lemma B.2). □

Proof of Theorem 3.1.

Step 1 *Construction of the "generalized limit actions"*

Let $(\pi, \kappa) \in B_0 = B^2_{\overline{\varepsilon}} = \bigcap_{j \ge 0} B_{\varepsilon_j}$. Define, on $D_0(\pi\kappa) = \omega_0^{-1}(\mathcal{D}_{\gamma_1, \gamma_2, \tau}, \pi\kappa) \cap D$,

$$\check{\iota}_j(\pi\kappa) := \widehat{\iota}_j(\pi\kappa) \circ \widehat{\iota}_{j-1}(\pi\kappa) \circ \cdots \circ \widehat{\iota}_1(\pi\kappa) \quad j \ge 1.$$

Then $\check{\iota}_j(\pi\kappa)$ converge uniformly to a $\check{\iota}(\pi, \kappa) = (\check{\iota}_1(\pi, \kappa), \check{\iota}_2(\pi, \kappa))$ verifying

$$\sup_{D_0(\pi\kappa)} |\check{\iota}_1(\pi\kappa) - \mathrm{id}\,| \le 6n \frac{\overline{M}_1}{\overline{\overline{M}}} \widetilde{\rho}_0 \widehat{E}_0, \qquad \sup_{D_0(\pi\kappa)} |\check{\iota}_2(\pi\kappa) - \mathrm{id}\,| \le 6n \frac{\overline{M}_i}{\overline{\overline{M}}} \widetilde{\rho}_0 \widehat{E}_0. \tag{78}$$

Moreover, as

$$\sup |\widehat{\iota}_j(\pi\kappa) - \widehat{\iota}(\pi\kappa)| \le 6n \widehat{E}_j \widetilde{\rho}_j < \frac{6n}{\widehat{c}} \widehat{\rho}_j < \rho_j$$

we have

$$D_*(pq) := \check{\iota}(\pi\kappa)(D_0(\pi\kappa)) \subset \bigcap_j D_j(\pi\kappa)_{\rho_j}. \tag{79}$$

In particular, taking $j = 0$,

$$D_*(\pi\kappa) \subset (D_0(\pi\kappa))_{6n\widehat{E}_0\widetilde{\rho}_0}. \tag{80}$$

Moreover,

$$\mathcal{L}(\check{\iota}(\pi\kappa) - \mathrm{id}\,) \le 2^8 n \widehat{E}.$$

So $\check{\imath}(\pi\kappa)$ is bi-Lipschitz, with

$$\mathcal{L}_-(\check{\imath}(\pi\kappa)) \geq 1 - 2^8 n\widehat{E}, \qquad \mathcal{L}_+(\check{\imath}(\pi\kappa)) \leq 1 + 2^8 n\widehat{E}.$$

Step 2 *Construction of* ϕ_{ω_*}. For each $j \geq 1$, the transformation

$$\Phi_j := \Psi_1 \circ \cdots \circ \Psi_j$$

is defined on $(W_j)_{\rho_j, s_j, \varepsilon_j}$. If

$$A_* := \bigcup_{|(\pi,\kappa)|<\overline{\varepsilon}} D_*(\pi\kappa) \times \{(\pi,\kappa)\}, \qquad W_* := A_* \times \mathbb{T}^n.$$

then, by (79), $W_* \subset \bigcap_j (W_j)_{\rho_j,s_j,\varepsilon_j}$. The sequence Φ_j converges uniformly on W_* to a map Φ. We then let

$$\begin{aligned}\phi_{\omega_*}(\vartheta,\pi,\kappa) &= \Big(v(\vartheta,\pi,\kappa;\omega_*), \vartheta + u(\vartheta,\pi,\kappa;\omega_*), \pi + w(\vartheta,\pi,\kappa;\omega_*), \kappa + y(\vartheta,\pi,\kappa;\omega_*)\Big)\\ &:= \Phi\left(\check{\imath}(\omega_0^{-1}(\omega_*,\pi\kappa)), \vartheta, \pi, \kappa\right)\end{aligned}$$

with $v(\vartheta,\pi,\kappa;\omega_*) := \big(v_1(\vartheta,\pi,\kappa;\omega_*), v_2(\vartheta,\pi,\kappa;\omega_*)\big)$. Since (59) imply, on W_*,[22]

$$\sup_{W_*} |\P_{I_1}\Phi - \mathrm{id}\,|_1 \leq 2n\frac{\widehat{M}_0}{M_0}\widehat{E}_0\widetilde{\rho}_0 \tag{81}$$

and similarly,

$$\begin{aligned}&\sup_{W_*} |\P_{I_2}\Phi - \mathrm{id}\,|_1 \leq 2n\widehat{E}_0\widetilde{\rho}_0\,, \quad \sup_{W_*} |\P_{\varphi}\Phi - \mathrm{id}\,|_\infty \leq 2\widehat{E}_0 s_0\,,\\ &\sup_{W_*} |\P_p\Phi - \mathrm{id}\,|_\infty \leq 2\widehat{E}_0\varepsilon_0, \quad \sup_{W_*} |\P_q\Phi - \mathrm{id}\,|_\infty \leq 2\widehat{E}_0\varepsilon_0\end{aligned} \tag{82}$$

then, in view of (78), (81), (82), the definition of W_* and the triangular inequality, we have (46). Equations (80), (81), (82) also imply

$$\mathrm{T}_{\omega_*} := \phi_{\omega_*}(\mathbb{T}^n, 0, 0) \subset (D_*(0))_{2\widehat{E}_0\widetilde{\rho}_0} \times \mathbb{T}^n \times B^2_{r'} \subset (D_0(0))_r \times \mathbb{T}^n \times B^2_{r'} \tag{83}$$

where

$$r = 8n\widehat{E}_0\widetilde{\rho}_0, \qquad r' = 2\widehat{E}_0\varepsilon_0$$

Finally, with similar arguments as in Step 1, by (84), the rescaled map

$$\check{\Phi} := \mathrm{id} + 1_{\widehat{\rho}_0^{-1}, s_0^{-1}, \varepsilon_0^{-1}}(\Phi - \mathrm{id}\,) \circ 1_{\widehat{\rho}_0, s_0, \varepsilon_0}$$

[22] $\P_z$ denotes the projection on the z-variables.

has Lipschitz constant

$$\mathcal{L}(\check{\Phi} - \mathrm{id}\,) \le 2^6(n+1)\widehat{E}_0\ . \tag{84}$$

In particular, $\check{\Phi}$, hence, Φ, and, finally, the map $(\vartheta, \pi, \kappa; \omega) \to \phi_\omega(\vartheta, \pi, \kappa)$ are bi-Lipschitz, hence, injective.

Step 3 *For any* $\omega_* \in \mathcal{D}_{\gamma_1,\gamma_2,\tau} \cap \omega_0(D, 0)$, T_{ω_*} *in* (83) *is a* n*-dimensional* H*-invariant torus with frequency* ω_*. This assertion is a trivial generalization of its analogue one in Chierchia and Pinzari (2010, Proof of Proposition 3, Step 3); therefore, its proof is omitted.

Step 4 *Measure Estimates (proof of* (45)) The proof of (45) proceeds as in Chierchia and Pinzari (2010, Proof of Proposition 3, Step 4), just replacing the quantities that in Chierchia and Pinzari (2010, Proof of Proposition 3, Step 4) are called

$$D_0, \quad D_*, \quad \check{\imath}, \quad \check{\Phi}, \quad \mathrm{K}$$

with the quantities here denoted as

$$D_0(0), \quad D_*(0), \quad \check{\imath}(0), \quad \check{\Phi}\big|_{(\pi,\kappa)=(0,0)}, \quad \mathrm{K}_0.$$

□

3.4 Normal Form Theory

Proposition 3.1 can be obtained from the more general Proposition 3.2, taking $m = 1$, $\mathbb{L} = \{0\}$ and changing coordinates as follows:

$$p = \frac{p_1 - \mathrm{i}q_1}{\sqrt{2}}\ , \qquad q = \frac{p_1 + \mathrm{i}q_1}{\sqrt{2}\mathrm{i}}\ .$$

We define c_m to be the smallest number such that, for any two functions, real-analytic in $W_{r,s,\varepsilon}$ and any choice of $\hat{r} < r$, $\hat{s} < s$, $\hat{\varepsilon} < \varepsilon$,

$$\|\{f, g\}\|_{r-\hat{r},s-\hat{s},\varepsilon-\hat{\varepsilon}} \le \frac{c_m}{\delta}\|f\|_{r,s,\varepsilon}\|g\|_{r,s,\varepsilon} \quad \text{with } \delta := \min\{\hat{r}\hat{s}, \hat{\varepsilon}^2\}.$$

Proposition 3.2 *Let* $\{0\} \subset \mathbb{L} \subset \mathbb{Z}$. *Proposition* 3.1 *holds true taking*

$$H(I, \varphi, p, q) = h\,(I_1, I_2, J(p, q)) + f(I, \varphi, p, q), \quad J(p, q) := \left(\frac{p_1^2 + q_1^2}{2}, \ldots, \frac{p_m^2 + q_m^2}{2}\right)$$

replacing c_1 *with* c_m, $\P_0$ *with* $\P_{\mathbb{L}}$ *and condition* (49) *with*

$$|\omega_1 \cdot k_1 + \omega_2 \cdot k_2| \ge \begin{cases} \alpha_1 \text{ if } k_1 \ne 0 \\ \alpha_2 \text{ if } k_1 = 0,\ k_2 \ne 0 \end{cases}$$

$$\forall\, k = (k_1, k_2) \in \mathbb{Z}^{n_1} \times \mathbb{Z}^{n_2+m} \setminus \mathbb{L} \neq (0,0),\ |k|_1 \leq K, \quad \forall (I_1, I_2, p, q) \in V_r \times B_\varepsilon^{2m} \tag{85}$$

where

$$\begin{aligned} \omega &= (\omega_1, \omega_2) \\ &:= \left(\partial_{I_1} h\,(I_1, I_2, J(p,q))\,,\ \partial_{(I_2, J(p,q))} h\,(I_1, I_2, J(p,q))\right). \end{aligned}$$

Lemma 3.3 *Let $\hat{r} < r/2$ $\hat{s} < s/2$, $\hat{\varepsilon} < \varepsilon/2$ and $\delta := \min\{\hat{r}\hat{s},\ \hat{\varepsilon}^2\}$. Let*

$$\begin{aligned} H(u, \varphi, p, q) &= \mathrm{h}(I, p, q) + g(u, \varphi, p, q) + f(u, \varphi, p, q) \qquad g(u, \varphi, p, q) \\ &= \sum_{i=1}^{m} g_i(u, \varphi, p, q) \end{aligned}$$

be real-analytic on $W_{v,s,\varepsilon}$. Assume that inequality (85) and

$$\|f\|_{v,s,\varepsilon} < \frac{\alpha_2 \delta}{c_m}$$

are satisfied. Then, one can find a real-analytic and symplectic transformation

$$\Phi:\ W_{v-2\hat{v}, s-2\hat{s}, \varepsilon-2\hat{\varepsilon}} \to W_{v,s,\varepsilon}$$

defined by the time-one flow[23] *$X_\phi^1 f := f \circ \Phi$ of a suitable ϕ verifying*

$$\|\phi\|_{v,s,\varepsilon} \leq \frac{\|f\|_{v,s,\varepsilon}}{\alpha_2}$$

such that

$$H_+ := H \circ \Phi = h + g + \P_{\mathbb{L}} T_K f + f_+$$

and, moreover, the following bounds hold

$$\begin{aligned} \|f_+\|_{v-2\hat{v}, s-2\hat{s}, \varepsilon-2\hat{\varepsilon}} \leq \left(1 - \frac{c_m}{\alpha_2 \delta}\|f\|_{v,s,\varepsilon}\right)^{-1} \Big[\frac{c_m}{\alpha_2 \delta} \|f\|^2_{v,s,\varepsilon} \\ + \max\left\{ e^{-K\hat{s}/2},\ \left(\frac{\varepsilon - \hat{\varepsilon}}{\varepsilon}\right)^{K/2} \right\} \|f\|_{v,s,\varepsilon} + \|\{\phi, g\}\|_{v-\hat{v}, s-\hat{s}, \varepsilon-\hat{\varepsilon}} \Big] \end{aligned}$$

[23] The time-one flow generated by ϕ is defined as the differential operator

$$X_\phi^1 := \sum_{k=0}^{\infty} \frac{\mathcal{L}_\phi^k}{k!}$$

where $\mathcal{L}_\phi^0 f := f$ and $\mathcal{L}_\phi^k f := \{\phi, \mathcal{L}_\phi^{k-1} f\}$, with $k = 1,\ 2,\ \cdots$.

Finally, for any real-analytic function F *on* $W_{v,s,\varepsilon}$,

$$\|F \circ \Phi - F\|_{v-2\hat{v},s-2\hat{s},\varepsilon-2\hat{\varepsilon}} \leq \frac{\|\{\phi, F\}\|_{v-\hat{v},s-\hat{s},\varepsilon-\hat{\varepsilon}}}{1 - \frac{c_m \|f\|_{v,s,\varepsilon}}{\alpha_2 \delta}}. \tag{86}$$

Sketch of proof Lemma 3.3 is a straightforward generalization of Pöschel (1993, Iterative Lemma). To obtain such generalization, just replace the norm defined in Pöschel (1993, Section 1) with the norm (48), where

$$f = \sum_{\substack{(k,\alpha,\beta)\in\mathbf{Z}^n\times\mathbb{N}^\ell\times\mathbb{N}^\ell \\ \alpha_i\neq\beta_i\forall i}} f_{k,\alpha,\beta}(I) e^{ik\cdot\varphi} \left(\frac{p - \mathrm{i}q}{\sqrt{2}}\right)^\alpha \left(\frac{p + \mathrm{i}q}{\mathrm{i}\sqrt{2}}\right)^\beta, \tag{87}$$

and bound the "ultraviolet remainders", namely the norm of the functions whose expansion (87) includes only terms with $|(k, \alpha - \beta)|_1 > K$, as follows. Observe that, if $|(k, \alpha - \beta)|_1 > K$, then either $|k|_1 > K/2$ or $|\alpha - \beta|_1 > K/2$. In the latter case, a fortiori, $|\alpha|_1 + |\beta|_1 \geq |\alpha - \beta|_1 > K/2$. Then we have, for such functions, $\|f\|_{r,s-\hat{s},\varepsilon-\hat{\varepsilon}} \leq \max\left\{e^{-K\hat{s}/2}, \left(\frac{\varepsilon-\hat{\varepsilon}}{\varepsilon}\right)^{K/2}\right\} \|f\|_{r,s,\varepsilon}$. Other details are omitted.

Proof of Proposition 3.2 Let

$$r_1 := r_0 - 2\hat{r}_0, \qquad s_1 := s_0 - 2\hat{s}_0, \quad \varepsilon_1 := \varepsilon_0 - 2\hat{\varepsilon}_0.$$

By Lemma 3.3, we find a canonical transformation $\Phi_1 = X_{\phi_1}$ which is real-analytic on $W_{r_1,s_1,\varepsilon_1}$ and conjugates $H = H_0$ to $H_1 = H_0 \circ \Phi_1 = h + g_1 + f_1$, where $g_1 = ¶_{\mathbb{L}} T_K f_0$ and

$$\begin{aligned}\|f_1\|_{v_1,s_1,\varepsilon_1} &\leq (1 - \frac{c_m E_0}{\alpha_2 \delta_0})^{-1}\Big[\frac{c_m E_0}{\alpha_2 \delta_0} + \max\left\{e^{-K\hat{s}_0/2}, (\frac{\varepsilon_0 - \hat{\varepsilon}_0}{\varepsilon_0})^{K/2}\right\}\Big] E_0 \\ &\leq 2\Big[\frac{c_m E_0}{\alpha_2 \delta_0} + e^{-K\hat{\sigma}_0/2}\Big] E_0\end{aligned}$$

having used

$$(\frac{\varepsilon_0 - \hat{\varepsilon}_0}{\varepsilon_0})^{K/2} = e^{\frac{K}{2}\log\left(1-\frac{\hat{\varepsilon}_0}{\varepsilon_0}\right)} \leq e^{-\frac{K}{2}\frac{\hat{\varepsilon}_0}{\varepsilon_0}}.$$

We now focus on the case

$$\frac{c_m E_0}{\alpha_2 \delta_0} < e^{-K\hat{\sigma}_0/2}$$

otherwise the lemma is[24] proved. Then, we have

$$\|f_1\|_{v_1,s_1,\varepsilon_1} \le 4\frac{c_m E_0^2}{\alpha_2\delta_0} =: E_1.$$

Note that

$$E_1 < \frac{E_0}{4}.$$

Assume now that, for some $j \ge 1$, it is $H_j = H_{j-1} \circ \Phi_j = h + g_j + f_j$, where

$$g_j = \sum_{h=0}^{j-1} \P_{\mathbb{L}} T_K f_h, \qquad \|f_j\|_{v_j,s_j,\varepsilon_j} \le E_j \le \min\left\{\frac{E_0}{4^j},\ 4\frac{c_m E_0^2}{\alpha_2\delta_0}\right\}. \tag{88}$$

We have just proved this is true when $j = 1$. Let $L := \left[\frac{K\hat{\sigma}_0}{8\log 2}\right]$. We prove that (88) is true for $j + 1$, for all $1 \le j \le L$. Let

$$\hat{r}_j := \frac{\hat{r}_0}{L}, \quad \hat{s}_j := \frac{\hat{s}_0}{L}, \quad \hat{\varepsilon}_j := \frac{\hat{\varepsilon}_0}{L} \quad \text{hence} \quad \delta_j = \frac{\delta_0}{L^2} \quad \forall\, 1 \le j \le L.$$

Note that, for all $1 \le j \le L$, it is $\hat{r}_j < \frac{r_j}{2}$:

$$r_j = r_1 - 2(j-1)\frac{\hat{r}_0}{L} \ge r_1 - 2(1 - 1/L)\hat{r}_0 = r_0 - 4\hat{r}_0 + 2\hat{r}_j > 2\hat{r}_j.$$

Similarly, $\hat{s}_j < \frac{s_j}{2}$, $\hat{\varepsilon}_j < \frac{\varepsilon_j}{2}$. Let then

$$r_{j+1} = r_j - 2\frac{\hat{r}_0}{L}, \quad s_{j+1} = s_j - 2\frac{\hat{s}_0}{L}, \quad \varepsilon_{j+1} = \varepsilon_j - 2\frac{\hat{\varepsilon}_0}{L}$$

so that $r_j = r_1 - 2(j-1)\frac{\hat{r}_0}{L}$, etc., for all $1 \le j \le L$. Then

$$c_m \frac{E_j}{\alpha_2\delta_j} \le 4\frac{c_0^2 E_0^2}{\alpha_2^2\delta_0^2} L^2 < \frac{1}{16} \tag{89}$$

[24] Indeed, in such case,

$$\|f_1\|_{v_1,s_1,\varepsilon_1} \le 4e^{-K\hat{\sigma}_0/2} \le e^{-K\hat{\sigma}_0/4}$$

because $e^{-K\hat{\sigma}_0/4} \le \frac{1}{4}$ having chosen $K\hat{\sigma}_0 \ge 8\log 2$.

and Lemma 3.3 applies again, and H_j can be conjugated to $H_{j+1} = H_j \circ \Phi_{j+1} = h + g_{j+1} + f_{j+1}$, with

$$g_{j+1} = g_j + \P_{\mathbb{L}} T_K f_j = \sum_{h=0}^{j} \P_{\mathbb{L}} T_K f_h$$
$$\|f_{j+1}\|_{r_{j+1},s_{j+1},\varepsilon_{j+1}} \le \big(1 - \frac{c_m}{\alpha_2 \delta_j} E_j\big)^{-1} \Big[\frac{c_m}{\alpha_2\delta_j} E_j^2 + \max\Big\{ e^{-K\hat{s}_j/2},\ \big(\frac{\varepsilon_j - \hat{\varepsilon_j}}{\varepsilon_j}\big)^{K/2} \Big\} E_j + \|\{\phi_j, g_j\}\|_{r_j-\hat{r}_j, s_j-\hat{s}_j, \varepsilon_j-\hat{\varepsilon_j}} \Big]$$

To bound the right hand side of the latter expression, we use (89) and observe that

$$e^{-K\widehat{s}_j/2} = e^{-\frac{K}{2L}\widehat{s}_0} \le \frac{1}{16}$$
$$\left(\frac{\varepsilon_j - \hat{\varepsilon_j}}{\varepsilon_j}\right)^{K/2} = \left(1 - \frac{\frac{\widehat{\varepsilon}_0}{L}}{\varepsilon_1 - 2(j-1)\frac{\widehat{\varepsilon}_0}{L}}\right)^{K/2} \le \left(1 - \frac{\widehat{\varepsilon}_0}{\varepsilon_1 L}\right)^{K/2} \le e^{-\frac{K\widehat{\varepsilon}_0}{2\varepsilon_1 L}} \le \frac{1}{16}$$

having used $e^{-\frac{K\widehat{s}_0}{2L}} \le e^{-\frac{K\hat{\sigma}_0}{2L}}$, $e^{-\frac{K\widehat{\varepsilon}_0}{2\varepsilon_1 L}} \le e^{-\frac{K\widehat{\varepsilon}_0}{2\varepsilon_0 L}} \le e^{-\frac{K\hat{\sigma}_0}{2L}}$ and $L \le \frac{K\hat{\sigma}_0}{8\log 2}$. Moreover, writing

$$g_j = \P_{\mathbb{L}} T_K f_0 + \mathbb{1}_{j\ge 2} \sum_{h=1}^{j-1} \P_{\mathbb{L}} T_K f_h =: f_0^{\mathbb{L},K} + f_{j-1}^{\mathbb{L},K}$$

with $f_0^{\mathbb{L},K}$ real-analytic on $W_{r_0,s_0,\varepsilon_0}$, while $f_{j-1}^{\mathbb{L},K}$ real-analytic on $W_{r_{j-1},s_{j-1},\varepsilon_{j-1}}$ and verifying

$$\|f_0^{\mathbb{L},K}\|_{r_0,s_0,\varepsilon_0} \le E_0, \quad \|f_{j-1}^{\mathbb{L},K}\|_{r_{j-1},s_{j-1},\varepsilon_{j-1}} \le \sum_{h=1}^{j-1} \frac{E_1}{4^{j-1}} \le \frac{4}{3} E_1$$

we get

$$\|\{\phi_j, g_j\}\|_{r_j-\hat{r}_j, s_j-\hat{s}_j, \varepsilon_j-\hat{\varepsilon_j}} \le \frac{c_m L}{\alpha_2\delta_0} E_0 E_j + \frac{4}{3}\frac{c_m L^2}{\alpha_2\delta_0} E_1 E_j$$
$$\le \left(\frac{c_m L}{\alpha_2\delta_0} E_0 + \frac{16}{3}\frac{c_m^2 L^2}{\alpha_2^2\delta_0^2} E_0^2\right) E_j$$
$$\le \left(\frac{1}{32} + \frac{1}{32}\right) E_j = \frac{E_j}{16}$$

Collecting all such bounds we get

$$E_{j+1} \le \frac{16}{15}\frac{3}{16}E_j < \frac{E_j}{4}.$$

The inductive claim $j \to j+1$ is thus proved, for all $1 \le j \le L$. Letting now $\Phi_* := \Phi_1 \circ \cdots \circ \Phi_{L+1}$ and

$$\begin{aligned} &H_* := H_{L+1} = h + g_{L+1} + f_{L+1} =: h + g_* + f_* \\ &r_* := r_{L+1} = r - 4\hat{r},\; s_* := s_{L+1} = s - 4\hat{s},\; \varepsilon_* := \varepsilon_{L+1} = \varepsilon - 4\hat{\varepsilon} \end{aligned}$$

and using $L + 1 > \frac{K\widehat{\sigma}_0}{8\log 2}$, we get

$$\begin{aligned} &\|f_*\|_{r_*,s_*,\varepsilon_*} \le \frac{E_0}{4^{L+1}} = e^{-2(L+1)\log 2}E_0 < e^{-\frac{K\widehat{\sigma}_0}{4}}E_0 \\ &\|g_* - ¶_{\mathbb{L}} T_K f_0\|_{r_*,s_*,\varepsilon_*} \le \frac{4}{3}E_1 < 8\frac{c_m E_0^2}{\alpha_2 \delta_0} \end{aligned}$$

as claimed. The bounds (52) are obtained from (86), by usual telescopic arguments. □

Acknowledgements G.P. is indebted to G. Gallavotti for many highlighting discussions. Theorem 1.1 is part of the MSc thesis of X.L. This research is funded by the PRIN project "New frontiers of Celestial Mechanics: theory and applications" and has been developed under the auspices of INdAM and GNFM. The authors confirm that the manuscript has no associated data.

Author Contributions GP conceptualized and developed the work and wrote the main manuscript. XL wrote a draft of Theorem 1.1, which is part of his MSc thesis at the University of Padova. All authors reviewed the manuscript.

Funding Open access funding provided by Università degli Studi di Padova within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Arnold, V.I.: Small denominators and problems of stability of motion in classical and celestial mechanics. Russ. Math. Surv. **18**(6), 85–191 (1963)

Arnold, V.I.: Instability of dynamical systems with many degrees of freedom. Dokl. Akad. Nauk SSSR **156**, 9–12 (1964)

Bonetto, F., Gallavotti, G., Gentile, G., Mastropietro, V.: Lindstedt series, ultraviolet divergences and Moser's theorem. Ann. Sc. Norm. Super Pisa. Cl. Sci. IV Ser. **26**(3), 545–593 (1998)

Celletti, A., Chierchia, L.: Construction of stable periodic orbits for the spin-orbit problem of celestial mechanics. Regul. Chaotic Dyn. **3**(3), 107–121 (1998). (**J. Moser at 70 (Russian)**)

Chierchia, L., Gallavotti, G.: Drift and diffusion in phase space. Ann. Inst. H. Poincaré Phys. Théor. **60**(1), 144 (1994)

Chierchia, L., Pinzari, G.: Properly–degenerate KAM theory (following V.I. Arnold). Discrete Contin. Dyn. Syst. Ser. S **3**(4), 545–578 (2010)

Chierchia, L., Pinzari, G.: The planetary N-body problem: symplectic foliation, reductions and invariant tori. Invent. Math. **186**(1), 1–77 (2011)

Chierchia, L., Procesi, M.: Kolmogorov–Arnold–Moser (KAM) Theory for Finite and Infinite Dimensional Systems, pp. 1–45. Springer, Berlin (2019)

Clarke, A., Féjoz, J., Guardia, M..: Why are inner planets not inclided? arXiv: 2210.11311v1 (2022)

Delshams, A., Kaloshin, V., de la Rosa, A., Seara, T.M.: Global instability in the restricted planar elliptic three body problem. Commun. Math. Phys. **366**(3), 1173–1228 (2019)

Deprit, A.: Elimination of the nodes in problems of n bodies. Celestial Mech. **30**(2), 181–195 (1983)

Féjoz, J.: Démonstration du 'théorème d'Arnold' sur la stabilité du système planétaire (d'après Herman). Ergod. Theory Dyn. Syst. **24**(5), 1521–1582 (2004)

Féjoz, J., Guardia, M., Kaloshin, V., Roldan, P.: Kirkwood gaps and diffusion along mean motion resonances in the restricted planar three body problem. J. Eur. Math. Soc. **18**, 2315 (2014)

Gallavotti, G.: Quasi-integrable mechanical systems. In: Phénomènes Critiques. Systèmes aléatoires, Théories de Jauge, Part I, II (Les Houches, 1984), pp. 539–624. North-Holland, Amsterdam (1986)

Gallavotti, G.: Twistless KAM tori. Commun. Math. Phys. **164**(1), 145–156 (1994)

Gentile, G., Gallavotti, G.: Majorant convergence for twistless KAM tori. Ergod. Theory Dyn. Syst. **15**(5), 587–869 (1995)

Guzzo, M., Efthymiopoulos, C., Paez, R.I.: Semi-analytic computations of the speed of Arnold diffusion along single resonances in a priori stable Hamiltonian systems. J. Nonlinear Sci. **30**(3), 851–901 (2020)

Jacobi, C.G.J.: Sur l'élimination des noeuds dans le problème des trois corps. Astron Nachr **1**, 81–102 (1842)

Laskar, J., Robutel, P.: Stability of the planetary three-body problem. I .Expansion of the planetary Hamiltonian. Celestial Mech. Dyn. Astronom. **62**(3), 193–217 (1995)

Pinzari, G.: Perihelia reduction and global Kolmogorov tori in the planetary problem. Mem. Amer. Math. Soc., vol. 255, No. 1218 (2018)

Pinzari, G.: On the co-existence of maximal and whiskered tori in the planetary three-body problem. J. Math. Phys. **59**(5), 052701 (2018)

Pöschel, J.: Nekhoroshev estimates for quasi-convex Hamiltonian systems. Math. Z. **213**(2), 187–216 (1993)

Radau, R.: Sur une transformation des équations différentielles de la dynamique. Ann. Sci. Ec. Norm. Sup. **5**, 311–375 (1868)

Rüssmann, H.: Invariant tori in non-degenerate nearly integrable Hamiltonian systems. Regul. Chaotic Dyn. **6**(2), 119–204 (2001)