JOURNAL OF MODERN DYNAMICS
VOLUME 5, NO. 4, 2011, 623–664

doi: 10.3934/jmd.2011.5.623

PLANETARY BIRKHOFF NORMAL FORMS

LUIGI CHIERCHIA AND GABRIELLA PINZARI
(Communicated by Anatole Katok)

ABSTRACT. Birkhoff normal forms for the (secular) planetary problem are investigated. Existence and uniqueness is discussed and it is shown that the classical Poincaré variables and the RPS–variables (introduced in [6]), after a trivial lift, lead to the same Birkhoff normal form; as a corollary the Birkhoff normal form (in Poincaré variables) is degenerate at all orders (answering a question of M. Herman). Non-degenerate Birkhoff normal forms for partially and totally reduced cases are provided and an application to long–time stability of secular action variables (eccentricities and inclinations) is discussed.

CONTENTS

1. INTRODUCTION

Consider the planetary $(1+n)$-body problem, i.e., the motions of $1+n$ point-masses, interacting only through gravity, with one body ("the Sun") having a much larger mass than the other ones ("the planets"). A fundamental feature

Received September 23, 2010; revised October 17, 2011.
2000 *Mathematics Subject Classification:* 70F10, 70K45, 70F15, 70F07, 70E55, 34C20, 34D10.
Key words and phrases: Planetary system, Birkhoff normal form, Birkhoff invariants, N-body problem, Long-time stability.
Partially supported by European Research Council under FP7 project "New connections between dynamical systems and Hamiltonian PDEs with small divisors phenomena" and by Prin 2009 project "Critical Point Theory and Perturbative Methods for Nonlinear Differential Equations."

of this Hamiltonian system (for negative decoupled energies) is the separation between fast degrees of freedom, roughly describing the relative distances of the planets, and the slow (or "secular") degrees of freedom, describing the relative inclinations and eccentricities (of the osculating Keplerian ellipses). A second remarkable feature of the planetary system is that the secular Hamiltonian has (in suitable "Cartesian variables") an elliptic equilibrium around zero inclinations and eccentricities. Birkhoff normal form theory[1] comes, therefore, naturally in. Such theory yields, in particular, information on the secular frequencies (first order Birkhoff invariants) and on the "torsion" (or "twist") of the secular variables (the determinant of the second-order Birkhoff invariants). Indeed, secular Birkhoff invariants are intimately related to the existence of maximal and lower-dimensional KAM tori[2], or, as we will show below (§ 6), one can infer long-time stability for the "secular actions" (essentially, eccentricities and mutual inclinations).

A natural question is therefore the construction of Birkhoff normal forms for the secular planetary Hamiltonian. Already Arnold in 1963 realized that this is not a straightforward task in view of *secular resonances*, i.e., rational relations among the first-order Birkhoff invariants holding identically on the phase space. Incidentally, Arnold was aware of the so-called rotational resonance (the vanishing of one of the "vertical" first-order Birkhoff invariants) but did not realize the presence of a second resonance of order $2n-1$ discovered by M. Herman (compare [10] and [1]). These resonances, apart from being an obstacle for the construction of Birkhoff normal forms, constituted also a problem for the application of KAM theory. This problem was overcome, in full generality, only in 2004 [10] using a weaker KAM theory involving only information on the first order Birkhoff invariants, waving the check of Kolmogorov's nondegeneracy (related to full torsion[3]); for a short description of the main ideas involved, see [6, Remark 11.1, (iii)].

In particular the question of the torsion of the secular Hamiltonian remained open. M. Herman investigated such question thoroughly using Poincaré variables [11] but declared not to know if some of the second-order Birkhoff invariants are zero even in the $n=2$ case (compare the Remark towards the end of p. 24 in [11]).

A different point of view is taken up in [6], where a new set of variables, called RPS ("Regularized Planetary Symplectic") variables, is introduced in order to study the symplectic structure of the phase space of the planetary system. Such variables are based on Deprit's action-angles variables ([8, 5]), which may be used for a symplectic reduction lowering by one the number of degrees of freedom. A further reduction is possible (at the expense of introducing a new singularity) leading to a totally reduced phase space, compare [6, §9] and § 5.1

[1] See [12] for generalities and Appendix A for the theory for rotational invariant systems.

[2] Compare [2, 16, 10, 7, 6] for maximal tori and [9, 3, 6] for lower-dimensional elliptic tori.

[3] That is, the nonvanishing of the determinant of the matrix formed by the second order Birkhoff invariants.

below. On the reduced phase spaces, one can construct Birkhoff normal forms ([6, Sect 7 and 9]; § 2, § 5.1 below). Following such strategy one can show that the matrix of second-order Birkhoff invariants (for the reduced system) is nondegenerate and prove full torsion. In particular, it is then possible to construct a large measure set of maximal nondegenerate KAM tori ([6, §11]).

In this paper we consider and clarify various aspects of Birkhoff normal forms for the planetary system. In particular we analyze the connection between the Birkhoff normal form in the classical setting (Poincaré variables) and in the new setting of [6]. It turns out that after lifting in a trivial way the RPS variables to the full-dimensional phase space, such variables and the Poincaré variables are related in a simple way, namely, through a symplectic map which leaves the action variables Λ (conjugate to the mean anomalies) fixed and so that the correspondence between the respective Cartesian variables is close to the identity map (and independent of the fast angles); compare Theorem 3.2 below. Since, up to such class of symplectic maps, the Birkhoff normal form is unique, one sees that the Birkhoff normal form in Poincaré variables is *degenerate at all orders*, answering negatively the question of M. Herman; see Theorem 2.1 below. We mention also that the construction of Birkhoff normal form for rotational invariant Hamiltonian (such as the secular planetary Hamiltonian) is simpler than the standard construction: in fact, one needs to assume nonresonance of the first-order Birkhoff invariant for those Taylor modes $k \neq 0$ such that $\sum_i k_i = 0$ (and not just $k \neq 0$); compare Appendix A. By this remark one sees that the secular resonances (both the rotational *and* the Herman resonance) do not really affect the construction of Birkhoff normal forms.

In § 5.1 we discuss the construction, up to any order, of the Birkhoff normal forms in the totally reduced setting (generalizing Proposition 10.1 in [6]) and, for completeness, we consider (§ 5.2) the planar planetary problem (in which case the Poincaré and the RPS variables coincide) and, after introducing a (total) symplectic reduction, we discuss Birkhoff normal forms in such reduced setting, comparing, in particular, with the detailed analysis in [11].

Finally, in § 6, we use the results of § 5.1 in order to prove that, in suitable open nonresonant phase space regions of relatively large Liouville measure, the eccentricities and mutual inclinations remain small and close to their initial values for times which are proportional to any prefixed inverse power of the distance from the equilibrium point (zero inclinations and zero eccentricities): such result is somewhat complementary to Nehoroševʼs original result [13], where exponential stability of the semi major axes was established, but no information on possible relatively large variation of the secular action was given.

2. PLANETARY BIRKHOFF NORMAL FORM

After the symplectic reduction of the linear momentum, the $(1+n)$-body problem with masses m_0, μm_1, $\cdots$, μm_n $(0 < \mu \ll 1)$ is governed by the $3n$-degrees-of-freedom Hamiltonian

$$
\begin{aligned}
\mathcal{H}_{\rm plt} &= \sum_{1\le i\le n}\left(\frac{|y^{(i)}|^2}{2M_i}-\frac{M_i\bar m_i}{|x^{(i)}|}\right)+\mu\sum_{1\le i<j\le n}\left(\frac{y^{(i)}\cdot y^{(j)}}{m_0}-\frac{m_im_j}{|x^{(i)}-x^{(j)}|}\right)\\
&=: h_{\rm plt}+\mu f_{\rm plt} \qquad (2.1)
\end{aligned}
$$

where $x^{(i)}$ represent the difference between the position of the $i^{\rm th}$ planet and the position of the Sun, $y^{(i)}$ are the associated symplectic momenta rescaled by μ, $x\cdot y=\sum_{1\le i\le 3}x_iy_i$ and $|x|:=(x\cdot x)^{1/2}$ denote, respectively, the standard inner product in $\mathbb{R}^3$ and the Euclidean norm;

$$
M_i:=\frac{m_0m_i}{m_0+\mu m_i}\ ,\quad \bar m_i:=m_0+\mu m_i\ . \tag{2.2}
$$

The phase space is the "collisionless" domain of $\mathbb{R}^{3n}\times\mathbb{R}^{3n}$

$$
\Big\{(y,x)=\big((y^{(1)},\dots,y^{(n)}),(x^{(1)},\dots,x^{(n)})\big)\ \text{s.t. } 0\ne x^{(i)}\ne x^{(j)}\ ,\ \forall\ i\ne j\Big\}\ ,
$$

endowed with the standard form

$$
\omega=\sum_{i=1}^n dy^{(i)}\wedge dx^{(i)}=\sum_{i=1}^n\sum_{j=1}^3 dy_j^{(i)}\wedge dx_j^{(i)}
$$

where $y_j^{(i)}$, $x_j^{(i)}$ denote the $j^{\rm th}$ component of $y^{(i)}$, $x^{(i)}$.

When $\mu=0$, the Hamiltonian (2.1) is integrable: its unperturbed limiting value $h_{\rm plt}$ is the sum of the Hamiltonians

$$
h_{\rm plt}^{(i)}=\frac{|y^{(i)}|^2}{2M_i}-\frac{M_i\bar m_i}{|x^{(i)}|}\ ,\qquad (y^{(i)},x^{(i)})\in\mathbb{R}^3\times\mathbb{R}^3_*:=\mathbb{R}^3\times(\mathbb{R}^3\setminus\{0\}) \tag{2.3}
$$

corresponding to uncoupled Two-Body Newtonian interactions.

In Poincaré coordinates – which will be reviewed in the next section – the Hamiltonian (2.1) takes the form

$$
\mathcal{H}_{\rm P}(\Lambda,\lambda,{\rm z})=h_{\rm K}(\Lambda)+\mu f_{\rm P}(\Lambda,\lambda,{\rm z})\ ,\quad {\rm z}:=(\eta,{\rm p},\xi,{\rm q})\in\mathbb{R}^{4n} \tag{2.4}
$$

where $(\Lambda,\lambda)\in\mathbb{R}^n\times\mathbb{T}^n$; the "Kepler" unperturbed term $h_{\rm K}$, coming from $h_{\rm plt}$ in (2.1), becomes

$$
h_{\rm K}:=\sum_{i=1}^n h_{\rm K}^{(i)}(\Lambda)=-\sum_{i=1}^n\frac{\bar m_i^2M_i^3}{2\Lambda_i^2}\ .
$$

Because of rotation (with respect the $k^{(3)}$-axis) and reflection (with respect to the coordinate planes) invariance of the Hamiltonian (2.1), the perturbation $f_{\rm P}$ in (2.4) satisfies well-known symmetry relations called *d'Alembert rules*, see (3.19)–(3.24) below. By such symmetries, in particular, the averaged perturbation

$$
f_{\rm P}^{\rm av}(\Lambda,{\rm z}):=\frac{1}{(2\pi)^n}\int_{\mathbb{T}^n}f_{\rm P}(\Lambda,\lambda,{\rm z})\,d\lambda \tag{2.5}
$$

is even around the origin $z = 0$ and its expansion in powers of z has the form[4]

$$f_P^{av} = C_0(\Lambda) + \mathcal{Q}_h(\Lambda) \cdot \frac{\eta^2 + \xi^2}{2} + \mathcal{Q}_v(\Lambda) \cdot \frac{p^2 + q^2}{2} + O(|z|^4) , \tag{2.6}$$

where $\mathcal{Q}_h$, $\mathcal{Q}_v$ are suitable quadratic forms. The explicit expression of such quadratic forms can be found, e.g., in [10, (36), (37)] (revised version).

By such expansion, the (secular) origin $z = 0$ is an *elliptic equilibrium* for f_P^{av} and corresponds to coplanar and cocircular motions. It is therefore natural to put (2.6) into Birkhoff normal form in a small neighborhood of the secular origin; see, e.g., [12] for general information on Birkhoff normal forms and Appendix A for Birkhoff theory for rotational invariant Hamiltonian systems.

As a preliminary step, one can diagonalize (2.6), i.e., find a symplectic transformation

$$\tilde{\Phi}_P : (\Lambda, \tilde{\lambda}, \tilde{z}) \in \tilde{\mathcal{M}}_P^{6n} \to (\Lambda, \lambda, z) \in \mathcal{M}_P^{6n} := \tilde{\Phi}_P(\tilde{\mathcal{M}}_P^{6n}) \tag{2.7}$$

(the domain $\tilde{\mathcal{M}}_P^{6n}$ will be specified in (2.12) below) defined by $\Lambda \to \Lambda$ and

$$\lambda = \tilde{\lambda} + \varphi(\Lambda, \tilde{z}),\ \eta = \rho_h(\Lambda)\tilde{\eta},\ \xi = \rho_h(\Lambda)\tilde{\xi},\ p = \rho_v(\Lambda)\tilde{p},\ q = \rho_v(\Lambda)\tilde{q} , \tag{2.8}$$

with ρ_h, $\rho_v \in SO(n)$ diagonalizing $\mathcal{Q}_h$, $\mathcal{Q}_v$. In this way, (2.6) takes the form

$$\tilde{\mathcal{H}}_P(\Lambda, \tilde{\lambda}, \tilde{z}) = \mathcal{H}_P \circ \tilde{\Phi}_P = h_K(\Lambda) + \mu \tilde{f}(\Lambda, \tilde{\lambda}, \tilde{z}) , \tag{2.9}$$

with the average over $\tilde{\lambda}$ of $\tilde{f}^{av}$ given by

$$\tilde{f}^{av}(\Lambda, \tilde{z}) = C_0(\Lambda) + \sum_{i=1}^n \sigma_i \frac{\tilde{\eta}_i^2 + \tilde{\xi}_i^2}{2} + \sum_{i=1}^n \varsigma_i \frac{\tilde{p}_i^2 + \tilde{q}_i^2}{2} + O(|\tilde{z}|^4), \quad \tilde{z} = (\tilde{\eta}, \tilde{\xi}, \tilde{p}, \tilde{q}). \tag{2.10}$$

The $2n$ real vector $\Omega := (\sigma, \varsigma) = (\sigma_1, \cdots, \sigma_n, \varsigma_1, \cdots, \varsigma_n)$ is formed by the eigenvalues of the matrices $\mathcal{Q}_h$ and $\mathcal{Q}_v$ in (2.6) and are called the *first-order Birkhoff invariants.*

It turns out that such invariants satisfy identically the following two *secular resonances*

$$\sum_{i=1}^n (\sigma_i + \varsigma_i) = 0 , \qquad \varsigma_n = 0 .$$

Such resonances strongly violate the usual nondegeneracy assumptions that are needed for the direct construction of Birkhoff normal forms.

The first resonance, discovered by M. Herman, is still quite mysterious (see, however, [1]), while the second resonance is related to the existence of two noncommuting integrals, given by the horizontal components C_1 and C_2 of the total angular momentum $C := \sum_{i=1}^n x^{(i)} \times y^{(i)}$ of the system (compare [2]).

Actually, the effect of rotation invariance is deeper: the vanishing of the eigenvalue ς_n is just "the first order" of a "rotational" proper degeneracy, as explained in the following theorem, which will be proved in § 4. Let $w := (u, v) =$

[4] $\mathcal{Q} \cdot u^2$ denotes the 2-indices contraction $\sum_{i,j} \mathcal{Q}_{ij} u_i u_j$ ($\mathcal{Q}_{ij}$, u_i denoting the entries of $\mathcal{Q}$, u).

$(u_1, \cdots, u_{2n}, v_1, \cdots, v_{2n})$, $\bar{w} := (u_1, \cdots, u_{2n-1}, v_1, \cdots, v_{2n-1})$ and

$$G(\Lambda, \bar{w}) := \sum_{i=1}^{n} \Lambda_i - \frac{1}{2} \sum_{i=1}^{2n-1} (u_i^2 + v_i^2) \ . \tag{2.11}$$

THEOREM 2.1. *For any $s \in \mathbb{N}$, there exists $\varepsilon > 0$, an open set $\mathscr{A} \subseteq \{a_1 < \cdots < a_n\}$ such that, if*

$$\mathscr{M}_{\rm B}^{6n} := \mathscr{A} \times \mathbb{T}^n \times B_\varepsilon^{4n-2} \times B^2_{2\sqrt{G}} \ ,$$

one can construct a symplectic map ("Birkhoff transformation"),

$$\Phi_{\rm B} \colon (\Lambda, l, w) \in \mathscr{M}_{\rm B}^{6n} \to (\Lambda, \tilde{\lambda}, \tilde{z}) \in \tilde{\mathscr{M}}_{\rm P}^{6n} := \Phi_{\rm B}(\mathscr{M}_{\rm B}^{6n}) \tag{2.12}$$

with the following properties. The pullback of the Hamiltonian (2.9) takes the form

$$\mathscr{H}_{\rm B}(\Lambda, l, w) := \tilde{\mathscr{H}}_{\rm P} \circ \Phi_{\rm B} = h_{\rm K}(\Lambda) + \mu f_{\rm B}(\Lambda, l, w) \tag{2.13}$$

where the average $f_{\rm B}^{\rm av}(\Lambda, w) := \dfrac{1}{(2\pi)^n} \displaystyle\int_{\mathbb{T}^n} f_{\rm B}\, dl$ is in Birkhoff normal form of order s:

$$f_{\rm B}^{\rm av}(\Lambda, w) = C_0 + \Omega \cdot r + {\rm P}_s(r) + {\rm O}(|w|^{2s+2}) \quad w := (u, v) \quad r_i := \frac{u_i^2 + v_i^2}{2} \ , \tag{2.14}$$

${\rm P}_s$ being homogeneous polynomial in r of order s, parameterized by Λ. Such normal form is unique up to symplectic transformations Φ which leave the Λ's fixed and with the $\tilde{z}$-projection independent of l and close to the identity in w, i.e.,

$$\Pi_{\tilde{z}} \Phi = w + {\rm O}(|w|^2) \ . \tag{2.15}$$

Furthermore, the normal form (2.13)–(2.14) is "infinitely degenerate", in the sense that $\mathscr{H}_{\rm B}$ does not depend on (u_{2n}, v_{2n}). In particular, there exists a unique polynomial $\bar{\rm P}_s \colon \mathbb{R}^{2n-1} \to \mathbb{R}$ (parameterized by Λ) such that

$$ {\rm P}_s(r) = \bar{\rm P}_s(\bar{r}) \quad \textit{where} \quad \bar{r} := (r_1, \cdots, r_{2n-1}) \ . \tag{2.16}$$

REMARK 2.2.

(i) Note that the w-projection of $\mathscr{M}_{\rm B}^{6n}$ corresponds to a neighborhood of $w = 0$, which is small only in the $4n-2$ components of w, while it is large (maximal) in the remaining 2 components (compare Appendix B for the natural radius $2\sqrt{G}$ in the variables (u_{2n}, q_{2n})). Indeed, to construct the normal form, by rotation invariance, it is not necessary to assume that *all* inclinations are small, but one can take the *mutual* inclinations to be small. This corresponds to consider $2n-1$ secular degrees of freedom (roughly, corresponding to n couples of eccentricities–perihelia and $n-1$ couples of inclinations–nodes) instead of $2n$. The *overall* inclination–node of the system (corresponding to the remaining 2 secular variables) is allowed to vary globally.

(ii) Theorem 2.1 depends strongly on the rotational invariance of the Hamiltonian (2.1), that is, on the fact that such Hamiltonian commutes with the three components of the angular momentum C. To exploit explicitly such invariance, we shall use a set of symplectic variables ("RPS variables"), introduced in [6] (in order to describe the symplectic structure of the planetary N-body problem and to check KAM nondegeneracies).

(iii) The RPS variables are obtained as a symplectic regularization of a set of action-angle variables, introduced by Deprit in 1983 ([8, 5]), which generalize to an arbitrary number n of planets the classical Jacobi's reduction of the nodes ($n = 2$). The remarkable property of the Deprit's variables is that there appear a conjugate couple (C_3 and ζ below) plus an action variable G which are integrals. Thus, the conjugate integrals are also cyclic and are responsible for the proper degeneracy of the planetary Hamiltonian. Furthermore, the RPS variables have a cyclic couple ((p_n, q_n) below), which foliates the phase space into symplectic leaves (the sets $\mathcal{M}^{6n-2}_{(p_n^\star, q_n^\star)}$ in (3.13) below), on which the planetary Hamiltonian keeps the same form. So, the construction of the "non degenerate part" of the normal form can be made up to any order (and is the same) on each leaf [6]. In particular, the even order of the remainder in (2.14) is due to invariance by rotations around the C-axis of the system. Finally, we prove that such normal form can be uniquely lifted to the degenerate normal form (2.14)–(2.16) on the phase space $\mathcal{M}^{6n}_{\mathrm{P}}$ in (2.7).

The proof is based on the remarkable link between RPS and Poincaré variables, described in the following section (see Theorem 3.2).

3. POINCARÉ AND RPS VARIABLES

In this section we first recall the definitions of the Poincaré and RPS variables[5] and then discuss how they are related. Recall that the Poincaré variables have been introduced to regularize around zero eccentricities and inclinations the Delaunay action-angle variables. Analogously, the RPS variables have been introduced to regularize around zero eccentricities and inclinations the Deprit action-angle variables.

- Fix $2n$ positive "mass parameters[6]" M_i, $\bar{m}_i$ and consider the *two-body* Hamiltonians $h_i(y^{(i)}, x^{(i)}) := h^{(i)}_{\mathrm{plt}}$ as in (2.3). Assume that $h_i(y^{(i)}, x^{(i)}) < 0$ so that the Hamiltonian flow $\phi^t_{h_i}(y^{(i)}, x^{(i)})$ evolves on a Keplerian ellipse $\mathcal{E}_i$ and assume that the eccentricity $e_i \in (0,1)$. Let a_i, P_i denote, respectively, the *semi major axis* and the *perihelion* of $\mathcal{E}_i$. Let $\mathrm{C}^{(i)}$ denote the i^{th} angular momentum $\mathrm{C}^{(i)} = x^{(i)} \times y^{(i)}$.
  - To define Delaunay variables, one needs the "Delaunay nodes"

$$\bar{\nu}_i = k^{(3)} \times \mathrm{C}^{(i)} \quad 1 \le i \le n\,, \tag{3.1}$$

[5] For full details, see [10], and references therein, and [6].

[6] The RPS variables will depend upon these mass parameters, which, in the planetary case, will obviously coincide with (2.2).

where $(k^{(1)}, k^{(2)}, k^{(3)})$ is the standard orthonormal basis in $\mathbb{R}^3$.

– To define Deprit variables, consider the "partial angular momenta"

$$S^{(i)} = \sum_{j=1}^{i} \mathrm{C}^{(j)}\ , \qquad S^{(n)} = \sum_{j=1}^{n} \mathrm{C}^{(j)} =: \mathrm{C}\ ; \tag{3.2}$$

(note that C is the total angular momentum of the system) and define the "Deprit nodes"

$$\begin{cases} \nu_{i+1} = S^{(i+1)} \times \mathrm{C}^{(i+1)}\ , & 1 \le i \le n-1 \\ \nu_1 = \nu_2 \\ \nu_{n+1} = k^{(3)} \times \mathrm{C} =: \bar{\nu}\ . \end{cases} \tag{3.3}$$

For $u, v \in \mathbb{R}^3$ lying in the plane orthogonal to a vector w, let $\alpha_w(u, v)$ denote the positively oriented angle (mod 2π) between u and v (orientation follows the "right hand rule").

- The classical *Delaunay action-angle variables* $(\Lambda, \Gamma, \Theta, \ell, \mathrm{g}, \theta)$ are defined as

$$\begin{cases} \Lambda_i = M_i \sqrt{\bar{m}_i a_i} \\ \ell_i = \text{mean anomaly of } x^{(i)} \text{ on } \mathscr{E}_i \end{cases} \quad \begin{cases} \Gamma_i = |\mathrm{C}^{(i)}| = \Lambda_i \sqrt{1 - e_i^2} \\ \mathrm{g}_i = \alpha_{\mathrm{C}^{(i)}}(\bar{\nu}_i, P_i) \end{cases} \quad \begin{cases} \Theta_i = \mathrm{C}^{(i)} \cdot k^{(3)} \\ \theta_i = \alpha_{k^{(3)}}(k^{(1)}, \bar{\nu}_i) \end{cases} \tag{3.4}$$

- The *Deprit action-angle variables* $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ are defined as follows. The variables Λ, Γ and ℓ are in common with the Delaunay variables (3.4), while

$$\gamma_i = \alpha_{\mathrm{C}^{(i)}}(\nu_i, P_i) \quad \Psi_i = \begin{cases} |S^{(i+1)}| & 1 \le i \le n-1 \\ \mathrm{C}_3 = \mathrm{C} \cdot k^{(3)} & i = n \end{cases}$$
$$\psi_i = \begin{cases} \alpha_{S^{(i+1)}}(\nu_{i+2}, \nu_{i+1}) & 1 \le i \le n-1 \\ \zeta = \alpha_{k^{(3)}}(k^{(1)}, \bar{\nu}) & i = n. \end{cases} \tag{3.5}$$

Define also $G = |\mathrm{C}| = |S^{(n)}|$.

Note that:

- Delaunay's variables are defined on an open set of full measure $\mathscr{P}^{6n}_{\mathrm{Del}^\star}$ of the Cartesian phase space $\mathscr{P}^{6n} = \mathbb{R}^{3n} \times \mathbb{R}^{3n}_*$, namely, on the set where $e_i \in (0,1)$ and the nodes $\bar{\nu}_i$ in (3.1) are well defined.
- Deprit's variables are defined on an open set of full measure $\mathscr{P}^{6n}_{\mathrm{Dep}^\star}$ of $\mathscr{P}^{6n}$ where $e_i \in (0,1)$ and the nodes ν_i in (3.3) are well defined.

- On $\mathscr{P}^{6n}_{\mathrm{Del}^\star}$ and $\mathscr{P}^{6n}_{\mathrm{Dep}^\star}$, the "Delaunay inclinations" i_i and the "Deprit inclinations" ι_i, defined through the relations

$$\cos \mathrm{i}_i = \frac{\mathrm{C}^{(i)} \cdot k^{(3)}}{|\mathrm{C}^{(i)}|}\ , \quad \cos \iota_i = \begin{cases} \dfrac{\mathrm{C}^{(i+1)} \cdot S^{(i+1)}}{|\mathrm{C}^{(i+1)}||S^{(i+1)}|} & 1 \le i \le n-1 \\ \dfrac{\mathrm{C} \cdot k^{(3)}}{|\mathrm{C}|} & i = n \end{cases}$$

  are well defined and we choose the branch of $\cos^{-1}$ so that i_i, $\iota_i \in (0,\pi)$.

Finally:

- The *Poincaré variables* are given by $(\Lambda, \lambda, \mathrm{z}) = (\Lambda, \lambda, \eta, \xi, \mathrm{p}, \mathrm{q})$, with the Λ's as in (3.4) and

$$\lambda_i = \ell_i + \mathrm{g}_i + \theta_i \tag{3.6}$$
$$\begin{cases} \eta_i = \sqrt{2(\Lambda_i - \Gamma_i)}\ \cos(\theta_i + \mathrm{g}_i) \\ \xi_i = -\sqrt{2(\Lambda_i - \Gamma_i)}\ \sin(\theta_i + \mathrm{g}_i) \end{cases}$$
$$\begin{cases} \mathrm{p}_i = \sqrt{2(\Gamma_i - \Theta_i)}\ \cos\theta_i \\ \mathrm{q}_i = -\sqrt{2(\Gamma_i - \Theta_i)}\ \sin\theta_i \end{cases}$$

- The RPS variables are given by $(\Lambda, \lambda, z) = (\Lambda, \lambda, \eta, \xi, p, q)$ with (again) the Λ's as in (3.4) and

$$\lambda_i = \ell_i + \gamma_i + \psi^n_{i-1} \tag{3.7}$$
$$\begin{cases} \eta_i = \sqrt{2(\Lambda_i - \Gamma_i)}\ \cos\big(\gamma_i + \psi^n_{i-1}\big) \\ \xi_i = -\sqrt{2(\Lambda_i - \Gamma_i)}\ \sin\big(\gamma_i + \psi^n_{i-1}\big) \end{cases}$$
$$\begin{cases} p_i = \sqrt{2(\Gamma_{i+1} + \Psi_{i-1} - \Psi_i)}\ \cos\psi^n_i \\ q_i = -\sqrt{2(\Gamma_{i+1} + \Psi_{i-1} - \Psi_i)}\ \sin\psi^n_i \end{cases}$$

  where

$$\Psi_0 = \Gamma_1\ , \quad \Gamma_{n+1} = 0\ , \quad \psi_0 = 0\ , \quad \psi^n_i = \sum_{i \le j \le n} \psi_j\ . \tag{3.8}$$

REMARK 3.1. From the definitions (3.7)–(3.8) it follows that the variables

$$\begin{cases} p_n = \sqrt{2(\Psi_{n-1} - \Psi_n)} \cos\psi_n = \sqrt{2(G - \mathrm{C}_3)} \cos\zeta \\ q_n = -\sqrt{2(\Psi_{n-1} - \Psi_n)} \sin\psi_n = -\sqrt{2(G - \mathrm{C}_3)} \sin\zeta \end{cases} \tag{3.9}$$

are defined only in terms of the integral C. Thus, they are integrals (hence, cyclic) in Hamiltonian systems which commute with the three components of the angular momentum C (or, equivalently, in systems which are invariant under rotations).

Let ϕ_{P} and ϕ_{RPS} denote the maps

$$\phi_{\mathrm{P}}\colon (y, x) \to (\Lambda, \lambda, \mathrm{z})\ , \qquad \phi_{\mathrm{RPS}}\colon (y, x) \to (\Lambda, \lambda, z)\ . \tag{3.10}$$

The main point of this procedure is that:

- *The map $\phi_{\rm P}$ can be extended to an analytic symplectic diffeomorphism on the set $\mathscr{P}^{6n}_{\rm Del}$ which is defined as $\mathscr{P}^{6n}_{{\rm Del}^\star}$, but with e_i and i_i allowed to be zero.*
- *The map $\phi_{\rm RPS}$ can be extended to an analytic symplectic diffeomorphism on the set $\mathscr{P}^{6n}_{\rm Dep}$ which is defined as $\mathscr{P}^{6n}_{{\rm Dep}^\star}$, but with e_i and ι_i allowed to be zero.*

The image sets $\mathscr{M}^{6n}_{\max,{\rm P}} := \phi_{\rm P}(\mathscr{P}^{6n}_{\rm Del})$ and $\mathscr{M}^{6n}_{\max,{\rm RPS}} := \phi_{\rm RPS}(\mathscr{P}^{6n}_{\rm Dep})$ are defined by elementary inequalities following from the definitions (3.6) and (3.7) (details in Appendix B). Note in particular that

- $e_i = 0$ corresponds to the Poincaré coordinates $\upeta_i = 0 = \upxi_i$ and the RPS coordinates $\eta_i = 0 = \xi_i$;
- $\mathrm{i}_i = 0$ corresponds to the Poincaré coordinates $\mathrm{p}_i = 0 = \mathrm{q}_i$;
- $\iota_i = 0$ corresponds to the the RPS coordinates $p_i = 0 = q_i$. In particular $p_n = 0 = q_n$ corresponds to the angular momentum C being parallel to the $k^{(3)}$-axis.
- Let $\bar z$ denote the set of variables

$$\bar z := (\eta, \xi, \bar p, \bar q) := \big((\eta_1,\dots,\eta_n),(\xi_1,\dots,\xi_n),(p_1,\dots,p_{n-1}),(q_1,\dots,q_{n-1})\big)\,. \tag{3.11}$$

  (roughly, $\bar z$ are related to eccentricities–perihelia, and *mutual* inclinations–nodes of the instantaneous ellipses $\mathscr{E}_i$). Then $\mathscr{M}^{6n}_{\max,{\rm RPS}}$ can be written as

$$\mathscr{M}^{6n}_{\max,{\rm RPS}} := \phi_{\rm RPS}(\mathscr{P}^{6n}_{\rm Dep}) = \{(\Lambda,\lambda,\bar z) \in \mathscr{M}^{6n-2}_{\max}\ , \quad p_n^2 + q_n^2 < 4G(\Lambda,\bar z)\} \tag{3.12}$$

  where $G(\Lambda,\bar z)$ is just the length of the total angular momentum expressed in RPS variables as given in (2.11) and $\mathscr{M}^{6n-2}_{\max}$ is a given subset of $\mathbb{R}^n_+ \times \mathbb{T}^n \times \mathbb{R}^{4n-2}$ (compare the end of Appendix B).
- We have already observed that for rotation-invariant systems the variables (p_n, q_n) are cyclic. In this case, the phase space $\mathscr{M}^{6n}_{\max,{\rm RPS}}$ is foliated into symplectic leaves

$$\mathscr{M}^{6n-2}_{(p_n^\star,q_n^\star)} := \phi_{\rm RPS}(\mathscr{P}^{6n}_{\rm Dep}) = \{(\Lambda,\lambda,z) \in \mathscr{M}^{6n}_{\max,{\rm RPS}} : p_n = p_n^\star,\ q_n = q_n^\star\}\,. \tag{3.13}$$

  In the next section, for the application to the planetary problem, we shall substitute the set $\mathscr{M}^{6n-2}_{\max}$ in the definition (3.12) of $\mathscr{M}^{6n}_{\max,{\rm RPS}}$ with a smaller set $\mathscr{M}^{6n-2}$: compare (4.2) below.

Consider the common domain of the maps $\phi_{\rm P}$ and $\phi_{\rm RPS}$ in (3.10), i.e., the set $\mathscr{P}^{6n}_{\rm Del} \cap \mathscr{P}^{6n}_{\rm Dep}$. In particular, on such set, $0 \le e_i < 1$, $0 \le \mathrm{i}_i < \pi$, $0 \le \iota_i < \pi$. On the $\phi_{\rm RPS}$-image of such domain consider the symplectic map

$$\phi^{\rm RPS}_{\rm P} : (\Lambda,\lambda,z) \to (\Lambda,\uplambda,\mathrm{z}) := \phi_{\rm P} \circ \phi^{-1}_{\rm RPS} \tag{3.14}$$

which maps the RPS variables onto the Poincaré variables. Such a map has a particularly simple structure:

THEOREM 3.2. *The symplectic map $\phi^{\rm RPS}_{\rm P}$ in (3.14) has the form*

$$\uplambda = \lambda + \varphi(\Lambda, z) \qquad \mathrm{z} = \mathscr{Z}(\Lambda, z) \tag{3.15}$$

where $\varphi(\Lambda,0)=0$ *and, for any fixed* Λ*, the map* $\mathcal{Z}(\Lambda,\cdot)$ *is 1:1, symplectic*[7] *and its projections satisfy, for a suitable* $\mathcal{V}=\mathcal{V}(\Lambda)\in \mathrm{SO}(n)$*, with* $\mathrm{O}_3=\mathrm{O}(|z|^3)$*,*

$$\Pi_\eta \mathcal{Z}=\eta+\mathrm{O}_3\ ,\ \Pi_\xi \mathcal{Z}=\xi+\mathrm{O}_3\ ,\ \Pi_{\mathrm{p}} \mathcal{Z}=\mathcal{V} p+\mathrm{O}_3\ ,\ \Pi_{\mathrm{q}} \mathcal{Z}=\mathcal{V} q+\mathrm{O}_3\ . \tag{3.16}$$

To prove Theorem 3.2, we need some information on the analytical expressions of the maps ϕ_{P} and ϕ_{RPS}.

- The analytical expression of the Cartesian coordinates $y^{(i)}$ and $x^{(i)}$ in terms of the Poincaré variables (3.6) is classical:

$$x^{(i)}=\mathcal{R}_{\mathrm{P}}^{(i)}\, x_{\mathrm{pl}}^{(i)}\ ,\quad y^{(i)}=\mathcal{R}_{\mathrm{P}}^{(i)}\, y_{\mathrm{pl}}^{(i)} \tag{3.17}$$

  where $x_{\mathrm{pl}}^{(i)}$, $y_{\mathrm{pl}}^{(i)}$ is the *planar Poincaré map* and $\mathcal{R}_{\mathrm{P}}^{(i)}$ is the *Poincaré rotation matrix*. Explicitly,
  - The *planar Poincaré map* is given by[8]

$$x_{\mathrm{pl}}^{(i)}=\big(\mathrm{x}_1^{(i)},\mathrm{x}_2^{(i)},0\big)\ ,\quad y_{\mathrm{pl}}^{(i)}=\big(\mathrm{y}_1^{(i)},\mathrm{y}_2^{(i)},0\big)=\beta_i\,\partial_{\lambda_i} x_{\mathrm{pl}}^{(i)}$$

    where

$$\begin{cases} \mathrm{x}_1^{(i)}:=\dfrac{1}{\bar m_i}\Big(\dfrac{\Lambda_i}{M_i}\Big)^2\Big(\cos u_i-\dfrac{\xi_i}{2\Lambda_i}\big(\eta_i\sin u_i+\xi_i\cos u_i\big)\\ \qquad -\dfrac{\eta_i}{\sqrt{\Lambda_i}}\sqrt{1-\dfrac{{\eta_i}^2+{\xi_i}^2}{4\Lambda_i}}\Big)\\ \mathrm{x}_2^{(i)}:=\dfrac{1}{\bar m_i}\Big(\dfrac{\Lambda_i}{M_i}\Big)^2\Big(\sin u_i-\dfrac{\eta_i}{2\Lambda_i}\big(\eta_i\sin u_i+\xi_i\cos u_i\big)\\ \qquad +\dfrac{\xi_i}{\sqrt{\Lambda_i}}\sqrt{1-\dfrac{{\eta_i}^2+{\xi_i}^2}{4\Lambda_i}}\Big)\\ \beta_i:=\dfrac{\bar m_i^2 M_i^4}{\Lambda_i^3}\end{cases}$$

    and $u_i=u_i(\Lambda_i,\lambda_i,\eta_i,\xi_i)=\lambda_i+\mathrm{O}(|(\eta_i,\xi_i)|)$ is the unique solution of the *(regularized) Kepler equation*

$$u_i-\frac{1}{\sqrt{\Lambda_i}}\sqrt{1-\frac{{\eta_i}^2+{\xi_i}^2}{4\Lambda_i}}\,\big(\eta_i\sin u_i+\xi_i\cos u_i\big)=\lambda_i\ ;$$

  - The *Poincaré rotation matrix* is given by

$$\mathcal{R}_{\mathrm{P}}^{(i)}=\begin{pmatrix}1-\mathrm{q}_i^2\mathrm{c}_i & -\mathrm{p}_i\mathrm{q}_i\mathrm{c}_i & -\mathrm{q}_i\mathrm{s}_i\\ -\mathrm{p}_i\mathrm{q}_i\mathrm{c}_i & 1-\mathrm{p}_i^2\mathrm{c}_i & -\mathrm{p}_i\mathrm{s}_i\\ \mathrm{q}_i\mathrm{s}_i & \mathrm{p}_i\mathrm{s}_i & 1-(\mathrm{p}_i^2+\mathrm{q}_i^2)\mathrm{c}_i\end{pmatrix}$$

    where $\mathrm{c}_i:=\frac{1}{2\Lambda_i-\eta_i^2-\xi_i^2}$ and $\mathrm{s}_i:=\sqrt{\mathrm{c}_i(2-(\mathrm{p}_i^2+\mathrm{q}_i^2)\mathrm{c}_i)}$.

[7] I.e., it preserves the two-form $d\eta\wedge d\xi+dp\wedge dq$.
[8] Compare, e.g., [3].

- The formulae of the Cartesian variables in terms of the RPS variables, differ from the formulae of the Poincaré map (3.17) just for the rotation matrix. Namely, one has

$$x^{(i)} = \mathcal{R}^{(i)}_{\rm RPS}\, x^{(i)}_{\rm pl}\,, \quad y^{(i)} = \mathcal{R}^{(i)}_{\rm RPS}\, y^{(i)}_{\rm pl}$$

where $x^{(i)}_{\rm pl}$, $y^{(i)}_{\rm pl}$ is the planar Poincaré map defined above. The expression of the RPS rotation matrices $\mathcal{R}^{(i)}_{\rm RPS}$ is a product of matrices

$$\mathcal{R}^{(i)}_{\rm RPS} = \mathcal{R}^*_n \mathcal{R}^*_{n-1} \cdots \mathcal{R}^*_i \mathcal{R}_i \tag{3.18}$$

where $\mathcal{R}_i$, $\mathcal{R}^*_i$ are 3×3 unitary matrices ($\mathcal{R}_1 \equiv {\rm id}$) given by

$$\mathcal{R}^*_i = \begin{pmatrix} 1-q_i^2\mathfrak{c}^*_i & -p_i q_i \mathfrak{c}^*_i & -q_i\mathfrak{s}^*_i \\ -p_i q_i\mathfrak{c}^*_i & 1-p_i^2\mathfrak{c}^*_i & -p_i\mathfrak{s}^*_i \\ q_i\mathfrak{s}^*_i & p_i\mathfrak{s}^*_i & 1-(p_i^2+q_i^2)\mathfrak{c}^*_i \end{pmatrix}, \qquad 1\le i\le n$$

$$\mathcal{R}_i = \begin{pmatrix} 1-q_{i-1}^2\mathfrak{c}_i & -p_{i-1} q_{i-1} \mathfrak{c}_i & -q_{i-1}\mathfrak{s}_i \\ -p_{i-1} q_{i-1}\mathfrak{c}_i & 1-p_{i-1}^2\mathfrak{c}_i & -p_{i-1}\mathfrak{s}_i \\ q_{i-1}\mathfrak{s}_i & p_{i-1}\mathfrak{s}_i & 1-(p_{i-1}^2+q_{i-1}^2)\mathfrak{c}_i \end{pmatrix}, \quad 2\le i\le n$$

where $\mathfrak{c}_i$, $\mathfrak{s}_i$, $\mathfrak{c}^*_j$, $\mathfrak{s}^*_j$ are analytic functions of $\frac{\eta_j^2+\xi_j^2}{2}$ and $\frac{p_j^2+q_j^2}{2}$'s, for $2\le i\le n$, $1\le j\le n$ even in z, with $\mathcal{R}_{i+1}$, $\mathcal{R}^*_j$ independent of (p_n,q_n), for $1\le j\le n-1$ (for the analytic expression, see [6, Appendix A.2]). Note that the only matrix in (3.18) depending on (p_n,q_n) is $\mathcal{R}^*_n$.

Extending results proven in [6], we now show that $\phi^{\rm RPS}_{\rm P}$ in (3.14) *"preserves rotations and reflections"* (Lemma 3.4 below).

Consider the transformations

$$\begin{aligned} \mathcal{R}_{1\leftrightarrow 2}\big(\Lambda,\ \lambda,\ {\rm z}\big) &= \Big(\Lambda,\ \frac{\pi}{2}-\lambda,\ \mathcal{S}_{1\leftrightarrow 2}{\rm z}\Big)\,; \\ \mathcal{R}^-_3\big(\Lambda,\ \lambda,\ {\rm z}\big) &= \big(\Lambda,\ \lambda,\ \mathcal{S}^-_{34}{\rm z}\big) \\ \mathcal{R}_g\big(\Lambda,\ \lambda,\ {\rm z}\big) &= \big(\Lambda,\ \lambda+g,\ \mathcal{S}_g{\rm z}\big) \end{aligned} \tag{3.19}$$

where, denoting the imaginary unit by i,

$$\begin{cases} \mathcal{S}_{1\leftrightarrow 2}(\eta,\xi,{\rm p},{\rm q}) := (-\xi,-\eta,{\rm q},{\rm p}) \\ \mathcal{S}^-_{34}(\eta,\xi,{\rm p},{\rm q}) := (\eta,\xi,-{\rm p},-{\rm q}) \\ \mathcal{S}_g : \big(\eta_j+{\rm i}\xi_j, \eta_j+{\rm i}\xi_j\big) \to \big(e^{-{\rm i}g}(\eta_j+{\rm i}\xi_j)\,,\ e^{-{\rm i}g}({\rm p}_j+{\rm i}{\rm q}_j)\big)\,. \end{cases} \tag{3.20}$$

Such transformations correspond, in Cartesian coordinates, to, respectively, reflection with respect to the plane $x_1=x_2$, the plane $x_3=0$ and a positive rotation of g around the $k^{(3)}$-axis:

$$\begin{aligned} &\mathcal{R}_{1\leftrightarrow 2} : x^{(i)} \to \big(x_2^{(i)},\ x_1^{(i)},\ x_3^{(i)}\big)\,, && y^{(i)} \to \big(-y_2^{(i)},\ -y_1^{(i)},\ -y_3^{(i)}\big) \\ &\mathcal{R}^-_3 : x^{(i)} \to \big(x_1^{(i)},\ x_2^{(i)},\ -x_3^{(i)}\big)\,, && y^{(i)} \to \big(y_1^{(i)},\ y_2^{(i)},\ -y_3^{(i)}\big) \\ &\mathcal{R}_g : x^{(i)} \to {\rm R}_3(g)\, x^{(i)}\,, && y^{(i)} \to {\rm R}_3(g)\, y^{(i)} \end{aligned} \tag{3.21}$$

where $\mathrm{R}_3(g)$ denotes the matrix

(3.22) $$\mathrm{R}_3(g) := \begin{pmatrix} \cos g & -\sin g & 0 \\ \sin g & \cos g & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad g \in \mathbb{T} .$$

For future use, consider also the following transformations, which are obtained obtained by suitably combining $\mathscr{R}_{1\leftrightarrow 2}$ and $\mathscr{R}_g$:

(3.23) $$\begin{cases} \mathscr{R}_1^-\Big(\Lambda,\ \lambda,\ \mathrm{z}\Big) := \mathscr{R}_{-\frac{\pi}{4}}\mathscr{R}_{1\leftrightarrow 2}\mathscr{R}_{\frac{\pi}{4}} = \Big(\Lambda,\ \pi-\lambda,\ \mathscr{S}_{14}^-\mathrm{z}\Big) \\ \mathscr{R}_2^-\Big(\Lambda,\ \lambda,\ \mathrm{z}\Big) := \mathscr{R}_{\frac{\pi}{4}}\mathscr{R}_{1\leftrightarrow 2}\mathscr{R}_{-\frac{\pi}{4}} = \Big(\Lambda,\ -\lambda,\ \mathscr{S}_{23}^-\mathrm{z}\Big) \end{cases}$$

where

(3.24) $$\mathscr{S}_{14}^-(\eta,\xi,\mathrm{p},\mathrm{q}) := (-\eta,\xi,\mathrm{p},-\mathrm{q})\ , \quad \mathscr{S}_{23}^-(\eta,\xi,\mathrm{p},\mathrm{q}) := (\eta,-\xi,-\mathrm{p},\mathrm{q})\ .$$

Note in particular:

REMARK 3.3 (D'Alembert rules)**.** Being $\mathscr{H}_{\mathrm{plt}}$-invariant under rotations around $k^{(3)}$ and under reflections with respect to the coordinate planes, the averaged perturbation $f_{\mathrm{P}}^{\mathrm{av}}$ does not change under the transformations $\mathrm{z} \to \mathscr{S}\mathrm{z}$, where $\mathscr{S}$ is as in (3.20) or in (3.24).

In particular, by D'Alembert rules, the expansion (2.6) follows.

LEMMA 3.4. *The map $\phi_{\mathrm{P}}^{\mathrm{RPS}}$ in (3.14) satisfies $\phi_{\mathrm{P}}^{\mathrm{RPS}}\mathscr{R} = \mathscr{R}\phi_{\mathrm{P}}^{\mathrm{RPS}}$, for any $\mathscr{R} = \mathscr{R}_{1\leftrightarrow 2}$, $\mathscr{R}_1^-$, $\mathscr{R}_2^-$, $\mathscr{R}_3^-$, $\mathscr{R}_g$ as in (3.19)–(3.24).*

Proof. It is enough to prove Lemma 3.4 for the transformations in (3.19) and (3.20). But this follows from the fact that *both in Poincaré variables and in* RPS *variables* the transformations in (3.21) have the form in (3.19)–(3.20). □

Proof of Theorem 3.2. For the proof of (3.15) (since $\phi_{\mathrm{P}}^{\mathrm{RPS}}$ is a regular map), we can restrict to the open dense set where none of the eccentricities e_i or of the nodes ν_{i+1} or $\bar\nu_i$ vanishes. In such set the angles γ_i, g_i, θ_i and ψ_i are well defined. By the definitions of λ_i in (3.6) and of λ_i in (3.7), one has

$$\lambda_i - \lambda_i = \Big(\ell_i + \mathrm{g}_i + \theta_i\Big) - \Big(\ell_i + \gamma_i + \psi_{i-1}^n\Big) = (\mathrm{g}_i - \gamma_i) + \theta_i - \psi_{i-1}^n\ .$$

The shifts $\mathrm{g}_i - \gamma_i = \alpha_{\mathrm{C}^{(i)}}(\bar\nu_i, P_i) - \alpha_{\mathrm{C}^{(i)}}(\nu_i, P_i) = \alpha_{\mathrm{C}^{(i)}}(\bar\nu_i, \nu_i)$ (compare their definitions in (3.4) and (3.5)), as well as the angles θ_i and ψ_j depend only on the angular momenta $\mathrm{C}^{(1)}, \cdots, \mathrm{C}^{(n)}$; hence, they do not depend upon λ.

With similar arguments one proves the second equation in (3.15).

Injectivity of $\mathcal{Z}(\Lambda,\cdot)$ follows from the definitions. That, for any fixed Λ, $\mathcal{Z}(\Lambda,\cdot)$ is symplectic, is a general property of any map of this form which is the projection over z of a symplectic transformation $(\Lambda,\lambda,\mathrm{z}) \to (\Lambda,\lambda,z)$ which leaves Λ unchanged.

Note now that $\phi_{\mathrm{P}}^{\mathrm{RPS}}$ preserves the quantities

$$|z|^2 = |\mathrm{z}|^2 = 2(|\Lambda|_1 - \mathrm{C}_3)\ ,$$

and the quantities

$$\eta_i^2 + \xi_i^2 = \upeta_i^2 + \upxi_i^2 = 2(\Lambda_i - \Gamma_i) \tag{3.25}$$

Therefore, it also preserves

$$|(\mathrm{p}, \mathrm{q})|^2 = |(p, q)|^2 \ . \tag{3.26}$$

From the previous equalities one has that $\phi_{\mathrm{P}}^{\mathrm{RPS}}$ sends injectively $(\eta_i, \xi_i) = 0$ to $(\upeta_i, \upxi_i) = 0$ and $(p, q) = 0$ to $(\mathrm{p}, \mathrm{q}) = 0$.

From the analytical expressions of ϕ_{P} and ϕ_{RPS} it follows that, when $(p, q) = 0$, the Poincaré variables $(\upeta, \upxi)$ and $\uplambda$ and the Deprit's (η, ξ) and λ respectively coincide. Therefore, from (3.15) and (3.25), we have $\varphi(\Lambda, 0) = 0$ and the first two equations in (3.16) follow. The fact that the remainder is $\mathrm{O}(|z|^3)$ is because $\mathscr{Z}(\Lambda, \cdot)$ is odd in z, as we shall now check. In fact, using Lemma 3.4 with $\mathscr{R} = \mathscr{R}_1^-$ or $\mathscr{R} = \mathscr{R}_2^-$, one finds that the $(\upeta, \mathrm{q})$-projection of $\mathscr{Z}(\Lambda, \cdot)$ is odd in (η, q), even in (ξ, p); the $(\upxi, \mathrm{p})$-projection of $\mathscr{Z}$ is odd in (ξ, p), even in (η, q). In particular, $\mathscr{Z}(\Lambda, \cdot)$ is odd in z.

Equation (3.26) and the fact that $\mathscr{Z}$ is odd imply that $(\mathrm{p}, \mathrm{q}) = \mathscr{R}(p, q) + \mathrm{O}(|z|^3)$, with $\mathscr{R} \in \mathrm{SO}(2n)$. Since p is odd in (ξ, p) and q is odd in (η, q), one has that $\mathscr{R}$ is block diagonal: $\mathscr{R} = \mathrm{diag}\,[\mathscr{V}_{\mathrm{p}}, \mathscr{V}_{\mathrm{q}}]$. The fact that $\mathscr{V}_{\mathrm{p}} = \mathscr{V}_{\mathrm{q}} := \mathscr{V}$ follows from Lemma 3.4, taking $\mathscr{R} = \mathscr{R}_{1 \leftrightarrow 2}$. □

4. PROOF OF THE NORMAL FORM THEOREM

For the proof of Theorem 2.1, we need some results from [6], to which we refer for details.

Let $\mathscr{H}_{\mathrm{RPS}}$ denote the planetary Hamiltonian expressed in RPS variables:

$$\mathscr{H}_{\mathrm{RPS}}(\Lambda, \lambda, \bar{z}) := \mathscr{H}_{\mathrm{plt}} \circ \phi_{\mathrm{RPS}}^{-1} = h_{\mathrm{K}}(\Lambda) + \mu f_{\mathrm{RPS}}(\Lambda, \lambda, \bar{z}) \tag{4.1}$$

where $\mathscr{H}_{\mathrm{plt}}$ is as in (2.1) and ϕ_{RPS} as in (3.10).

Note that, as $\mathscr{H}_{\mathrm{plt}}$ is rotation-invariant, the variables p_n, q_n in (3.9) are cyclic for $\mathscr{H}_{\mathrm{RPS}}$. Hence, the perturbation function f_{RPS} depends only on the remaining variables $(\Lambda, \lambda, \bar{z})$, where $\bar{z}$ is as in (3.11).

To avoid collisions, consider the ("partially reduced") variables in a subset of the maximal set $\mathscr{M}_{\max}^{6n-2}$ in (3.12) of the form

$$(\Lambda, \lambda, \bar{z}) \in \mathscr{M}^{6n-2} := \mathscr{A} \times \mathbb{T}^n \times B^{4n-2} \tag{4.2}$$

where $\mathscr{A}$ is a set of well separated semi major axes

$$\mathscr{A} := \left\{\Lambda : \ \underline{a}_j < a_j < \overline{a}_j \quad \text{for} \quad 1 \le j \le n\right\} \tag{4.3}$$

where $\underline{a}_1, \cdots, \underline{a}_n, \overline{a}_1, \cdots, \overline{a}_n$, are positive numbers satisfying $\underline{a}_j < \overline{a}_j < \underline{a}_{j+1}$ for any $1 \le j \le n$, $\overline{a}_{n+1} := \infty$; B^{4n-2} is a small $(4n-2)$-dimensional ball around the "secular origin" $\bar{z} = 0$.

As in the Poincaré setting, the Hamiltonian $\mathscr{H}_{\mathrm{RPS}}$ enjoys D'Alembert rules (namely, the symmetries in (3.20) and in (3.24)). Indeed, since the map $\phi_{\mathrm{P}}^{\mathrm{RPS}}$

in (3.14) commutes with any transformations $\mathscr{R}$ as in (3.19)–(3.24) and $\mathscr{H}_{\rm P}$ is $\mathscr{R}$-invariant, one has that $\mathscr{H}_{\rm RPS}$ is $\mathscr{R}$-invariant:

$$\mathscr{H}_{\rm RPS}\circ\mathscr{R}=\mathscr{H}_{\rm P}\circ\phi_{\rm P}^{\rm RPS}\circ\mathscr{R}=\mathscr{H}_{\rm P}\circ\mathscr{R}\circ\phi_{\rm P}^{\rm RPS}=\mathscr{H}_{\rm P}\circ\phi_{\rm P}^{\rm RPS}=\mathscr{H}_{\rm RPS}\ .$$

This implies that the averaged perturbation $f_{\rm RPS}^{\rm av}$ also enjoys D'Alembert rules and thus has an expansion analog to (2.6), but independent of (p_n, q_n):

$$f_{\rm RPS}^{\rm av}(\Lambda,\bar z)=C_0(\Lambda)+\mathscr{Q}_h(\Lambda)\cdot\frac{\eta^2+\xi^2}{2}+\bar{\rm Q}_{\rm v}(\Lambda)\cdot\frac{\bar p^2+\bar q^2}{2}+{\rm O}(|\bar z|^4) \tag{4.4}$$

with $\mathscr{Q}_h$ of order n and $\bar{\rm Q}_{\rm v}$ of order $(n-1)$. Note that the matrix $\mathscr{Q}_h$ in (4.4) is the same as in (2.6), since, when $p=(\bar p,p_n)=0$ and $q=(\bar q,q_n)=0$, Poincaré and RPS variables coincide.

The first step is to construct a normal form defined on a suitable lower-dimensional domain

$$(\Lambda,\check\lambda,\check z)\in\check{\mathscr{M}}^{6n-2}:=\mathscr{A}\times\mathbb{T}^n\times\check B^{4n-2} \tag{4.5}$$

(where $\check B^{4n-2}$ is an open ball in $\mathbb{R}^{4n-2}$ around $\check z=0$).

The existence of such normal form for the Hamiltonian (4.4) at any order s defined over a set of the form (4.5) is a corollary of [6, §7]. Indeed (by [6]), one can first conjugate $\mathscr{H}_{\rm RPS}=h_{\rm K}+\mu f_{\rm RPS}$ to a Hamiltonian

$$\tilde{\mathscr{H}}_{\rm RPS}=\mathscr{H}_{\rm RPS}\circ\tilde\phi=h_{\rm K}+\mu\tilde f_{\rm RPS}\ , \tag{4.6}$$

so that the average $\tilde f_{\rm RPS}^{\rm av}$ has the quadratic part into diagonal form:

$$\tilde f_{\rm RPS}^{\rm av}(\Lambda,\tilde z)=C_0(\Lambda)+\sum_{i=1}^n\sigma_i\frac{\tilde\eta_i^2+\tilde\xi_i^2}{2}+\sum_{i=1}^{n-1}\bar\varsigma_i\frac{\tilde p_i^2+\tilde q_i^2}{2}+{\rm O}(|\tilde z|^4) \tag{4.7}$$

where $\tilde z=(\tilde\eta,\tilde\xi,\tilde p,\tilde q)$ and σ_i, $\bar\varsigma_i$ denote[9] the eigenvalues of the matrices $\mathscr{Q}_h$ and $\bar{\rm Q}_{\rm v}$ in (4.4). Here, $\tilde\phi$ denotes the "symplectic diagonalization" which lets $\Lambda\to\Lambda$ and

$$\lambda=\tilde\lambda+\tilde\varphi(\Lambda,\tilde z)\ ,\ \eta={\rm U_h}(\Lambda)\tilde\eta\ ,\ \xi={\rm U_h}(\Lambda)\tilde\xi\ ,\ \bar p=\bar{\rm U}_{\rm v}(\Lambda)\tilde p\ ,\ \bar q=\bar{\rm U}_{\rm v}(\Lambda)\tilde q\ , \tag{4.8}$$

where ${\rm U_h}\in{\rm SO}(n)$ and $\bar{\rm U}_{\rm v}\in{\rm SO}(n-1)$ put $\mathscr{Q}_h$ and $\bar{\rm Q}_{\rm v}$ into diagonal form and will be chosen later. Note that $\tilde\phi$ leaves the set $\mathscr{M}^{6n-2}$ in (4.2) unchanged.

Next, we can use Birkhoff theory for rotation-invariant Hamiltonians, which allows to construct Birkhoff normal form for rotation-invariant Hamiltonian for which there are no resonances (up to a certain prefixed order) for those Taylor indices k such that $\sum k_i=0$ (rather than $k\neq0$ as in standard Birkhoff theory; compare Appendix A below). Indeed, as shown in [6, Proposition 7.2], the first-order Birkhoff invariants $\bar\Omega=(\sigma,\bar\varsigma)\in\mathbb{R}^n\times\mathbb{R}^{n-1}$ do not satisfy any resonance (up to any prefixed order s) over a (s-dependent) set $\mathscr{A}$ chosen as in (4.3), other

[9] In [6], the matrix $\mathscr{Q}_h$ is denoted by ${\rm Q_h}$; the $(n-1)$ components of $\bar\varsigma$ are denoted by ς_i. Beware that here we denote by ς_i also the n components of ς in (2.10). Actually, it will turn out that $\varsigma_i=\bar\varsigma_i$ (for $i\le n-1$): compare (i) in Remark 4.1 below.

than $\sum_{i=1}^{n}\sigma_i+\sum_{i=1}^{n-1}\bar\varsigma_i=0$ and $\bar\varsigma_n=0$. Thus, one can find a Birkhoff normalization $\breve\phi$ defined on the set (4.5), which conjugates $\tilde{\mathscr{H}}_{\rm RPS}=h_{\rm K}+\mu\tilde f_{\rm RPS}$ to

$$\breve{\mathscr{H}}_{\rm RPS}:=\tilde{\mathscr{H}}_{\rm RPS}\circ\breve\phi=h_{\rm K}+\mu\breve f_{\rm RPS}\ ,\tag{4.9}$$

where $\breve f^{\rm av}_{\rm RPS}$ is in the form (2.14), with r of dimension $n+(n-1)=2n-1$ and $\bar\Omega=(\sigma,\bar\varsigma)$ replacing Ω and $\bar{\rm P}_{\rm s}$ as in (2.16).

It is a remarkable fact, proved in [6], that both the transformations $\tilde\phi$ *and* $\breve\phi$ above leave $G(\Lambda,\bar z)$ in (2.11) unchanged

$$G\circ\tilde\phi=G\circ\breve\phi=G\ ,\tag{4.10}$$

(i.e., they commute with $\mathscr{R}_g$). Therefore, if we denote

$$\mathscr{M}^{6n}:=\{(\Lambda,\lambda,(\bar z,p_n,q_n)):\ (\Lambda,\lambda,\bar z)\in\mathscr{M}^{6n-2}\ ,\ p_n^2+q_n^2<4G(\Lambda,\bar z)\}$$
$$\mathscr{M}^{6n}_{\rm B}:=\{(\Lambda,\breve\lambda,(\breve z,p_n,q_n)):\ (\Lambda,\breve\lambda,\breve z)\in\breve{\mathscr{M}}^{6n-2}\ ,\ p_n^2+q_n^2<4G(\Lambda,\breve z)\}$$

where $\mathscr{M}^{6n-2}$ and $\breve{\mathscr{M}}^{6n-2}$ are as in (4.2) and (4.5), respectively, $\tilde\phi$ and $\breve\phi$ can be lifted to symplectic transformations

$$\tilde\Phi_{\rm RPS}:\mathscr{M}^{6n}\to\mathscr{M}^{6n}\ ,\quad \breve\Phi_{\rm RPS}:\mathscr{M}^{6n}_{\rm B}\to\mathscr{M}^{6n}\tag{4.11}$$

through the identity map on (p_n,q_n). Moreover:

(i) since $\mathscr{H}_{\rm RPS}$ is (p_n,q_n)-independent,

$$\mathscr{H}_{\rm RPS}\circ\tilde\Phi_{\rm RPS}=\tilde{\mathscr{H}}_{\rm RPS}\ ,\quad \tilde{\mathscr{H}}_{\rm RPS}\circ\breve\Phi_{\rm RPS}=\breve{\mathscr{H}}_{\rm RPS}\tag{4.12}$$

where $\tilde{\mathscr{H}}_{\rm RPS}$ and $\breve{\mathscr{H}}_{\rm RPS}$ are as in (4.6) and in (4.9), respectively;

(ii) $\tilde\Phi_{\rm RPS}$ is given by (4.8), with $(\bar p,p_n)$, $(\bar q,q_n)$, $(\tilde p,p_n)$, $(\tilde q,q_n)$, ${\rm U_v}:={\rm diag}\,[\bar{\rm U}_{\rm v},1]$ replacing $\bar p$, $\bar q$, $\tilde p$, $\tilde q$, $\bar{\rm U}_{\rm v}$, respectively;

(iii) $\breve\Phi_{\rm RPS}$ is of the form (2.15) (but with w and $\tilde{\rm z}$ replaced by $(\breve z,p_n,q_n)$ and $(\tilde z,p_n,q_n)$, respectively), since a similar property holds for $\breve\phi$.

Proof of Theorem 2.1. We prove only existence of the normal form; uniqueness follows from the same argument of standard Birkhoff normal form theory: compare [12].

Let $\tilde{\mathscr{H}}_{\rm P}$ as in (2.9), where $\tilde\Phi_{\rm P}$ is as in (2.7)–(2.8), for suitable fixed matrices ρ_h, ρ_v diagonalizing $\mathscr{Q}_h$, $\mathscr{Q}_v$ in (2.6). If $\mathscr{V}$ is as in (3.16), Eqs. (2.6), (4.4) and Theorem 3.2 imply that

$$\mathscr{V}^{\rm t}\mathscr{Q}_v\mathscr{V}={\rm Q_v}:={\rm diag}\,[\bar{\rm Q}_{\rm v},0]\ .\tag{4.13}$$

Thus, ${\rm Q_v}$ is diagonalized by the matrix $\mathscr{V}^{\rm t}\rho_v$. We can therefore choose ${\rm U_h}$ and $\bar{\rm U}_{\rm v}$ in (4.8) taking

$${\rm U_h}:=\rho_{\rm h}\ ,\quad {\rm U_v}:={\rm diag}\,[\bar{\rm U}_{\rm v},1]=\mathscr{V}^{\rm t}\rho_{\rm v}\ .\tag{4.14}$$

Analogously, let $\tilde\Phi_{\rm RPS}$, $\breve\Phi_{\rm RPS}$ as in (4.11), $\phi^{\rm RPS}_{\rm P}$ as in (3.14). Consider the transformation

$$\Phi_{\rm B}:=\Phi'_{\rm B}\circ\breve\Phi_{\rm RPS}\tag{4.15}$$

where

$$\Phi'_{\rm B}:=\tilde\Phi_{\rm P}^{-1}\circ\phi^{\rm RPS}_{\rm P}\circ\tilde\Phi_{\rm RPS}\ .\tag{4.16}$$

By (4.12), Φ_B transforms $\tilde{\mathcal{H}}_P$ into

$$\begin{aligned}\mathcal{H}_B &:= \tilde{\mathcal{H}}_P \circ \Phi_B \\ &= \mathcal{H}_P \circ \tilde{\Phi}_P \circ \Phi_B \\ &= \mathcal{H}_P \circ \tilde{\Phi}_P \circ \tilde{\Phi}_P^{-1} \circ \phi_P^{RPS} \circ \tilde{\Phi}_{RPS} \circ \check{\Phi}_{RPS} \\ &= \mathcal{H}_P \circ \phi_P^{RPS} \circ \tilde{\Phi}_{RPS} \circ \check{\Phi}_{RPS} \\ &= \mathcal{H}_{RPS} \circ \tilde{\Phi}_{RPS} \circ \check{\Phi}_{RPS} \\ &= \tilde{\mathcal{H}}_{RPS} \circ \check{\Phi}_{RPS} \\ &= \check{\mathcal{H}}_{RPS} = h_K + \mu \check{f}_{RPS} := h_K + \mu f_B\end{aligned}$$

where $f_B^{av} = \check{f}_{RPS}^{av}$ has just the claimed form.

To conclude, we have to check (2.15). It is sufficient to prove such equality (with w replaced by $(\tilde{z}, p_n, q_n)$) for the transformation Φ'_B in (4.15) (by item (iii) above). But this is an immediate consequence of (2.8), (3.16), (4.14), (4.16) and item (ii) above. □

REMARK 4.1. As a byproduct of the previous proof, we find that the matrices $\mathcal{Q}_v$ in (2.6) and $Q_v = \mathrm{diag}[\bar{Q}_v, 0]$ in (4.13) have the same eigenvalues, so the invariants ς_i and $\bar{\varsigma}_i$ in (2.6) and (4.7) coincide (for $i \le n-1$).

5. FURTHER REDUCTIONS AND BIRKHOFF NORMAL FORMS

In this section we discuss complete symplectic reduction by rotations, together with the respective Birkhoff normal forms, both in the spatial and planar cases (indeed, as in the three-body case, the planar case cannot be simply deduced from the spatial one in view of singularities). The Birkhoff normal form constructed in the spatial case (§ 5.1) is at the basis of the dynamical application given in § 6.

5.1. The totally reduced spatial case.

Proposition 5.1 below is a generalization at arbitrary order s of [6, Proposition 10.1]; the proof is reported, for completeness, in Appendix C.

Let us consider the system $\mathcal{H}_B = h_K + \mu f_B$ given by Theorem 2.1. Since the couple $(p_n, q_n) = (u_{2n}, v_{2n})$ does not appear into $\mathcal{H}_B$, we shall regard $\mathcal{H}_B$ as a function of $(6n-2)$ variables $(\Lambda, l, \bar{w})$, where

$$\bar{w} = (\bar{u}, \bar{v}) := (u_1, \cdots, u_{2n-1}, v_1, \cdots, v_{2n-1})$$

is taken in the set $\mathcal{M}_B^{6n-2} := \mathcal{A} \times \mathbb{T}^n \times B_\varepsilon^{4n-2}$. Without changing names to functions, we have a Hamiltonian of the form (compare (2.13)–(2.14))

$$\text{(5.1)} \qquad \begin{cases} \mathcal{H}_B(\Lambda, l, \bar{w}) = h_K + \mu f_B(\Lambda, l, \bar{w}) \quad \text{with} \\ f_B^{av}(\Lambda, \bar{w}) = C_0 + \bar{\Omega} \cdot \bar{r} + \frac{1}{2} \bar{\tau} \cdot \bar{r}^2 + \bar{\mathcal{P}}_3 + \cdots + \bar{\mathcal{P}}_s + \mathcal{P}(\Lambda, \bar{w}) \end{cases}$$

with $\bar{\mathcal{P}}_j$ homogeneous polynomials of degree j in $\bar{r}_i := \frac{\bar{u}_i^2 + \bar{v}_i^2}{2}$ and $\mathcal{P}(\Lambda, \bar{w}) = O(|\bar{w}|^{2s+2})$. We recall that $\mathcal{H}_B$ has been constructed, starting from the Hamiltonian $\mathcal{H}_{RPS}$ in (4.1), as $\mathcal{H}_B = \mathcal{H}_{RPS} \circ \tilde{\phi} \circ \check{\phi}$ where $\tilde{\phi}$, $\check{\phi}$ are given, respectively, in

(4.6) and (4.9). Recall also that, since $\tilde{\phi}$ and $\check{\phi}$ satisfy (4.10), the function G in (2.11) is an integral for $\mathcal{H}_{\rm B}$.

Incidentally, note that, since $\tilde{\phi}$ and $\check{\phi}$ leave Λ's unvaried, their respective $\bar{z}$, $\tilde{z}$-projections actually preserve the Euclidean length of $\tilde{z}$, $\check{z}$:

$$|\Pi_{\tilde{z}} \circ \tilde{\phi}(\Lambda, \lambda, \bar{z})| = |\bar{z}| \ , \quad |\Pi_{\tilde{z}} \circ \check{\phi}(\Lambda, \lambda, \check{z})| = |\check{z}| \ . \tag{5.2}$$

The Hamiltonian (5.1) is thus preserved under the G-flow, i.e., under the transformations, which we still denote by $\mathcal{R}_g$, defined as in (3.19)–(3.20), with (Λ, λ, z) replaced by $(\Lambda, l, \bar{w})$. It is therefore natural to introduce the symplectic transformation

$$\hat{\phi}: \quad \begin{cases} (\Lambda, G, \hat{l}, \hat{g}, \hat{w}) \to (\Lambda, l, \bar{w}) \\ \hat{w} = (\hat{u}, \hat{v}) \ , \quad \hat{u} = (\hat{u}_1, \cdots, \hat{u}_{2n-2}) \ , \quad \hat{v} = (\hat{v}_1, \cdots, \hat{v}_{2n-2}) \end{cases}$$

which acts as the identity on Λ and, on the other variables, is defined by the following formulae

$$\hat{\phi}: \ l_j = \hat{l}_j + \hat{g}; \quad u_j + {\rm i}\, v_j = \begin{cases} e^{-{\rm i}\hat{g}}(\hat{u}_j + {\rm i}\,\hat{v}_j) \ , & j \neq 2n-1 \\ e^{-{\rm i}\hat{g}}\sqrt{\varrho^2 - |\hat{w}|^2} \ , & j = 2n-1 \end{cases} \tag{5.3}$$

where $\varrho = \varrho(\Lambda, G)$ is defined by

$$\varrho^2 := 2\Big(\sum_{1 \le j \le n} \Lambda_j - G\Big) \ . \tag{5.4}$$

The map $\hat{\phi}$ is well defined for $(G, \hat{g}, \Lambda, \hat{l}, \hat{w}) \in \mathbb{R}_+ \times \mathbb{T} \times \hat{\mathcal{M}}^{6n-4}$, where $\hat{\mathcal{M}}^{6n-4}$ is the subset of $(\Lambda, \hat{l}, \hat{w}) \in \mathcal{A} \times \mathbb{T}^n \times \mathbb{R}^{4(n-1)}$ described by the following inequalities

$$|\hat{w}| < \varrho < \varepsilon \ . \tag{5.5}$$

As it immediately follows from (5.3), the action variable G is the integral (2.11). Hence, its conjugate variable $\hat{g}$ is *cyclic* for the Hamiltonian, parametrized by G,

$$\hat{\mathcal{H}} := \mathcal{H}_{\rm B} \circ \hat{\phi} = \mathcal{H}_{\rm RPS} \circ \tilde{\phi} \circ \check{\phi} \circ \hat{\phi} = h_{\rm K} + \mu \hat{f} \ . \tag{5.6}$$

and we may regard $\hat{\mathcal{H}}$ as a Hamiltonian of $(3n-2)$ degrees of freedom. Note, however, that $\hat{\mathcal{H}}$ is no longer in normal form.

Now, let $\mathcal{A}$ and ε be, respectively, as in (4.3) and (5.5), and, for $0 < \hat{\delta} < \delta < \varepsilon$, define the following sets[10]

$$\check{\mathcal{A}} = \check{\mathcal{A}}(\hat{\delta}, \delta) := \{\Lambda \in \mathcal{A} : \ \hat{\delta} < \varrho < \delta\} \ , \tag{5.7}$$

$$\check{\mathcal{M}}^{6n-4} = \check{\mathcal{M}}^{6n-4}(\hat{\delta}, \delta) := \Big\{\Lambda \in \check{\mathcal{A}}(\hat{\delta}, \delta) \ , \ \check{\lambda} \in \mathbb{T}^n \ , \ |\check{w}| \le \frac{1}{4}\hat{\delta}\Big\} \ . \tag{5.8}$$

PROPOSITION 5.1 (Birkhoff normal form for the fully reduced spatial planetary system)**.** *For any integer* $s \ge 2$*, there exists* $0 < \delta^* < \varepsilon$ *such that whenever* $0 < \hat{\delta} < \delta < \delta^*$ *one can find a real-analytic symplectic transformation*

$$\phi_s : (\Lambda, \check{\lambda}, \check{w}) \in \check{\mathcal{M}}^{6n-4}(\hat{\delta}, \delta) \to (\Lambda, \hat{\lambda}, \hat{w}) \in \hat{\mathcal{M}}^{6n-4}$$

[10] The number 1/4 in (5.8) is arbitrary: one could replace it by any $0 < \vartheta < 1$.

such that the planetary Hamiltonian $\hat{\mathscr{H}}$ *in* (5.6) (*regarded as a function of* $(6n-4)$ *variables, parametrized by* G) *takes the form*

$$\text{(5.9)}\qquad \begin{cases} \check{\mathscr{H}} = \hat{\mathscr{H}} \circ \phi_s(\Lambda, \check{\lambda}, \check{w}) = h_{\rm K}(\Lambda) + \mu \check{f}(\Lambda, \check{\lambda}, \check{w}) \quad \textit{with} \\ \check{f}^{\rm av} = \check{\rm P}_s + {\rm O}(|\check{w}|^{2s+1})\ , \quad \check{\rm P}_s := \check{C}_0 + \check{\Omega} \cdot \check{r} + \frac{1}{2}\check{\tau} \cdot \check{r}^2 + \check{\mathscr{P}}_3 + \cdots + \check{\mathscr{P}}_s \end{cases}$$

where $\check{w} = (\check{u}, \check{v}) = (\check{u}_1, \cdots, \check{u}_{2n-2}, \check{v}_1, \cdots, \check{v}_{2n-2})$ *and the* $\check{\mathscr{P}}_j$*'s are homogeneous polynomials of degree* j *in* $\check{r}_i = \frac{\check{u}_i^2 + \check{v}_i^2}{2}$, *with coefficients depending on* Λ.

The first-order Birkhoff invariants $\check{\Omega}_i$ *of such normal form do not satisfy identically any resonance and the matrix* $\check{\tau}$ *of the second-order Birkhoff invariants is nonsingular. The transformation* ϕ_s *may be chosen to be* δ^{2s+1}*-close to the identity.*

5.2. **The totally reduced planar case.** Let us now restrict to the planar setting, that is, when the coordinates $y^{(i)}$, $x^{(i)}$ in (2.1) are taken in $\mathbb{R}^2$ instead of $\mathbb{R}^3$. Also in this case, in view of the presence of the integral $\sum_{i=1}^n x_1^{(i)} y_2^{(i)} - x_2^{(i)} y_1^{(i)}$, a (total) symplectic reduction is available (compare, also, [9]).

In the case of the planar problem, the instantaneous ellipses $\mathscr{E}_i$ defined in § 3 become coplanar and both the Poincaré variables $(\Lambda, \lambda, {\rm z})$ and RPS variables (Λ, λ, z) reduce to the *planar Poincaré variables.* Analytically, the planar Poincaré variables can be derived from (3.6) by setting $\theta_i = 0$ and disregarding the p and q.

To avoid introducing too many symbols, *we keep denoting the planar Poincaré variable*

$$(\Lambda, \lambda, {\rm z}) = (\Lambda, \lambda, \eta, \xi) \in \mathscr{M}^{4n} := \mathscr{A} \times \mathbb{T}^n \times B^{2n} \subseteq \mathbb{R}_+^n \times \mathbb{T}^n \times \mathbb{R}^{2n}$$

where $\mathscr{A}$ can be taken as in (4.3) above and B^{2n} the $(2n)$-dimensional open ball around the origin, whose radius (related to eccentricities, as in the spatial case), is chosen so small to avoid collisions; beware that ${\rm z} = (\eta, \xi)$, here, is $2n$-dimensional. The planetary Hamiltonian in such variables is given by $\mathscr{H}_{\rm pl}(\Lambda, \lambda, {\rm z}) = h_{\rm Kep}(\Lambda) + \mu f_{\rm pl}(\Lambda, \lambda, {\rm z})$ obtained from $\mathscr{H}_{\rm P}$ in (2.4) by putting, simply, ${\rm p} = 0 = {\rm q}$; clearly, also the expression of the averaged perturbation, $f_{\rm pl}^{\rm av}$, can be derived in the same way from (2.5).

Since, in particular, the *"horizontal" first-order Birkhoff invariants* σ *do not satisfy resonances of any finite order* s on[11] $\mathscr{A}$, the Birkhoff-normalization up to any order can be constructed in the planar case and it coincides with the expression of $f_{\rm B}^{\rm av}$ in (2.14), where one has to take

$$w = (u, v) =: \big((\check{\eta}, \check{\rm p}), (\check{\xi}, \check{\rm q})\big) = \big((\check{\eta}, 0), (\check{\xi}, 0)\big).$$

We recall in fact that the transformation $\Phi_{\rm B}$ in Theorem 2.1 sends injectively $\check{\rm p} = 0 = \check{\rm q}$ to ${\rm p} = 0 = {\rm q}$ and hence the restriction $\Phi_{\rm B}|_{\check{\rm p}=0=\check{\rm q}}$ performs the desired normalization in the planar case.

Let us denote by

$$\text{(5.10)}\qquad \check{\mathscr{H}}_{\rm pl}(\Lambda, \check{\lambda}, \check{\rm z}) = h_{\rm K}(\Lambda) + \mu \check{f}_{\rm pl}(\Lambda, \check{\lambda}, \check{\rm z})\ , \quad (\Lambda, \check{\lambda}, \check{\rm z}) \in \check{\mathscr{M}}_{\rm pl}^{4n} := \mathscr{A} \times \mathbb{T}^n \times B_{\check{\epsilon}}^{2n}$$

[11] Compare [10] or, equivalently, use again [6, Proposition 7.2]).

the planar Birkhoff-normalized system, that is, the system such that the averaged perturbation $\check f^{\rm av}(\Lambda,\check z)$ is in Birkhoff normal form: the Birkhoff normal form of order 4 is given by

$$\check f^{\rm av}_{\rm pl}(\Lambda,\check z) = C_0(\Lambda) + \sum_{1\le i\le n} \sigma_i(\Lambda)\check r_i + \frac{1}{2}\sum_{1\le i,j\le n} \bar\tau_{ij}(\Lambda)\check r_i\check r_j + {\rm O}(|\check z|^6) \tag{5.11}$$

with $\check r_i := \frac{\check\eta_i^2+\check\xi_i^2}{2}$.

The asymptotic evaluation of the first-order invariants σ and especially of planar torsion $\bar\tau$ in (5.11) for general $n\ge 2$ can be found in the paper by J. Féjoz [10] and in the notes by M. Herman [11]. However, since the asymptotics considered in such papers is slightly different from the one considered in [6] for the general spatial case[12], we collect here the asymptotic expressions of σ and $\bar\tau$ as they follow from [6] (compare also below for a short proof):

- The first-order Birkhoff invariants σ into (5.11) satisfy

$$\sigma_j = \begin{cases} -\dfrac{3}{4} m_1 m_2 \dfrac{a_1}{a_2^2\Lambda_1}\Big(\dfrac{a_1}{a_2} + {\rm O}(\dfrac{a_1}{a_2})^2\Big)\,, & j=1\\[2ex] -\dfrac{3m_j}{4\Lambda_j a_j^3}\displaystyle\sum_{1\le i<j} m_i a_i^2\big(1+{\rm O}(a_j^{-2})\big)\,, & 2\le j\le n. \end{cases} \tag{5.12}$$

- The second-order Birkhoff invariants $\bar\tau$ into (5.11) satisfy[13], for $n=2$,

$$\bar\tau = m_1 m_2 \frac{a_1^2}{a_2^3}\begin{pmatrix} \frac{3}{4\Lambda_1^2} & -\frac{9}{4\Lambda_1\Lambda_2}\\ -\frac{9}{4\Lambda_1\Lambda_2} & -\frac{3}{\Lambda_2^2}\end{pmatrix}(1+{\rm O}(a_2^{-5/4}))\,, \tag{5.13}$$

and[14] for $n\ge 3$,

$$\bar\tau = \begin{pmatrix} \tilde\tau + {\rm O}(\delta) & {\rm O}(\delta)\\ {\rm O}(\delta) & \bar\tau_{nn} + {\rm O}(\delta^2)\end{pmatrix} \quad \text{where} \quad \delta := a_n^{-3} \tag{5.14}$$

---

[12] In [10, 11] the semi major axes $a_1 < \cdots < a_n$ are taken well spaced in the following sense: at each step, namely, when a new planet (labeled by "1") is added to the previous $(n-1)$ (labeled from 2 to n) $a_2, \cdots, a_n$ are taken O(1) and $a_1 \to 0$. In [6] one takes $a_1, \cdots, a_{n-1}$ =O(1) and $a_n\to\infty$. The reason for the different choice lies in technicalities related to the evaluation of the "vertical torsion" (i.e., the entries of the torsion matrix in (2.14) with indices from $n+1$ to $2n$) in the spatial case. The asymptotics in [10] and [11] does not allow (as in [6]) to evaluate at each step the new torsion simply picking the dominant terms, because of increasing errors (of O(1)): compare the discussion in [11, end of p. 23]. To overcome these technicalities (and to avoid too many computations), Herman introduces a modification of the Hamiltonian and a new fictitious small parameter δ, also used in [10]. Note that, since Herman computes the asymptotics using Poincaré variables, by the presence of the 0-eigenvalue ς_n, he could not use the limit $a_n\to\infty$, being such limit singular (not continuous) for the matrices ρ_v in (2.8).

[13] The evaluation of the planar three-body torsion (5.13) is due to Arnold. Compare [2, p.138, Eq. (3.4.31)], noticing that in [2] the second-order Birkhoff invariants are defined as one half the $\bar\tau_{ij}$'s and that a_2^4 should be a_2^7. Compare also with [11, beginning of p. 21], (where a factor a_2^3 at denominator of each entry is missing).

[14] Compare (5.14) and (5.15) with the inductive formulae obtained in the other asymptotics in [11, end of p. 21].

with $\tilde{\tau}$ of rank $(n-1)$ and

$$\bar{\tau}_{nn} = -3\frac{m_n}{\Lambda_n^2}\sum_{1\le j<n} m_j \frac{1}{a_n}\Big(\frac{a_j^2}{a_n^2} + \mathrm{O}\Big(\frac{a_j^4}{a_n^4}\Big)\Big)\,. \tag{5.15}$$

- Eq. (5.12) implies in particular nonresonance of the σ_j's into a domain of the form of (4.3) (with $\underline{a}_j$, $\overline{a}_j$ depending on s).
- Using (5.13)–(5.15) and $\Lambda_i^2 = m_i^2 m_0 a_i(1+\mathrm{O}(\mu))$, one finds that, for $n \ge 2$ and $0 < \delta_\star < 1$ there exist[15] $\bar{\mu} > 0$, $0 < \underline{a}_1 < \overline{a}_1 < \cdots < \underline{a}_n < \overline{a}_n$ such that, on the set $\mathscr{A}$ defined in (4.3) and for $0 < \mu < \bar{\mu}$, the matrix $\bar{\tau}$ is nonsingular: $\det\bar{\tau} = \bar{d}_n(1+\delta_n)$, where $|\delta_n| < \delta_\star$ and

$$\bar{d}_n = (-1)^{n-1}\frac{117}{48}\Big(\frac{3}{m_0}\Big)^{n-1}\frac{m_2}{m_0 m_n}\frac{a_1^3}{a_2^3 a_n^2}\prod_{j=2}^{n}\frac{1}{a_j^2}\,. \tag{5.16}$$

Proof of (5.12)–(5.15). Eqs. (5.12)–(5.15) can be obtained, e.g., as a particular case of more general formulae, proved in [6]: For Equation (5.12), for $n = 2$, use [6, Eq. (7.5)], and "Herman resonance" $\sigma_1 = -\varsigma - \sigma_2$; in the case $n \ge 3$, compare the asymptotic expression of σ_n after [6, Eq. (7.7)]. Equation (5.13) corresponds[16] to [6, Eq. (8.33)]. Equation (5.14) is obtained from [6, Eq. (8.45)] picking only the entries which are relative to the horizontal variables $\frac{\check{\eta}_i^2+\check{\xi}_i^2}{2}$. In particular, the matrix $\tilde{\tau}$ of (5.14) is the horizontal part (that is, the upper left $(n-1)\times(n-1)$ submatrix) of the matrix $\hat{\tau}$ of [6, Eq. (8.45)]. For Eq. (5.15), note that $\bar{\tau}_{nn}$ is the upper left entry of the 2×2 matrix $\bar{\tau}$ in [6] and use the asymptotics for $\mathrm{r}_1(a_2, a_1)$ given in [6, Eq. (8.32)]. □

We describe, now, briefly a *(total) symplectic reduction for the planar problem* and discuss the relative Birkhoff normal form. The discussion is based on tools and arguments similar to those used in § 5.1 above for the spatial case.

Indeed, quite analogously to the spatial case, the Hamiltonian (5.10) is preserved under the G-flow, where now G denotes the function in (2.11) with $\bar{z} = (\eta, \xi, 0, 0)$. Therefore, as in (5.3), one introduces the symplectic transformation $\hat{\phi}_{\rm pl}$ which lets $\Lambda \to \Lambda$ and

$$\hat{\phi}_{\rm pl}:\ \check{\lambda}_j = \hat{\lambda}_j + \hat{g}\,,\qquad (\check{\eta}_j + \mathrm{i}\check{\xi}_j) = \begin{cases} e^{-\mathrm{i}\hat{g}}(\hat{\eta}_j + \mathrm{i}\hat{\xi}_j)\,, & \text{for } j \ne n\,,\\ e^{-\mathrm{i}\hat{g}}\sqrt{\varrho^2 - |\hat{z}|^2}\,, & \text{for } j = n\,,\end{cases}$$

where ϱ^2 is as in (5.4) and $\hat{z}$ has components $(\hat{\eta}_1, \cdots, \hat{\eta}_{n-1}, \hat{\xi}_1, \cdots, \hat{\xi}_{n-1})$.

Again, in order for $\hat{\phi}_{\rm pl}$ to be well defined, the domain $\hat{\mathscr{M}}^{4n}_{\rm pl}$ of $(G, \hat{g}, \Lambda, \hat{\lambda}, \hat{z})$ will be taken of the form

$$(\Lambda, G) \in \mathscr{A}\times\mathbb{R}_+\,,\quad (\hat{\lambda}, \hat{g}) \in \mathbb{T}^{n+1}\,,\quad \hat{z} \in \mathbb{R}^{2n}\,,\quad |\hat{z}| < \varrho(\Lambda, G) < \epsilon^* \le \check{\epsilon}\,, \tag{5.17}$$

[15] $\bar{\mu}$ is taken small only to simplify (5.16), but a similar evaluation hold with $\bar{\mu} = 1$. Note that the normal planar torsion is not sign-definite [Herman]. A similar results holds also in the spatial case [6, Eq. (8.38)].

[16] In [6, Eq. (8.33)], $\bar{\tau}$ is denoted by $\tau_{\rm pl}$.

where $\check{\mathcal{H}}$, $\check{\mathrm{e}}$ are as in (5.10). We denote by $\hat{\mathcal{H}}_{\rm pl} := \check{\mathcal{H}}_{\rm pl} \circ \hat{\phi}_{\rm pl}$ the planar "reduced Hamiltonian".

Adapting the proof of Proposition 5.1 above to the planar case, we then have:

- *For any $s \in \mathbb{N}$, one can always find a set of symplectic variables $(\Lambda, \check{\lambda}, \check{\mathrm{z}})$ varying on some domain $\check{\mathcal{M}}_{\rm pl}^{4n-2} \subseteq \mathbb{R}^n \times \mathbb{T}^n \times \mathbb{R}^{2n-2}$ of the form (5.7)–(5.8) with $6n-4$ replaced by $4n-2$, such that, in such variables, the reduced Hamiltonian $\hat{\mathcal{H}}_{\rm pl}$ is put into the form $\check{\mathcal{H}}_{\rm pl} = h_{\rm K} + \mu \check{f}_{\rm pl}$, with the averaged perturbation $\check{f}_{\rm pl}^{\rm av}$ in normal form of order $2s$. The first- and second-order Birkhoff invariants are given by*

$$\left\{ \begin{array}{l} \hat{\sigma}_i(\Lambda; G) = \sigma_i(\Lambda) - \sigma_n(\Lambda) + \mathrm{O}(\varrho^2) \ , \\ \hat{\bar{\tau}}_{ij}(\Lambda; G) = \bar{\tau}_{ij}(\Lambda) - \bar{\tau}_{in}(\Lambda) - \bar{\tau}_{jn}(\Lambda) + \bar{\tau}_{nn}(\Lambda) + \mathrm{O}(\varrho^2) \ . \end{array} \right. \tag{5.18}$$

Using (5.12)–(5.15), one immediately sees that

- *The invariants $\hat{\sigma}$ and $\hat{\bar{\tau}}$ in (5.18) are asymptotically close (for $a_1, \cdots, a_{n-1} = \mathrm{O}(1)$, $a_n \to \infty$ and $\varrho \to 0$) to the unreduced σ_i and $\bar{\tau}_{ij}$ (for $i, j \le n-1$).*

Therefore, the following corollary follows at once.

COROLLARY 5.2. *Fix $n \ge 2$ and $0 < \delta_\star < 1$, $s \ge 4$. Then there exist $\bar{\mu} > 0$, $0 < \underline{a}_1 < \overline{a}_1 < \cdots < \underline{a}_n < \overline{a}_n$ such that for any $\mu < \bar{\mu}$ and for any $\Lambda \in \mathcal{A}_{\rm G}$, where $\mathcal{A}_{\rm G}$ is the set in (5.17), the first-order Birkhoff invariants $\hat{\sigma}$ are nonresonant up to order s and the matrix $\hat{\tau}$ is nonsingular:* $\det \hat{\tau} = \check{d}_n(1+\delta_n)$, *with $|\delta_n| < \delta_\star$ and*

$$\check{d}_n = \left\{ \begin{array}{ll} m_1 m_2 \dfrac{a_1^2}{a_2^3} \dfrac{3}{4\Lambda_1^2} \ , & n = 2 \ , \\ \bar{d}_{n-1} \ , & n \ge 3 \ , \end{array} \right.$$

where $\bar{d}_n$ is as in (5.16).

6. LONG-TIME STABILITY OF PLANETARY ACTIONS

In the 70's N. N. Nehorošev [13] proved exponential stability of the semi major axes in the planetary problem: *during the motion, the semi major axes*[17] *$a_i(t)$ stay close to their initial values for exponentially long times,* i.e.,

$$|a_i(t) - a_i(0)| < C\mu^b \ , \quad \forall\ |t| \le \frac{1}{C\mu} \exp\left(\frac{1}{C\mu^a}\right) ,$$

for suitable positive constants C, a, b, provided μ is sufficiently small and that the initial values $a_i(0)$ are in the well separated regime (4.3). The numbers C, a and b given by Nehorošev, were later improved in [14].

Note that, while the semi major axes stay close to their initial values, the "secular" Poincaré variables $\mathrm{z} = (\eta, \xi, \mathrm{p}, \mathrm{q})$ in (3.6) (also used by Nehorošev in describing the motion) may, in principle, vary on a relatively large ball B_r^{4n} around the origin: indeed, in [13] and [14] no information is given on possible "order one" variations of eccentricities and relative inclinations.

[17] Which are related to the Poincaré variables Λ as in (3.4).

Here, we prove a complementary result, namely, that in a suitable *partially*[18] *nonresonant* open set in phase space, the secular actions related to eccentricities and inclinations stay close to their initial values for arbitrarily long times compared to the distance from the secular equilibrium. More precisely, we have:

THEOREM 6.1. *Let $\mathscr{A}$ be as in (4.3); let $s \geq 2$, $\tau > n-1$ and δ^* be as in Proposition 5.1. Then there exist $c^\star > 1$ and $0 < \epsilon^\star < \delta^*/2$ such that, for $0 < \hat\epsilon < \epsilon < \epsilon^\star$, $(c^\star \hat\epsilon)^3 < \mu < (\hat\epsilon/c^\star)^{3/2}$ and $\kappa > 0$ there is an open set $\mathscr{A}_\star \subseteq \mathscr{A}$ of Lebesgue measure*

$$\text{meas}\,\mathscr{A}_\star \geq \Big(1 - \frac{c^\star}{\kappa}\sqrt{\hat\epsilon}\Big)\,\text{meas}\,\mathscr{A}, \tag{6.1}$$

such that the following holds. Let $\mathscr{M}_{\rm pn}$, and $\mathscr{M}'_{\rm pn}$ be the phase space regions in (5.7), (5.8) given, respectively, by $\check{\mathscr{M}}^{6n-4}(\hat\epsilon,\epsilon)$ with $\mathscr{A}$ replaced by $\mathscr{A}_\star$ and by $\check{\mathscr{M}}^{6n-4}(\hat\epsilon/2, 2\epsilon)$ with with $\mathscr{A}$ replaced by $\mathscr{A}_\star$ and $1/4$ replaced by $3/4$. Then any trajectory generated by $\check{\mathscr{H}}$ with initial datum in $\mathscr{M}_{\rm pn}$ remains in $\mathscr{M}'_{\rm pn}$ and satisfies[19]

$$\max_i\{|\Lambda_i(t) - \Lambda_i(0)|\} < \hat\epsilon^2\,, \quad \max_j\{|\check r_j(t) - \check r_j(0)|\} < \kappa\hat\epsilon^2 \tag{6.2}$$

for all

$$|t| \leq \mathfrak{t} := \frac{\kappa}{c^\star \mu\,\hat\epsilon^{2s-1}}. \tag{6.3}$$

In particular, the action variables $\check r_j$ satisfy $\max_j\{|\check r_j(t) - \check r_j(0)|\} < \hat\epsilon^{9/4}$ provided $\check r_j(0) \leq \epsilon^2$ and $\Lambda_j(0)$ belong to a set of density $(1 - c^\star \epsilon^{1/4})$.

REMARK 6.2. Stability estimates hold up exponentially long times in *completely nonresonant* regions, i.e., essentially in an open neighborhood of KAM tori. Let $\mathscr{K} \subseteq \mathscr{M}_{\rm pn}$ denote the Kolmogorov set (i.e., the union of KAM tori) of $\check{\mathscr{H}}$. Then *for initial data on the open set $\mathscr{K}_d$ around $\mathscr{K}$, hence, of measure*[20]

$$\text{meas}\,\mathscr{K}_d \geq \text{meas}\,\mathscr{K} \geq (1 - \sqrt{\hat\epsilon})\,\text{meas}\,\mathscr{M}_{\rm pn}$$

one can replace (6.3) with $|t| \leq \mathfrak{t}_{\rm exp}(d) := \dfrac{\kappa\hat\epsilon^2}{c^\star d^{\sigma'}} e^{\frac{1}{c^\star d^\sigma}}$ (for some $0 < \sigma < 1 < \sigma'$).

Here is a sketch of proof. The set $\mathscr{K}_d$ is a high-order nonresonant set, being equivalent to the direct product $\mathscr{N}_d \times \mathbb{T}^{3n-2}$, where $\mathscr{N}_d$ is $(\alpha, K) \sim (d^{1-\sigma}, d^{-\sigma})$ nonresonant for the frequency map $(\Lambda, \check r) \to \varpi_1(\Lambda, \check r) = \partial_{(\Lambda,\check r)}(h_{\rm K}(\Lambda) + \mu \check{\rm P}_2(\Lambda, \check r))$. Here, $h_{\rm K}$ and $\check{\rm P}_2$ are as in (5.9).

By Averaging Theory, one can find an open set $\check{\mathscr{A}}_1 \subseteq \check{\mathscr{A}}$, a number $0 < c < 1$ and a real-analytic symplectic transformation

$$\Phi\colon \big((\Lambda, r), \vartheta\big) \in \check{\mathscr{A}}_1 \times \mathscr{I}^{2n-2}_{\hat\delta/8} \times \mathbb{T}^{3n-2} \to \Phi\big((\Lambda, r), \vartheta\big) \in \check{\mathscr{M}}^{6n-4}\,,$$

[18] I.e., Λ-nonresonant, but possibly resonant in the secular variables.

[19] Recall that $\check w = (\check u, \check v) = (\check u_1, \cdots, \check u_{2n-2}, \check v_1, \cdots, \check v_{2n-2})$ and that $\check r_j = \frac{\check u_j^2 + \check v_j^2}{2}$.

[20] See [6].

where $\check{\mathcal{M}}^{6n-4}$ is as in (5.7)–(5.8) and $\mathcal{I}_\delta$ is the interval $\mathcal{I}_\delta = (c\delta, \delta) \subseteq \mathbb{R}_+$, which conjugates the Hamiltonian (5.9) (with $s = 2$) to a new Hamiltonian of the form

$$H\big((\Lambda, r), \vartheta\big) := \check{\mathcal{H}} \circ \Phi\big((\Lambda, r), \vartheta\big) = h_{\rm K}(\Lambda) + \mu \check{\rm P}_2(\Lambda, r) + {\rm O}\big(\mu^2; (\Lambda, r), \vartheta\big) \ .$$

Consider the frequency map $(\Lambda, \check r) \to \varpi_1(\Lambda, \check r) := \partial_{(\Lambda, \check r)}(h_{\rm K} + \mu \check{\rm P}_2)$ and, for any $0 < \gamma_2 \le \gamma_1$ and $\tau' > 3n-2$, consider the generalized $(\gamma_1, \gamma_2, \tau')$-Diophantine numbers of the form[21]

$$\mathcal{D}_{\gamma_1, \gamma_2, \tau'} := \bigcap_{0 \ne k = (k_1, k_2) \in \mathbb{Z}^n \times \mathbb{Z}^{2n-2}} \left\{ \omega \in \mathbb{R}^{3n-2} : \ |\omega \cdot k| \ge \begin{cases} \frac{\gamma_1}{|k|^{\tau'}} & \text{if } k_1 \ne 0 \\ \frac{\gamma_2}{|k_2|^{\tau'}} & \text{otherwise} \end{cases} \right\} .$$

By KAM theory[22], for any $\omega \in \mathcal{D}_{\gamma_1, \gamma_2, \tau'}$ lying in the ϖ_1-image of $\check{\mathcal{A}}_1 \times \mathcal{I}^{2n-2}_{\hat\delta/8}$, one can find a Lagrangian, analytic torus $\mathcal{T}_\omega := \phi(\mathbb{T}^{3n-2}; \omega) \in \mathcal{K}$, defined by an embedding

$$\phi(\cdot, \omega) \colon \vartheta \in \mathbb{T}^{3n-2} \to \phi(\vartheta; \omega) = (v(\vartheta; \omega), \vartheta + u(\vartheta; \omega)) \in \check{\mathcal{A}}_1 \times \mathcal{I}^{2n-2}_{\hat\delta/8} \times \mathbb{T}^{3n-2}$$

with $\vartheta \to \vartheta + u(\vartheta; \omega)$ a diffeomorphism of $\mathbb{T}^{3n-2}$, such that, on $\mathcal{T}_\omega$ the Hamiltonian flow is $\dot\vartheta = \omega$. Being $\mathcal{T}_\omega$ Lagrangian, the embedding $\phi(\cdot; \omega)$ can be lifted to a symplectic transformation $(y, \vartheta) \to \bar\phi(y, \vartheta; \omega)$ defined around $\mathcal{T}_\omega$ such that $\bar\phi(0, \vartheta; \omega) = \phi(\vartheta; \omega)$ which – since $\mathcal{T}_\omega = \phi(\mathbb{T}^{3n-2}; \omega) = \bar\phi(0, \mathbb{T}^{3n-2}; \omega)$ is invariant and is run with frequency ω – puts H in *Kolmogorov normal form*

$$K_\omega := H \circ \bar\phi(y, \vartheta, \omega) = c(\omega) + \omega \cdot y + Q(y, \vartheta; \omega) \tag{6.4}$$

namely, with $c(\omega)$ independent of ϑ and $Q(y, \vartheta; \omega) = {\rm O}(y^2)$. Note incidentally that the matrix $\int_{\mathbb{T}^{3n-2}} Q_{yy}\, d\vartheta$, being close to the block-diagonal matrix $\mathcal{Q}_0 = {\rm diag}\,[\partial^2 h_{\rm K} \circ \varpi^{-1}(\omega), \check\tau \circ \varpi^{-1}(\omega)]$, satisfies the so-called *Kolmogorov condition* to be not singular, which, together with (6.4), says that the tori of $\mathcal{K}$ are indeed *Kolmogorov tori*. From (6.4) using standard Averaging Theory (since ω is Diophantine), one sees that, if $|y| \le d = {\rm const}\, \frac{\gamma_2}{K^{\tau'+1}}$, one can conjugate K_ω to

$$K^{\rm av}_\omega = c(\omega) + \omega \cdot y + \bar Q(y; \omega) + \hat Q(y, \vartheta; \omega) \, .$$

where $\bar Q$ does not depend on ϑ and

$$|\hat Q(y, \vartheta)| \le {\rm const}\, d^2 e^{-\hat c K} = {\rm const}\, d^2 e^{-(\frac{c}{d})^{1/(\tau'+1)}} .$$

This implies the claim with $\sigma = 1/(\tau'+1)$ and $\sigma' = 2$.

Proof of Theorem 6.1. Let $\kappa > 0$ and $\vartheta \in (0,1)$. Let, also, ϵ, θ and μ be such that

$$\hat\epsilon < \epsilon < \epsilon^\star < \min\{\frac{3}{64}\delta^*, \ \frac{3}{64}\delta_*\} \ , \quad \theta \in (2,3] \ , \quad \Big(c\frac{64}{3}\vartheta\hat\epsilon\Big)^\theta < \mu < \Big(\frac{\frac{64}{3}\vartheta\hat\epsilon}{c}\Big)^{\frac{\theta}{\theta-1}} \ , \tag{6.5}$$

with c and δ_* to be defined below; finally, let

$$\hat\epsilon < \epsilon < \epsilon^\star \ , \quad \vartheta\hat\epsilon < \tilde\epsilon < \hat\epsilon \ , \quad \frac{\vartheta\hat\epsilon}{\tilde\epsilon} < \tilde\vartheta < 1 \ , \quad \check\epsilon^2 := \epsilon^2 + \hat\epsilon^2 - \tilde\epsilon^2 \tag{6.6}$$

---

[21]The set $\mathcal{D}_{\gamma_1, \gamma_2, \tau'}$ has been used for the first time in [2]. For $\gamma_1 = \gamma_2$ it corresponds to the usual Diophantine set.

[22]Compare [4, Theorem 1.4].

Note that, by the choice of $\epsilon^\star$ in (6.5), $\check\epsilon$ verifies $\epsilon<\check\epsilon<2\epsilon$.

Pick two positive numbers $\bar\gamma_0$ and $\bar\eta$, with $\bar\gamma_0^{-1}$ and $\bar\eta$ so small that

$$\bar\eta\le\frac12\,,\quad \Big(1+\frac{2}{\bar\gamma_0}+\bar\eta\Big)\vartheta\hat\epsilon\le\tilde\vartheta\tilde\epsilon\,,\quad \frac{nc}{\bar\gamma_0}\Big(\frac{64}{3}\vartheta\hat\epsilon\Big)^2\le\frac14(\hat\epsilon^2-\tilde\epsilon^2)\,,\tag{6.7}$$

and, moreover,

$$\frac{(\vartheta\bar\eta)^2}{4(n-1)}+\frac{c}{\bar\gamma_0}\frac{64}{3}\vartheta^2\Big(2(1+\bar\eta+\frac{1}{\bar\gamma_0})+\frac{c}{\bar\gamma_0}\frac{64}{3}\Big)<\kappa\,.\tag{6.8}$$

The number c in (6.5) and (6.8) will be defined below, independently of $\bar\eta$, $\bar\gamma_0$, θ, κ, $\hat\epsilon$ and ϵ. Note that, because of the definition of $\tilde\epsilon$ in (6.6), the numbers $\bar\gamma_0$ and $\bar\eta$ depend on ϑ, $\tilde\vartheta$, κ, but *not* upon $\hat\epsilon$ and, moreover, that the number $\bar\gamma_0$ can be chosen to be

$$\bar\gamma_0=\frac{\mathrm{const}\,(\vartheta)}{\kappa}\,.\tag{6.9}$$

Now, let $\check{\mathcal M}^{6n-4}_{\rm reg}:=\mathcal A\times\mathbb T^n\times B^{4(n-1)}_{\frac{64}{3}\vartheta\hat\epsilon}$; let $\check{\mathcal H}$ be as in Proposition 5.1 and let $\check{\mathcal H}_{\rm reg}\colon\check{\mathcal M}^{6n-4}_{\rm reg}\to\mathbb R$ be an *analytic* extension of $\check{\mathcal H}$ on $\check{\mathcal M}^{6n-4}_{\rm reg}$, namely a real-analytic Hamiltonian on $\check{\mathcal M}^{6n-4}_{\rm reg}$ such that

$$\check{\mathcal H}_{\rm reg}=\check{\mathcal H}=h_{\rm K}+\mu\check f\quad\text{on}\quad\check{\mathcal M}^{6n-4}_{\tilde\vartheta}(\tilde\epsilon,\check\epsilon)\,,\tag{6.10}$$

where, for $\tilde\vartheta$, $\tilde\epsilon$ and $\check\epsilon$ as in (6.6),

(6.11)
$$\check{\mathcal M}^{6n-4}_{\tilde\vartheta}(\tilde\epsilon,\check\epsilon):=\{\Lambda\in\mathcal A,|\check w|<\tilde\vartheta\tilde\epsilon,\tilde\epsilon<\varrho<\check\epsilon\}\times\mathbb T^n\subseteq\mathcal A\times\mathbb T^n\times B^{4(n-1)}_{\tilde\vartheta\tilde\epsilon}\subseteq\check{\mathcal M}^{6n-4}_{\rm reg}.$$

Since $\check f^{\rm av}$ is in $(2s)$-Birkhoff normal form (5.9) and the polynomial $\check{\rm P}_s=\check C_0+\check\Omega\cdot\check r+\frac12\check\tau\cdot\check r^2+\check{\mathcal P}_3+\cdots+\check{\mathcal P}_s$ is obviously analytic on $\check{\mathcal M}^{6n-2}_{\rm reg}$, we can choose $\check{\mathcal H}_{\rm reg}$ of the form $\check{\mathcal H}_{\rm reg}=h_{\rm K}+\mu\check f_{\rm reg}$ with $\check f^{\rm av}_{\rm reg}=\check{\rm P}_s+{\rm O}(|\check w|^{2s+1})$, having *the same normal form* $\check{\rm P}_s$ as $\check f^{\rm av}$.

By (6.10), all the motions of $\check{\mathcal H}_{\rm reg}$ which remain confined to $\check{\mathcal M}^{6n-4}_{\tilde\vartheta}(\tilde\epsilon,\check\epsilon)$ are indeed motions of $\check{\mathcal H}$.

Put $n_1:=n$, $n_2:=2(n-2)$, $H_0:=h_{\rm K}$, $P:=\check f_{\rm reg}$, $\rho_0:=\frac{\bar\gamma_0}{\bar c_0}\max\{\sqrt{\frac{\mu}{\hat\epsilon}},\sqrt{\hat\epsilon}\}$, $V:=\mathcal A^-_{\rho_0}$, $\bar\epsilon:=\frac{64}{3}\vartheta\hat\epsilon$, $a:=\frac{1}{2\theta(\tau+1)}$, where $\bar c_0$ will be defined below and $\mathcal A^-_{\rho_0}$ denotes the set $\{\Lambda\in\mathcal A:\ B_{\rho_0}(\Lambda)\subseteq\mathcal A\}$. Note that $\mathcal A^-_{\rho_0}$ is nonempty for small $\epsilon^\star$, because of the choice of μ in (6.5). Let $\epsilon_\star$ be as in Proposition D.1 in Appendix D and take, in (6.5), $\delta^\star:=\epsilon^\star$, so that, by the above choice of $\bar\epsilon$, $\bar\epsilon=\frac{64}{3}\vartheta\hat\epsilon<\frac{64}{3}\epsilon^\star<\epsilon_\star$; compare (D.2) in Proposition D.1. Note that

1. $\check f^{\rm av}_{\rm reg}$ has the same Birkhoff normal form as $\check f^{\rm av}$, hence, in particular, the first-order Birkhoff invariants are nonresonant;
2. that assumptions (D.2) of Proposition D.1 are trivially implied by (6.5) and the above choice of a and θ.

This allows to apply Proposition D.1 with n_1, n_2, H_0, P, $\cdots$ as above.

We then find suitable c_0, $c_\star$, $\rho_\star$, $\mathcal A_\star\subseteq\mathcal A^-_{\rho_0}\subseteq\mathcal A$, $\phi_\star$ as in the statement of Proposition D.1. Take in (6.5) and (6.8), $c:=c_\star$ and, in the definition of ρ_0,

$\bar{c}_0 := c_0$, so that $\rho_0 = \rho_\star$. Note also that: $\rho_\star := \frac{\bar{\gamma}_0}{c_0}\max\{\sqrt{\frac{\mu}{\hat{\epsilon}}},\ \sqrt{\hat{\epsilon}}\} \ge \hat{c}\sqrt{\hat{\epsilon}}$; by (D.7), the definition of ρ_0, the assumption on μ in (6.5) and, finally (6.9), $\mathcal{A}_\star$ is easily seen to satisfy (6.1); the transformation $\phi_\star$ acts as

$$\phi_\star\colon (\mathcal{A}_\star)_{\rho_\star} \times \mathbb{T}^n_{s_0/24} \times B^{4(n-1)}_{2\vartheta\hat{\epsilon}} \to (V_\star)_{31\rho_\star} \times \mathbb{T}^{n_1}_{s_0/6} \times B^{4(n-1)}_{\frac{64}{3}\vartheta\hat{\epsilon}} \tag{6.12}$$

and transforms $\check{\mathcal{H}}_{\rm reg}$ into $\mathcal{H}_\star := \check{\mathcal{H}}_{\rm reg} \circ \phi_\star$ with

$$\begin{aligned}\mathcal{H}_\star(\Lambda_\star, l_\star, w_\star)\\ = h_{\rm K}(\Lambda_\star) + \mu N_\star(\Lambda_\star, r_\star) + \mu\mathcal{P}_\star(\Lambda_\star, u_\star, v_\star) + \ \mu c_\star e^{-(\frac{1}{c_\star\mu})^a} f_\star(\Lambda_\star, l_\star, w_\star).\end{aligned} \tag{6.13}$$

In applying Proposition D.1, take in (D.5) $\gamma_0 = \bar{\gamma}_0$ and $\eta = \bar{\eta}$, where $\bar{\gamma}_0$, $\bar{\eta}$ satisfy (6.7)–(6.8) above, with $c = c_\star$. By (D.5), the transformation $\phi_\star$ satisfies

$$\phi_\star\Big((\mathcal{A}_\star)_{\rho_\star/2} \times \mathbb{T}^n_{s_0(1+\frac{1}{\bar{\gamma}_0})/48} \times B^{4(n-1)}_{\vartheta\hat{\epsilon}(1+\frac{1}{\bar{\gamma}_0})}\Big) \supseteq (\mathcal{A}_\star)_{\rho_\star/4} \times \mathbb{T}^n_{s_0/48} \times B^{4(n-1)}_{\vartheta\hat{\epsilon}} \tag{6.14}$$

and, by the first inequality in (6.7),

$$\begin{aligned}\phi_\star\Big((\mathcal{A}_\star)_{\rho_\star(1+\bar{\eta})/2} \times \mathbb{T}^n_{s_0(1+\bar{\eta}+\frac{1}{\bar{\gamma}_0})/48} \times B^{4(n-1)}_{\vartheta\hat{\epsilon}(1+\bar{\eta}+\frac{1}{\bar{\gamma}_0})}\Big)\\ \subseteq (\mathcal{A}_\star)_{3\rho_\star/4+\rho_\star\bar{\eta}/2} \times \mathbb{T}^n_{s_0(1+\bar{\eta}+\frac{2}{\bar{\gamma}_0})/48} \times B^{4(n-1)}_{\tilde{\vartheta}\tilde{\epsilon}}\ .\end{aligned} \tag{6.15}$$

Let ϑ, $\hat{\epsilon}$ and ϵ be as in (6.5) and define the set

$$\check{\mathcal{M}}^{6n-4}_{\star\vartheta}(\hat{\epsilon},\epsilon) := \{\Lambda \in \mathcal{A}_\star\ ,\ |\check{w}| < \vartheta\hat{\epsilon}\ ,\ \hat{\epsilon} < \varrho < \epsilon\} \times \mathbb{T}^n\ ; \tag{6.16}$$

note that $\check{\mathcal{M}}^{6n-4}_{\star\vartheta}(\hat{\epsilon},\epsilon) \subseteq \mathcal{A}_\star \times \mathbb{T}^n \times B^{4(n-1)}_{\vartheta\hat{\epsilon}} \subseteq \check{\mathcal{M}}^{6n-4}_{\rm reg}$.

From the above definitions (see (6.6), (6.7), (6.11)) the following inclusions follow

$$\check{\mathcal{M}}^{6n-4}_{*\vartheta}(\hat{\epsilon},\epsilon) \subseteq \check{\mathcal{M}}^{6n-4}_{\tilde{\vartheta}}(\tilde{\epsilon},\check{\epsilon}) \subseteq \check{\mathcal{M}}^{6n-4}_{\rm reg}\ .$$

We prove that motions of $\check{\mathcal{H}}_{\rm reg}$ with initial data $(\Lambda(0), \check{l}(0), \check{w}(0))$ in $\check{\mathcal{M}}^{6n-4}_{\star\vartheta}(\hat{\epsilon},\epsilon)$ remain in $\check{\mathcal{M}}^{6n-4}_{\tilde{\vartheta}}(\tilde{\epsilon},\check{\epsilon})$ for $|t| \le \mathfrak{t}$. At the end, to obtain the statement of the theorem, we shall take $\theta = 3$, $\vartheta = 1/4$,$\tilde{\epsilon} = \hat{\epsilon}/2$ and $\tilde{\vartheta} = 3/4$.

Consider now motions of $\check{\mathcal{H}}_{\rm reg}$ with initial data in $\check{\mathcal{M}}^{6n-4}_{\star\vartheta}(\hat{\epsilon},\epsilon)$. Taking the real part in (6.14), all such motions are the $\phi_\star$-images of some subset of motions of $\mathcal{H}_\star$ with initial data $(\Lambda_\star(0), l_\star(0), w_\star(0)) \in (\mathcal{A}_\star)_{\rho_\star/2} \times \mathbb{T}^n \times B^{4(n-1)}_{\vartheta\hat{\epsilon}(1+\frac{1}{\bar{\gamma}_0})}$.

Using (6.3), (6.13), one finds that for $|t| \le \mathfrak{t}$, the actions $\Lambda_\star(t)$ and $r_\star(t)$, where $r_\star := \frac{u_\star^2+v_\star^2}{2}$, with $w_\star = (u_\star, v_\star)$, satisfy, for a possibly smaller value of

$\epsilon^\star$,

$$
\begin{aligned}
|(\Lambda_\star)_i(t)-(\Lambda_\star)_i(0)| &\le \mu c_\star e^{-(\frac{1}{c_\star\mu})^a}|t| \\
&\le \mu c_\star e^{-(\frac{1}{c_\star\mu})^a}\mathfrak{t} \le \frac{c_\star}{c^\star}\frac{e^{-(\frac{1}{c_\star})^a(\frac{c_\star}{\hat\epsilon})^{\theta/(\theta-1)}}}{\hat\epsilon^{2s-1}} \\
&\le \min\{\hat c\frac{\bar\eta}{2}\sqrt{\hat\epsilon}, \frac{\hat\epsilon^2-\tilde\epsilon^2}{4n}\} \\
&\le \min\{\frac{\bar\eta}{2}\rho_\star, \frac{\hat\epsilon^2-\tilde\epsilon^2}{4n}\}\ .
\end{aligned}
\tag{6.17}
$$

Similarly, taking the derivatives of (6.13) with respect to $w_\star=(u_\star,v_\star)$ and using that, on the domain of $\phi_\star$ in (6.12), $|\mathscr{P}_\star|\le\bar c(2\vartheta\hat\epsilon)^{2s+1}$, for some constant $\bar c$ depending only on $\mathscr{P}$, one finds that, for a possibly larger value of $c^\star$ in (6.3),

$$
|(r_\star)_j(t)-(r_\star)_j(0)|\le\mu\Big(\bar c(2\vartheta\hat\epsilon)^{2s+1}+c_\star e^{-(\frac{1}{c_\star\mu})^a}\Big)\mathfrak{t}\le\frac{(\vartheta\hat\epsilon\bar\eta)^2}{4(n-1)}\ .
\tag{6.18}
$$

(6.17)–(6.18) imply that for $|t|\le\mathfrak{t}$, the motion $t\to(\Lambda_\star(t),l_\star(t),w_\star(t))$ remains confined to the set $(\mathscr{A}_\star)_{\rho_\star(1+\bar\eta)/2}\times\mathbb{T}^n\times B^{4(n-1)}_{\vartheta\hat\epsilon(1+\bar\eta+\frac{1}{\bar\gamma_0})}$. In particular,

$$
|w_\star|_\infty\le|w_\star|_2<\vartheta\hat\epsilon\Big(1+\bar\eta+\frac{1}{\bar\gamma_0}\Big)\ .
\tag{6.19}
$$

By (6.15) and the fact that $\bar\eta\le\frac12$, the $\phi_\star$-images $t\to(\Lambda(t),\check l(t),\check w(t))$ of such motions remain confined to $(\mathscr{A}_\star)_{3\rho_\star/4+\rho_\star\bar\eta/2}\times\mathbb{T}^n\times B^{4(n-1)}_{\bar\vartheta\bar\epsilon}\subseteq(\mathscr{A}_\star)_{\rho_\star}\times\mathbb{T}^n\times B^{4(n-1)}_{\bar\vartheta\bar\epsilon}$. We now prove that such trajectories are confined to $\check{\mathscr{M}}^{6n-4}_{\bar\vartheta}(\tilde\epsilon,\check\epsilon)$, and hence, by (6.10), they are actually motions of $\check{\mathscr{H}}$. By the definition of $\check{\mathscr{M}}_{\bar\vartheta}(\tilde\epsilon,\check\epsilon)$, we have to prove that

$$
\tilde\epsilon<\varrho(\Lambda(t),G)<\check\epsilon\ ,\qquad\forall\ |t|\le\mathfrak{t}\ .
\tag{6.20}
$$

Using (D.6), (6.7) and that, by (6.5), $\mu<(\hat\epsilon/c_\star)^{3/2}$, one finds the following bound for the Λ-projection of $\phi_\star$:

$$
|\Lambda-\Lambda_\star|_1\le\frac{nc_\star}{\bar\gamma_0}\mu^{a/2}\sqrt{\mu\frac{64}{3}\vartheta\hat\epsilon}\le\frac14(\hat\epsilon^2-\tilde\epsilon^2)\ .
$$

By this inequality and the first bound in (6.17), we have

$$
2|\Lambda(t)-\Lambda(0)|_1\le2|\Lambda_\star(t)-\Lambda_\star(0)|_1+2\sup|\Lambda_\star-\Lambda|_1\le\hat\epsilon^2-\tilde\epsilon^2\ ,
$$

proving the first inequality in (6.2). Moreover, since, by (6.16), $\hat\epsilon<\varrho(\Lambda(0),G)<\epsilon$,

$$
\begin{aligned}
\tilde\epsilon^2&=\hat\epsilon^2-(\hat\epsilon^2-\tilde\epsilon^2)\\
&<\varrho(\Lambda(0),G)^2+2|\Lambda(t)|_1-2|\Lambda(0)|_1\\
&=\varrho(\Lambda(t),G)<\epsilon^2+\hat\epsilon^2-\tilde\epsilon^2=\check\epsilon^2\ ,
\end{aligned}
$$

which proves (6.20). To conclude, it remains to prove the bound in (6.2) for the actions $\check r_j$.

Assumption (6.5) and the bounds in (D.6) imply that $w_\star$ and $\check{w}$ are at most at distance

$$|\check{w} - w_\star|_\infty \le \frac{c_\star}{\bar{\gamma}_0}\frac{64}{3}\vartheta\hat{\epsilon}\ . \tag{6.21}$$

It follows from (6.19) and (6.21) that

$$|\check{w}(t)|_\infty \le |w_\star|_\infty + |w_\star(t) - \check{w}(t)|_\infty < \vartheta\hat{\epsilon}(1+\bar{\eta}+\frac{1}{\bar{\gamma}_0}) + \frac{c_\star}{\bar{\gamma}_0}\frac{64}{3}\vartheta\hat{\epsilon},$$

giving finally, by (6.8) and (6.18),

$$\begin{aligned}|\check{r}(t) - \check{r}(0)|_\infty &\le |r_\star(t) - (r_\star)(0)|_\infty + |\check{w} - w_\star|_\infty(|w_\star|_\infty + |\check{w}|_\infty)\\ &\le \frac{(\vartheta\hat{\epsilon}\bar{\eta})^2}{4(n-1)} + \frac{c_\star}{\bar{\gamma}_0}\frac{64}{3}(\vartheta)^2(\hat{\epsilon})^2\Big(2(1+\bar{\eta}+\frac{1}{\bar{\gamma}_0}) + \frac{c_\star}{\bar{\gamma}_0}\frac{64}{3}\vartheta\Big)\\ &\le \kappa\hat{\epsilon}^2 < \hat{\epsilon}^2\ .\end{aligned}$$

□

Theorem 6.1 actually implies stability of eccentricities $e_1, \cdots, e_n$ and of the mutual inclinations $\hat{\imath}_1, \cdots, \hat{\imath}_{n-2}$, where e_i and $\hat{\imath}_j$ are defined as[23]

$$e_i = \sqrt{1 - (\frac{|\mathrm{C}^{(i)}|}{\Lambda_i})^2}\ , \quad \cos\hat{\imath}_j = \frac{\mathrm{C}^{(j+1)}\cdot S^{(j)}}{|\mathrm{C}^{(j+1)}||S^{(j)}|}\ ,$$

$\mathrm{C}^{(j+1)}$ and $S^{(j)}$ being as in (3.2). Indeed, we have the following

COROLLARY 6.3. *For any $c>0$, there exists $C>0$ such that, for all motions starting in the set $\mathcal{M}_\star$ of Theorem 6.1, e_i and $\hat{\imath}_j$ satisfy*

$$\max\{|e_i(t) - e_i(0)|,\ |\hat{\imath}_j(t) - \hat{\imath}_j(0)|\} \le c\epsilon\ , \quad \forall\ |t| \le \frac{C}{\mu\epsilon^{2s-1}}\ . \tag{6.22}$$

Proof. For ease of computations, we shall consider the functions

$$\mathfrak{e}_i := e_i^2 \quad \text{and} \quad \mathfrak{i}_j := 1 - \cos^2\hat{\imath}_j$$

and we shall check that, for any $\bar{c}>0$, one has

$$\max\{|\mathfrak{e}_i(t) - \mathfrak{e}_i(0)|,\ |\mathfrak{i}_j(t) - \mathfrak{i}_j(0)|\} \le \bar{c}\epsilon^2\ , \tag{6.23}$$

which implies, clearly, (6.22). The proof of (6.23) comes from the relation between $\mathfrak{e}_i$, $\mathfrak{i}_j$ and the variables $(\Lambda, \check{l}, \check{w})$; in particular, on how $\mathfrak{e}_i$ and $\mathfrak{i}_j$ are related to the stable actions $\Lambda_1, \cdots, \Lambda_n, \check{r}_1, \cdots, \check{r}_{2n-2}$.

Recall that the RPS variables $(\Lambda, \lambda, \bar{z})$ are related to the variables $(\Lambda, \check{l}, \check{w})$ by $(\Lambda, \lambda, \bar{z}) = \phi(\Lambda, \check{l}, \check{w})$ with

$$\phi := \tilde{\phi}\circ\check{\phi}\circ\hat{\phi}\circ\check{\phi} \tag{6.24}$$

[23] Note that in the completely reduced setting the number of *independent* inclinations is $(n-2)$. Indeed, the *overall* inclination of C has no physical meaning by rotation invariance and the inclination $\hat{\imath}_{n-1}$ between $S^{(n-1)}$ and $\mathrm{C}^{(n)}$ is a function of $\Lambda_1, \cdots, \Lambda_n, e_1, \cdots, e_n, \hat{\imath}_1, \cdots, \hat{\imath}_{n-2}$ and G.

where $\tilde\phi$, $\breve\phi$ and $\hat\phi$ are as in § 5.1 and where we have denoted by $\check\phi$ the $(6n-2)$-dimensional transformation obtained from the $(6n-4)$-dimensional transformation ϕ_s given by Proposition 5.1, lifted on G and $\hat g$ in the obvious way (see the proof of Theorem 5.1 in Appendix C). Let us remark the following facts:

(i) The transformation $\tilde\phi$ in (6.24), is defined in (4.8). Its Λ-projection is the identity and, we claim, its $\bar z$-projection of $\tilde\phi$ is $\Lambda_2^{-5/2}$-close to the identity. Indeed, such projection is defined by the matrices $\mathrm{U_h}$ and $\bar{\mathrm{U}}_\mathrm{v}$ in (4.8), which make the quadratic part in (4.4) diagonal. By induction: For $n=2$, $\bar{\mathrm{Q}}_\mathrm{v}$ is of order 1, so $\bar{\mathrm{U}}_\mathrm{v}=1$, and $\mathrm{Q_h}$ is 2×2. Its explicit expression can be found in [6, Appendix B]. Using such expression one readily checks that, for $n=2$, $\mathrm{U_h}$ is actually $\Lambda_2^{-5/2}$-close to the identity. For $n\ge 2$, as proven in [6, Eq. (8.10), with δ just after Eq. (7.7)], the matrices $\mathrm{U_h^+}$ and $\bar{\mathrm{U}}_\mathrm{v}^+$ at rank n are related to the corresponding ones $\mathrm{U_h}$ and $\bar{\mathrm{U}}_\mathrm{v}$ at rank $(n-1)$ by $\mathrm{U_h^+}=\mathrm{diag}\,[\mathrm{U_h},1]+\mathrm{O}(\Lambda_n^{-6})$, $\mathrm{U_v^+}=\mathrm{diag}\,[\mathrm{U_v},1]+\mathrm{O}(\Lambda_n^{-6})$ and the claim follows.

(ii) $\breve\phi$ is the Birkhoff transformation defined in (4.9) which acts as the identity on Λ (Appendix A), and is $\mathrm{O}(|\bar w|^3)$-close to the identity in the $\bar w$-variables (parity). By items (iii) and (iv) below, the projection $\Pi_{\bar z}\circ(\breve\phi\circ\hat\phi\circ\check\phi)$ is ϵ^3-close to the identity, where ϵ is any number such that $\varrho(\Lambda,G)<\epsilon$;

(iii) $\hat\phi$ is explicitly given in (5.3); recall that the Euclidean length $|\bar w|^2$ is sent into $\varrho(\Lambda,G)^2$, with $\varrho(\Lambda,G)^2$ as in (5.4);

(iv) $\check\phi$ is constructed in (the proof of) Proposition 5.1. In particular, it leaves (Λ,G) fixed and is ϵ^{2s+1}-close to the identity in $\check w$;

(v) In terms of the RPS variables $(\Lambda,\lambda,\bar z)$, the functions $\mathfrak{e}_i=\mathfrak{e}_i(\Lambda,\rho,\mathrm{r})$, $\mathfrak{i}_j=\mathfrak{i}_j(\Lambda,\rho,\mathrm{r})$ are *rational functions* of Λ_i and of $\rho_i:=\dfrac{\eta_i^2+\xi_i^2}{2}$ and $\mathrm{r}_j:=\dfrac{p_j^2+q_j^2}{2}$ explicitly given by

$$\mathfrak{e}_i=\frac{\rho_i}{\Lambda_i}\Big(2-\frac{\rho_i}{\Lambda_i}\Big)\,,\quad \mathfrak{i}_j=2\mathrm{r}_j\mathfrak{c}_{j+1}\,,\quad \mathfrak{c}_{j+1}:=\frac{2\mathscr{L}_j-|z_{j-1}|^2-\mathrm{r}_j}{2(\Lambda_{j+1}-\rho_{j+1})(2\mathscr{L}_{j+1}-|z_j|^2)}$$

where

$$\mathscr{L}_i:=\sum_{1\le j\le i}\Lambda_j,\quad z_i=(\eta_1,\dots,\eta_{i+1},\xi_1,\dots,\xi_{i+1},p_1,\dots,p_i,q_1,\dots,q_i).$$

Such expressions may be found from (3.7) above; compare also [6, Appendix A.2], for more details.

From (i)–(v) above there follows that $\mathfrak{e}_i$, $\mathfrak{i}_j$, expressed in the variables $(\Lambda,\check r)$ have the form, respectively $\mathfrak{e}_i(\Lambda,\check r)+\tilde{\mathfrak{e}}_i(\Lambda,\check r)$, $\mathfrak{i}_j(\Lambda,\check r)+\tilde{\mathfrak{i}}_j(\Lambda,\check r)$ where $\tilde{\mathfrak{e}}_i$, $\tilde{\mathfrak{i}}_j$ are functions of order $\mathrm{O}(\epsilon^2\Lambda_2^{-5/2}+\epsilon^3)$. This, by (6.2), implies (6.23) and hence (6.22). □

APPENDIX A. BIRKHOFF NORMAL FORMS AND SYMMETRIES

In this appendix we analyze the properties of Birkhoff-normalizations $\breve\phi$ used in (4.9) for, respectively, partial and total reduction in case of symmetries.

Let us consider[24] again the transformation $\mathscr{R}_g$, $\mathscr{R}_{1\leftrightarrow 2}$ and $\mathscr{R}_3^-$ in (3.19)–(3.20), but generalized replacing λ, η, ξ, p, q with $\tilde\lambda \in \mathbb{T}^n$, $(\tilde\eta,\tilde\xi) \in \mathbb{R}^{2m_1}$, $(\tilde p,\tilde q) \in \mathbb{R}^{2m_2}$, for some n, m_1 and $m_2 \in \mathbb{N}$. Put $m := m_1 + m_2$. Let $\mathscr{A}$ be an open, bounded set of parameters in $\mathbb{R}^n$; consider a function f: $\mathscr{A} \times B^{2m}_{\tilde\epsilon} \to \mathbb{R}$ of the form of $\tilde f^{\rm av}_{\rm RPS}$ in (4.7), with the numbers n, $n-1$ into the summands replaced by m_1, m_2.

PROPOSITION A.1. *Let* f *be* $\mathscr{R}_g$, $\mathscr{R}_{1\leftrightarrow 2}$ *and* $\mathscr{R}_3^-$*-invariant. Assume that the first-order Birkhoff invariants* $\bar\Omega = (\sigma,\bar\varsigma)$ *satisfy, for some integer* s*,*

$$\inf_{\mathscr{A}} |\bar\Omega\cdot k| > 0, \quad \forall\ k \in \mathbb{Z}^m:\ \sum_{i=1}^m k_i = 0\ , \quad 0 < |k|_1 := \sum_{i=1}^m |k_j| \le 2s\ .$$

Then there exists $0 < \check\epsilon \le \tilde\epsilon$ *and a symplectic transformation*

$$\check\phi:\ (\Lambda,\check\lambda,\check z) = (\Lambda,\check\lambda,(\check\eta,\check\xi,\check p,\check q)) \in \mathscr{A}\times\mathbb{T}^n\times B^{2m}_{\check\epsilon} \to (\Lambda,\tilde\lambda,\tilde z) \in \mathscr{A}\times\mathbb{T}^n\times B^{2m}_{\tilde\epsilon}$$

which puts f *into Birkhoff normal form up to the order* $2s$*. Furthermore,* $\check\phi$ *leaves the* Λ*-variables unchanged, acts as a* $\check\lambda$*-independent shift on* λ*, is* $\check\lambda$*-independent on the remaining variables, preserves the function* $G(\Lambda,\check z) := |\Lambda|_1 - |\check z|_2^2/2$ *and finally verifies*

$$\check\phi\circ\mathscr{R} = \mathscr{R}\circ\check\phi \tag{A.1}$$

for any $\mathscr{R} = \mathscr{R}_g$, $\mathscr{R}_{1\leftrightarrow 2}$, $\mathscr{R}_3^-$*. Moreover,* (A.1) *holds for any of such* $\check\phi$*'s.*

REMARK A.2.

(i) Since $\check\phi$ commutes with $\mathscr{R}_3^-$, its $(\tilde p,\tilde q)$-projection

$$\Pi_{(\tilde p,\tilde q)}\check\phi = (\check p,\check q) + \mathrm{O}(|(\check p,\check q)|^3)$$

is odd in $(\check p,\check q)$; its $(\tilde\eta,\tilde\xi)$ and $\tilde\lambda$-projections

$$\Pi_{\tilde\lambda}\check\phi = \check\lambda + \check\varphi(\Lambda,\check z)\ , \qquad \Pi_{(\tilde\eta,\tilde\xi)}\check\phi = (\check\eta,\check\xi) + \mathrm{O}(|(\check\eta,\check\xi)|^3)$$

are even in $(\check p,\check q)$. Using also the commutation with $\mathscr{R}_\pi$, one finds that the $(\tilde\eta,\tilde\xi)$-projection of $\check\phi$ is odd in $(\check\eta,\check\xi)$.

(ii) It is not difficult to derive $\mathscr{R}_g$, $\mathscr{R}_{1\leftrightarrow 2}$ and $\mathscr{R}_3^-$-invariance of $\tilde f^{\rm av}_{\rm RPS}$ from that of $f_{\rm RPS}$ in (4.1) (or see the comments between [6, Eq. (7.24) and Eq. (7.25)]).

(iii) Proposition A.1 is closely related[25] to [6, Proposition 7.3]. The difference being that, in [6], (A.1) was proven only for $\mathscr{R}_g$. To extend the proof in [6], we briefly recall the setting, referring to [6] for full details.

Proof of Proposition A.1. We recall that $\check\phi$ can be constructed in $s-1$ steps, as a product $\phi_2\circ\cdots\circ\phi_{2s-2}$. The first step is as follows. To uniform notations, put

[24] Clearly, Proposition A.1 below is general. However, to avoid to introduce too many symbols, we use notations (i.e., $\mathscr{A}$, n, $\bar\Omega = (\sigma,\bar\varsigma)$, $\check\epsilon$, $\check\phi$, $\tilde\lambda$, $\tilde z = (\tilde\eta,\tilde\xi,\tilde p,\tilde q)$, $\check\lambda$, $\check z = (\check\eta,\check\xi,\check p,\check q)$) already used in the paper, which make the application transparent: compare the second item in Remark A.2 below.

[25] In [6] $\mathscr{A}$, $\bar\Omega$, Λ, $\tilde\lambda$, $\tilde z$, $\check\lambda$, $\check z$, $\mathscr{R}_g$, $\check\epsilon$, $\tilde\epsilon$ are denoted $\mathscr{B}$, Ω, I, φ, w, $\check\varphi$, $\check w$, $\mathscr{R}^g$, $\check r$, r, respectively.

$w = (u, v) := \big((\tilde\eta, \tilde p), (\tilde\xi, \tilde q)\big)$. One introduces the "Birkhoff coordinates"

$$(t, t^*) = \big((t_1, \dots, t_m), (t_1^*, \dots, t_m^*)\big) : \begin{cases} t_j = \frac{u_j - \mathrm{i} v_j}{\sqrt{2}} \\ t_j^* = \frac{u_j + \mathrm{i} v_j}{\sqrt{2}\mathrm{i}} \end{cases} .$$

Consider then the polynomial of degree 4 (f is even in w since it is $\mathcal{R}_g$-invariant) into the expansion of f in powers of w:

$$\mathcal{P}_4 = \sum_{|\alpha|_1 + |\alpha^*|_1 = 4} c^{(4)}_{\alpha,\alpha^*} \prod_{1\le j\le m} t_j^{\alpha_j} t_j^{*\alpha_j^*} . \tag{A.2}$$

Let ϕ_2 be the time-one flow generated by the Hamiltonian

$$K_4\big(\Lambda, (t, t^*)\big) = \sum_{|\alpha|_1 = |\alpha^*|_1 = 2} \frac{c^{(4)}_{\alpha,\alpha^*}}{\mathrm{i}\bar\Omega \cdot (\alpha - \alpha^*)} \prod_{1\le j\le n} t_j^{\alpha_j} t_j^{*\alpha_j^*} . \tag{A.3}$$

Since f is $\mathcal{R}_g$-invariant, K_4 is so, hence G is an integral for the K_4-flow; taking this flow at time $\theta = 1$, we have that ϕ_2 preserves G. Note that f being $\mathcal{R}_{1\leftrightarrow 2}$ – invariant implies that the coefficients $c^{(4)}_{\alpha,\alpha^*}$ in (A.2) satisfy $c^{(4)}_{\alpha,\alpha^*} = c^{(4)}_{\alpha^*,\alpha}$. So, the function K_4 in (A.3) is skew-symmetric in (t, t^*): $K_4\big(\Lambda, (t, t^*)\big) = -K_4\big(\Lambda, (t^*, t)\big)$. Writing the motion equations of K_4 with initial datum $(\Lambda, \pi/2 - \lambda, t^*, t)$, the claim follows. The function $\mathrm{f}_2 := \mathrm{f} \circ \phi_2 = \mathrm{f}(\Lambda, \cdot) \circ \mathrm{Z}_2(\Lambda, \cdot)$ where $\mathrm{Z}_2(\Lambda, \cdot)$ is the projection on (t, t^*) of ϕ_2, is now in normal form of order 4 and it is easy to seen to be again $\mathcal{R}_{1\leftrightarrow 2}$ – invariant; so that the procedure can be iterated. The commutation with $\mathcal{R}_3^-$ is proved similarly. The (standard) proof of independence of (A.1) upon the choice of $\breve\phi$ is omitted. □

APPENDIX B. DOMAINS OF POINCARÉ AND RPS VARIABLES

In this appendix, for completeness, we describe analytically the global domains $\mathcal{M}^{6n}_{\max,\mathrm{P}}$, $\mathcal{M}^{6n}_{\max,\mathrm{RPS}}$.

- The domain $\mathcal{M}^{6n}_{\max,\mathrm{P}}$ is the subset of $(\Lambda, \lambda, \mathrm{z}) \in \mathbb{R}^n_+ \times \mathbb{T}^n \times \mathbb{R}^{4n}$ where their respective action variables satisfy

$$0 < \Gamma_i \le \Lambda_i \ , \quad -\Gamma_i < \Theta_i \le \Gamma_i$$

  where the action variables Γ_i, Θ_i are regarded as functions of the Poincaré variables in (3.6) i.e.,

$$\Gamma_i = \Lambda_i - \frac{\eta_i^2 + \xi_i^2}{2} \ , \quad \Theta_i = \Lambda_i - \frac{\eta_i^2 + \xi_i^2}{2} - \frac{\mathrm{p}_i^2 + \mathrm{q}_i^2}{2}$$

- The domain $\mathcal{M}^{6n}_{\max,\mathrm{RPS}}$ is the subset of $(\Lambda, \lambda, z) \in \mathbb{R}^n_+ \times \mathbb{T}^n \times \mathbb{R}^{4n}$ where[26] the action variables satisfy

$$\begin{cases} 0 < \Gamma_i \le \Lambda_i \ , & 1 \le i \le n \ , \\ |\Psi_{i-1} - \Gamma_{i+1}| < \Psi_i \le \Psi_{i-1} + \Gamma_{i+1} \ , & 1 \le i \le n-1 \ , \\ -\Psi_{n-1} < \Psi_n \le \Psi_{n-1} \ . \end{cases} \tag{B.1}$$

[26] Recall: $\Gamma_i = |\mathrm{C}^{(i)}| = \Lambda_i \sqrt{1 - e_i^2}$; $\Psi_{n-1} = |C|$; $\Psi_n := \mathrm{C}_3 = \mathrm{C} \cdot k^{(3)}$; $\Psi_i = |S^{(i+1)}| = |S^{(i)} + \mathrm{C}^{(i+1)}|$.

Here, Γ_i, Ψ_i are regarded as functions of the RPS-variables as in (3.7):

$$\text{(B.2)}\qquad \begin{cases} \Gamma_i = \Lambda_i - \frac{\eta_i^2+\xi_i^2}{2}\,, & 1 \le i \le n\,,\\ \Psi_i = \sum_{j=1}^{i+1}\Lambda_j - \sum_{j=1}^{i+1}\frac{\eta_j^2+\xi_j^2}{2} - \sum_{j=1}^{i}\frac{p_j^2+q_j^2}{2}\,, & 1\le i\le n-1\,,\\ \Psi_n = \Psi_{n-1} - \frac{p_n^2+q_n^2}{2}\,. \end{cases}$$

Note in particular that the only inequality in (B.1) involving (p_n, q_n) is the third one. Using (compare (B.2))

$$\Psi_n + \frac{p_n^2+q_n^2}{2} = \Psi_{n-1} = |\mathrm{C}| = G(\Lambda, z) = |\Lambda|_1 - \frac{|z|_2^2}{2}$$

one has that such inequality is just the second one in (3.12), i.e., $\sqrt{p_n^2+q_n^2} < 2\sqrt{G}$. The set $\mathcal{M}^{6n-2}_{\max}$ in (3.12) is then defined by the first two inequalities into (B.1), with $\Gamma_1, \cdots, \Gamma_n, \Psi_1, \cdots, \Psi_{n-1}$ functions of Λ and z as in (B.2).

APPENDIX C. PROOF OF PROPOSITION 5.1

The proof is obtained as a generalization of [6, Proposition 10.1]: in [6] the proof is divided into four steps, and here we just remark how to modify such steps, in order to get the generalization at arbitrary order. For the purpose of this proof we shall use the notations adopted in [6], which we now recall. The variables $(\Lambda, \hat{l}, \hat{w}) = \big(\Lambda, \hat{l}, (\hat{u}, \hat{w})\big)$ defined in (5.3) are denoted there by $(\Lambda, \hat{\lambda}, \hat{z})$, with again $\hat{z} = (\hat{u}, \hat{v})$. The variables $\check{r}_1, \cdots, \check{r}_{2n-2}$ correspond to $\check{\mathrm{R}}_1, \cdots, \check{\mathrm{R}}_{2n-2}$ in [6, Proposition 10.1]. Moreover, in [6], the variables (Λ, G), $(\hat{\lambda}, \hat{g})$ are called I, $\hat{\varphi}$, respectively, and the same convention is next used during the proof: $\varphi^\star$, φ^*, $\check{\varphi}$ are names for $(\lambda^\star, g^\star)$, and so on. Note also that functions $\hat{\mathcal{H}}$, $\hat{f}$, in (5.6) and the function $\mathcal{P}$ in (5.1) for $2s = 4$ are called, in [6], $\hat{\mathcal{H}}_{\mathrm{G}}$, $\hat{f}_{\mathrm{G}}$, $\check{\mathcal{P}}$, while the average $\hat{f}^{\mathrm{av}}$ is denoted $\hat{f}_{\mathrm{G,av}}$, compare [6, Eqs. (7.30), (9.7), (10.1)].

Step 1. Fix $s \in \mathbb{N}$, $\vartheta \in (0,1)$. We shall prove Proposition 5.1 with ϑ at the place of $1/4$ in (5.8); at the end we shall take $\vartheta = 1/4$. Let $\eta \in (0,1)$ be so small such that the number $\vartheta + 2s\eta$ is still in $(0,1)$, $\delta^* < \varepsilon$, where ε is as in (5.5).

Take the number θ in [6, Eq. (10.15)] to be $\theta := \vartheta + 2s\eta$. Replace the function $f^{(\varrho)}$ defined just after [6, Eq. (10.16)] by the function

(C.1)
$$f^{(\varrho)} = \hat{C}_0(\Lambda, \varrho) + \varrho^2\Big(\hat{\Omega}\cdot\check{\mathrm{R}} + \frac{\varrho^2}{2}\hat{\tau}(\Lambda)\cdot\check{\mathrm{R}}^2 + \hat{\mathcal{P}}_3(\check{\mathrm{R}};\Lambda) + \cdots + \hat{\mathcal{P}}_s(\check{\mathrm{R}};\Lambda) + \varrho^{2s}\mathcal{Q}(\Lambda, \check{z}, \varrho)\Big).$$

where[27]

$$f^{(\varrho)} - \varrho^{2s+2}\mathcal{Q} = \hat{C}_0(\Lambda, \varrho) + \varrho^2\Big(\hat{\Omega}\cdot\check{\mathrm{R}} + \frac{\varrho^2}{2}\hat{\tau}(\Lambda)\cdot\check{\mathrm{R}}^2 + \cdots + \mathcal{P}_s(\check{\mathrm{R}};\Lambda)\Big)$$

[27] Note incidentally that the monomials $\hat{\mathcal{P}}_1 := \hat{\Omega}\cdot\check{\mathrm{R}}$, $\hat{\mathcal{P}}_2 := \frac{1}{2}\hat{\tau}\check{\mathrm{R}}^2, \cdots, \hat{\mathcal{P}}_s$ in (C.1) are related to the corresponding monomials $\bar{\mathcal{P}}_1 := \bar{\Omega}\cdot\bar{r}$, $\bar{\mathcal{P}}_2 := \frac{1}{2}\bar{\tau}\bar{r}^2, \cdots, \bar{\mathcal{P}}_s$ in (5.1) simply replacing in $\bar{\mathcal{P}}_j$ $\bar{r}_i$ with $\check{r}_i$ for $i \ne 2n-1$ and $\bar{r}_{2n-1}$ with $\varrho^2 - \sum_1^{2n-2}\check{r}_j$. Such invariants may be taken to be, up to $\mathrm{O}(\varrho^2)$, as the first approximation of the invariants $\check{\Omega}$, $\check{\tau}, \cdots, \check{\mathcal{P}}_s$ in (5.9).

is a polynomial in the variables $\check{\mathrm{R}}_i = \frac{\check{u}_i^2+\check{v}_i^2}{2}$, which is of degree $2s$ in $(\check{u}, \check{v})$. Next, comparing to [6, Eq. (10.26)], the remainder $\alpha^{2s+2}\mathcal{Q}$ in (C.1) is

$$\alpha^{2s+2}\mathcal{Q}(\Lambda, \check{z}, \alpha) = \mathcal{P}(\Lambda, \alpha\check{\phi}_{\check{z}}^{(1)}(\check{z})) ,$$

with $\mathcal{P}$ as in (5.1) and, quite analogously to [6, Eq. (10.17)], $\check{\phi}_{\check{z}}^{(1)}$ denotes the projection on $\bar{z}$ of the transformation (5.3) with $\hat{g} = 0$, ϱ replaced by 1 and $\bar{w}$ replaced by $\bar{z}$. Note that the functions $\hat{\Omega}$ and $\hat{\tau}$ are ϱ^2-close to the functions defined in [6, Eq. (10.6)–(10.7)]. In particular, $\hat{\Omega}$ do not satisfy resonances up to order $2s$, for small δ^*. Replace then the definition of the function F just before [6, Eq. (10.19)] with

$$\begin{aligned} F(\check{z}, \alpha) &:= \partial_{\check{z}}(f^{(\alpha)} - \hat{C}_0(\Lambda, \alpha))\alpha^{-2} \\ &= \partial_{\check{z}}\Big(\hat{\Omega}(\Lambda, \alpha)\cdot\check{\mathrm{R}} + \frac{\alpha^2}{2}\hat{\tau}(\Lambda)\cdot\check{\mathrm{R}}^2 + \cdots + \alpha^{2s-2}\mathcal{P}_s(\mathrm{R}; \Lambda) + \alpha^{2s}\mathcal{Q}(\Lambda, \check{z}, \alpha)\Big) . \end{aligned}$$

Then, quite similarly, for small values of α, by the Implicit-Function Theorem, one finds an equilibrium point $\check{z}_{\mathrm{e}}(\Lambda, \alpha)$ for F which satisfies, instead of [6, Eq. (10.21)], the following estimate (with possibly a bigger value of c_4)

$$|\check{z}_{\mathrm{e}}| \le 2\mathrm{m}|F(0, \alpha)| \le c_4\alpha^{2s} ,$$

with m as in [6, Eq. (10.19)]. Thus, the function $\hat{f}_{\mathrm{G,av}}$ has an equilibrium point $\hat{z}_{\mathrm{e}}(\Lambda, G) := \varrho(\Lambda, G)\check{w}_{\mathrm{e}}(\Lambda, \varrho(\Lambda, G))$ satisfying $|\hat{z}_{\mathrm{e}}(\Lambda, G)| \le C\varrho(\Lambda, G)^{2s+1}$, with a suitable constant C independent of Λ and G.

Next, instead of taking $\varrho < \epsilon_2$, where ϵ_2 is an upper bound for ϱ with the property at the end of [6, Step 1], take $\varrho(\Lambda, G) \le \delta^*$, where δ^* is so small that, for $\varrho(\Lambda, G) \le \delta^*$, the following inequality holds

$$|z_{\mathrm{e}}(\Lambda, G)| \le C\varrho(\Lambda, G)^{2s+1} \le \eta\varrho(\Lambda, G) . \tag{C.2}$$

Step 2. Define a change of variables $(I, \varphi^\star, z^\star) \to (I, \hat{\varphi}, \hat{z})$ defined by [6, Eq. (10.22)] and by the last equation at the end of [6, Step 2], but modify the choice of the domain of $\phi^\star$ as follows

$$I \in \mathcal{A} \times \mathbb{R}_+ , \quad \varphi^\star \in \mathbb{T}^{n+1} , \quad |z^\star| \le (\vartheta + (2s-1)\eta)\varrho = (\theta - \eta)\varrho \le \delta^* \tag{C.3}$$

By the triangular inequality, (C.2) and Equation [6, Eq. (10.22)], $\phi^\star$ is well defined on such domain. Exploiting the definition of $\phi^\star$ and (C.2) one finds that $\phi^\star$ (acts as the identity on $I = (\Lambda, G)$, as a $\varphi^\star$-independent shift on $\varphi^\star$ and moreover) verifies

$$|\phi^\star(I, \varphi^\star, z^\star) - (I, \varphi^\star, z^\star)| \le C\varrho(\Lambda, G)^{2s+1} ,$$

with C independent of $\varphi^\star$ and $z^\star$. Finally, letting $\mathcal{H}^\star := \hat{\mathcal{H}} \circ \phi^\star = h_K + \mu f^\star$, one has that the averaged perturbation becomes[28]

$$(f^\star)^{\mathrm{av}}(I, z^\star) := (f \circ \phi^\star)^{\mathrm{av}} = \hat{f}^{\mathrm{av}} \circ \phi^\star$$
$$= C^\star(I) + \Omega^\star(\Lambda) \cdot \mathrm{R}^\star + \frac{1}{2} \tau^\star(\Lambda) \cdot (\mathrm{R}^\star)^2 + \cdots + \mathcal{P}_s^\star(\mathrm{R}^\star, \Lambda)$$
$$+ \mathcal{Q}^\star(I, z^\star)\ , \tag{C.4}$$

for suitable $\Omega^\star$, $\tau^\star$, $\cdots$, $\mathcal{P}_s^\star$, which are ϱ^2-close to $\hat{\Omega}$, $\hat{\tau}$, $\cdots$, $\mathcal{P}_s$ in (C.1) and $\mathcal{Q}^\star$ defined as in [6, Eqs. (10.25)–(10.26)], with $\check{\mathcal{P}}$ replaced by the function $\mathcal{P}$ in (5.1). In particular, $\Omega^\star$ do not satisfy resonances up to order $2s$, provided δ^* is suitably small.

Step 3. Replace [6, Eqs. (10.27)–(10.30)] as follows. Denote by

$$\mathcal{Q}^\star(I, z^\star) = \sum_{k \in \{0, \cdots, 2s\}, k \neq 1} \mathcal{Q}_k^\star + \mathrm{O}(|z^\star|^{2s+1}) \tag{C.5}$$

the Taylor expansion around $z^\star = 0$ of $\mathcal{Q}$ in (C.4). In the case $2s = 4$, $\mathcal{Q}_0^\star$, $\mathcal{Q}_2^\star$, $\mathcal{Q}_3^\star$, $\mathcal{Q}_4^\star$ correspond to the functions $\mathcal{Q}_0^\star$, $\mathrm{Q}^\star$, $\mathrm{C}^\star$, $\mathrm{F}^\star$ of [6, Eq. (10.27)]. By the definition of $\mathcal{Q}^\star$, it is not difficult to see that $\mathcal{Q}_k^\star$ are $\varrho^{(2s-k+2)}$-close to zero. Since $\mathcal{Q}_2^\star$ is ϱ^{2s}-close to zero, for a possibly smaller δ^*, one can find a symplectic transformation $\phi^*: (I, \varphi^*, z^*) \to (I, \varphi^\star, z^\star)$ which leaves I unvaried, as a φ^*-independent shift on φ^*, is linear on w^* and puts $\Omega^\star \cdot \mathrm{R}^\star + \mathcal{Q}_2^\star$ into the normal form $\Omega^\star \cdot \mathrm{R}^\star$, where Ω^* are ϱ^{2s}-close to $\Omega^\star$ and hence do not satisfy resonances up to order $2s$ for a possibly smaller δ^*. Such transformation ϕ^* is easily seen to be ϱ^{2s+1}-close to the identity and the transformed hamiltonian $\mathcal{H}^* := \mathcal{H}^\star \circ \phi^* = h_K(\Lambda) + \mu f^*(\Lambda.l^*, w^*)$ g^*-independent and has the quadratic part of $(f^*)^{\mathrm{av}} = (f^\star)^{\mathrm{av}} \circ \phi^*$ in diagonal form. Finally, since $\phi^\star$ is ϱ^{2s+1}-close to the identity, with an eventually small δ^* for which $|z^\star - z^*| \le C\varrho^{2s+1} \le \eta\varrho$, one can take as domain of ϕ^* the set

$$I \in \mathcal{A} \times \mathbb{R}_+\ , \quad \varphi^* \in \mathbb{T}^{n+1}\ , \quad |z^*| \le (\vartheta + (2s-2)\eta)\varrho = (\theta - 2\eta)\varrho(\Lambda, G) \le \delta^*\ , \tag{C.6}$$

which implies that $z^\star$ satisfies (C.3). Moreover, ϕ^* puts $f_{\mathrm{av}}^\star$ into the form

$$f_{\mathrm{av}}^* = f_{\mathrm{av}}^\star \circ \phi^* = C^*(I) + \Omega^* \cdot \mathrm{R}^* + \frac{1}{2}\tau^* \cdot (\mathrm{R}^*)^2 + \cdots \mathcal{P}_s^*(\mathrm{R}^*, \Lambda)$$
$$+ \sum_{k \in \{3, \cdots, 2s\}} \mathcal{Q}_k^*(I, z^*) + \mathrm{O}(|z^\star|^{2s+1})$$

where $\mathcal{Q}_k^*$ are monomials of degree k in z^*, which are ϱ^{2s}-close to $\mathcal{Q}_k^\star$ in (C.5) and hence ϱ^{2s+2-k}-close to zero. This implies in particular that Ω^* are $(2s)$ nonresonant and the matrix τ^* is ϱ^2-close to $\hat{\tau}$ in (C.1), hence, nonsingular. Note that, in the case $2s = 4$, $C^\star$, Ω^*, τ^* correspond to the functions $\check{C}_0$, $\check{\Omega}$, $\hat{\tau}$ in the last equation in [6, Step 3]; $\mathcal{Q}_3^*$, $\mathcal{Q}_4^*$ to the functions C^*, F^*.

[28]The operation of composition with $\phi^\star$ commutes with $\lambda^\star$-averaging, since $\phi^\star$ acts $\varphi^\star$-independent shift on $\varphi^\star$. This fact is common to the transformations ϕ^*, $\check{\phi}_{2s-2}$ below and it will not be mentioned anymore.

Step 4. Apply now a Birkhoff transformation $\check{\phi}_{2s-2}$ in $(2s-2)$ steps (which is possible thanks to nonresonance of Ω^*). From the claimed properties of the polynomials $\mathcal{Q}_k^*$ in Step 3 above, one has that $\check{\phi}_{2s-2}$ can be chosen to be ϱ^{2s+1}-close to the identity, and acting as a the identity on I, as a $\check{\varphi}$-independent shift on $\check{\varphi}$. Letting δ^* to be so small that $|\check{z}-z^*| \le C\varrho^{2s+1} \le (2s-2)\eta$, one has that the domain of $\check{\phi}_{2s-2}$ may be chosen to be $I \in \mathscr{A} \times \mathbb{R}_+ \check{\varphi} \in \mathbb{T}^{n+1}$, $|\check{z}| \le \vartheta\varrho(\Lambda, G) \le \delta^*$, so that z^* satisfies (C.6). This implies in particular that $\check{\phi} := \phi^\star \circ \phi^* \circ \check{\phi}_{2s-2}$ is well defined on the domain defined in (5.7) above, with $\vartheta = 1/4$ and arbitrary $\hat{\delta} < \delta \le \delta^*$. Moreover, the $(\Lambda, \check{\lambda}, \check{z})$-projection of $\check{\phi}$, $\phi_s := \Pi_{(\Lambda,\check{\lambda},\check{z})} \circ \check{\phi}$ is easily seen to be symplectic with respect to the 2-form $d\Lambda \wedge d\check{\lambda} + d\check{u} \wedge d\check{w}$ and satisfying the statement of the Theorem.

APPENDIX D. PROPERLY-DEGENERATE AVERAGING THEORY

In this Appendix we shall prove a result in Averaging Theory, which is needed in the proof of Theorem 6.1.

Let us fix some standard notations: $B_r^m(z)$ denotes the *complex* ball of radius r in $\mathbb{C}^m$, centered in z; the ball around the origin $B_r^m(0)$ is simply denoted by B_r^m. If $V \subseteq \mathbb{R}^m$ is an open set, V_ρ denotes the complex set $\bigcup_{x\in V} B_\rho^m(x)$ and $\mathbb{T}_s^m$ denotes the complex neighborhood of $\mathbb{T}^m$ given by $\{x \in \mathbb{C}^m : |\mathrm{Im}\, x_j| < s, 1 \le j \le m\}/(2\pi\mathbb{R}^m)$. Also, if $f(u,\varphi) = \sum_{k\in\mathbb{Z}^n} f_k(u) e^{\mathrm{i}k\cdot\varphi}$ is a real-analytic function on $W_{v,s} = V_v \times \mathbb{T}_s^n$, $\|f\|_{v,s}$ denotes its "sup-Fourier" norm:

$$\|f\|_{v,s} := \sum_{k\in\mathbb{Z}^n} \sup_{V_v} |f_k| e^{\mathrm{i}|k|s},$$

where $|k| := |k|_1 := \sum_{i=1}^n |k_i|$.

PROPOSITION D.1. *Let n_1, $n_2 \in \mathbb{N}$; let V be an open set in $\mathbb{R}^{n_1}$;*

$$W_{\rho_0,\epsilon_0,s_0} := V_{\rho_0} \times \mathbb{T}_{s_0}^{n_1} \times B_{\epsilon_0}^{2n_2};$$

let $H(I,\varphi,p,q;\mu)\colon W_{\rho_0,\epsilon_0,s_0} \to \mathbb{C}$ be a real-analytic Hamiltonian on $W_{\rho_0,\epsilon_0,s_0}$ of the form

$$H(I,\varphi,p,q;\mu) := H_0(I;\mu) + \mu P(I,\varphi,p,q;\mu) \tag{D.1}$$

where the average $P_{\mathrm{av}} := \int_{\mathbb{T}^{n_1}} P(I,\varphi,p,q;\mu)\, \frac{d\varphi}{(2\pi)^{n_1}}$ has an elliptic equilibrium in $p = q = 0$ for all $I \in V$. Assume that the map $I \to \partial^2 H_0(I;\mu)$ is a diffeomorphism of V; that the first-order Birkhoff invariants Ω of P_{av} do not satisfy resonances on V up to the order $2s$. Let $\tau > n-1$.

There exist positive numbers $c_\star$, c_0 such that, for all $0 < a < \frac{1}{4(\tau+1)}$ one can find a number $0 < \epsilon_\star < 1$ such that for all

$$\gamma_0 \ge 1\,, \quad 0 < \bar{\epsilon} < \epsilon_\star \quad \text{and} \quad (c_\star\bar{\epsilon})^{\frac{1}{2a(\tau+1)}} < \mu < \Big(\frac{\bar{\epsilon}}{c_\star\gamma_0}\Big)^{\frac{1}{1-2a(\tau+1)}} \tag{D.2}$$

one can find an open set $V_\star \subseteq V_{\rho_0/32}$ a positive number c and a real-analytic symplectic transformation

$$\phi_\star\colon (V_\star)_{\rho_\star} \times \mathbb{T}^{n_1}_{s_0/24} \times B^{2n_2}_{3\bar{\epsilon}/32} \to (V_\star)_{31\rho_\star} \times \mathbb{T}^{n_1}_{s_0/6} \times B^{2n_2}_{\bar{\epsilon}} \tag{D.3}$$

where $\rho_\star := \frac{\gamma_0}{c_0}\max\{\sqrt{\frac{\mu}{\bar\epsilon}},\ \sqrt{\bar\epsilon}\}\mu^{a/2} \le \frac{\rho_0}{32}$, *which carries* H *into* $H_\star := H\circ\phi_\star$, *where*

$$H_\star(I,\varphi,p,q) = H_0(I) + \mu N_\star(I,r) + \mu P_\star(I,p,q) + c\mu e^{-(\frac{1}{c_\star\mu})^a} f_\star(I,\varphi,p,q)\ , \tag{D.4}$$

where $N_\star$ *is a polynomial of degree* s *in* $r_i = \dfrac{{p_i}^2+{q_i}^2}{2}$ *whose coefficients are* $(\bar\epsilon, \mu/\bar\epsilon)$*-close to those of the Birkhoff normal form associated to* $P_{\rm av}$*;* $P_\star$ *has a zero of order* $(2s+1)$ *in* $(p,q)=0$ *for all* $I\in(V_\star)_{\rho_\star}$ *and* $f_\star$ *is uniformly bounded by* 1.

The transformation $\phi_\star$ *may be chosen so as to satisfy*

$$\begin{aligned}
&\phi_\star\Big((V_\star)_{\rho_\star/2}\times\mathbb{T}^{n_1}_{s_0(1+\frac{1}{\gamma_0})/48}\times B^{2n_2}_{3\bar\epsilon(1+\frac{1}{\gamma_0})/64}\Big) \supseteq (V_\star)_{\rho_\star/4}\times\mathbb{T}^{n_1}_{s_0/48}\times B^{2n_2}_{3\bar\epsilon/64}\\
&\phi_\star\Big((V_\star)_{\rho_\star(1+\eta)/2}\times\mathbb{T}^{n_1}_{s_0(1+\eta+\frac{1}{\gamma_0})/48}\times B^{2n_2}_{3\bar\epsilon(1+\eta+\frac{1}{\gamma_0})/64}\Big)\subseteq\\
&\qquad\qquad (V_\star)_{3\rho_\star/4+\rho_\star\eta/2}\times\mathbb{T}^{n_1}_{s_0(1+\eta+\frac{2}{\gamma_0})/48}\times B^{2n_2}_{3\bar\epsilon(1+\eta+\frac{2}{\gamma_0})/64}
\end{aligned}\tag{D.5}$$

for all $\eta\in(0,1)$ *and, moreover, if* $(I_\star,\varphi_\star,p_\star,q_\star)$ *is short for* $\phi_\star(I,\varphi,p,q)$*, the following bounds*

$$\begin{aligned}
|I_\star - I| &\le \frac{c_\star}{\gamma_0}\min\{\sqrt{\mu\bar\epsilon},\ \frac{\mu}{\sqrt{\bar\epsilon}}\}\mu^{a/2}\\
|\varphi_\star-\varphi| &\le \frac{c_\star}{\gamma_0}\mu^{a(6\tau+5)/2}\\
\max\{|p_\star-p|\ ,\ |q_\star-q|\} &\le \frac{c_\star}{\gamma_0}\max\{\bar\epsilon,\ \frac{\mu}{\sqrt{\bar\epsilon}}\}\ .
\end{aligned}\tag{D.6}$$

The set $V_\star$ *can be chosen to have Lebesgue measure*

$$\mathrm{meas}\, V_\star \ge (1-\mathrm{f}_\star(\bar\epsilon,\mu)\mu^{-a(\tau+1/2)})\,\mathrm{meas}\, V\ , \tag{D.7}$$

with $\mathrm{f}_\star(\bar\epsilon,\mu) := \sqrt{c_\star\gamma_0}\max\{\sqrt{\frac{\mu}{\bar\epsilon}},\ \sqrt{\bar\epsilon}\}$.

If, instead of (D.2)*, one assumes*

$$\gamma_0\ge 1\ ,\quad 0<\bar\epsilon<\epsilon_\star\quad \textit{and}\quad 0<\mu<\frac{1}{c_\star\gamma_0}\bar\epsilon\,(\log\bar\epsilon^{-1})^{-2(\tau+1)} \tag{D.8}$$

(*where* $\epsilon_\star$ *and* $c_\star$ *depend only on* s) *then* $\rho_\star$*,* (D.4)*,* (D.6) *and* (D.7) *are respectively replaced by* $\bar\rho_\star = \frac{\gamma_0}{c_0}\max\{\sqrt{\frac{\mu}{\bar\epsilon}},\ \sqrt{\bar\epsilon}\}$,

$$H_\star(I,\varphi,p,q) = H_0(I)+\mu N_\star(I,r)+\mu P_\star(I,p,q)+c\mu\epsilon^{2s+1}Q_\star(I,\varphi,p,q)$$

$$\left\{\begin{aligned}
&|I-I_*|\le \tfrac{c_*}{\gamma_0}\min\{\sqrt{\mu\bar\epsilon},\ \tfrac{\mu}{\sqrt{\bar\epsilon}}\}(\log\bar\epsilon^{-1})^{-1}\\
&|\varphi-\varphi_*|\le \tfrac{c_*}{\gamma_0}\min\{\bar\epsilon,\ \tfrac{\mu}{\bar\epsilon}\}(\log\bar\epsilon^{-1})^{-1}\\
&|p-p_*|\ ,\ |q-q_*|\le \tfrac{c_*}{\gamma_0}\min\{\bar\epsilon,\ \tfrac{\mu}{\bar\epsilon}\}(\log\bar\epsilon^{-1})^{-1}
\end{aligned}\right. \tag{D.9}$$

$$\mathrm{meas}\, V_\star \ge \big(1-\mathrm{f}_\star(\bar\epsilon,\mu)(\log\bar\epsilon^{-1})^{\tau+1}\big)\,\mathrm{meas}\, V$$

where $N_\star$ *and* $P_\star$, $\mathrm{f}_\star$ *are as above and* $|Q_\star|\le 1$.

The proof is based upon a technical result proven in [15] or [4].

LEMMA D.2 (Averaging Theory). *Let $\bar K, \bar s, s > 0$ be such that $\bar K s \ge 6$; let $\alpha > 0$ and $\ell \in \mathbb{N}$. Let $H(u,\varphi) = h(I) + f(u,\varphi)$, with $f(u,\varphi) = \sum_k f_k(u) e^{ik\cdot\varphi}$, be real-analytic on $W_{v,\bar s+s} := A_r \times B_{r_p} \times B'_{r_q} \times \mathbb{T}^\ell_{\bar s+s}$, where $A\times B\times B' \subseteq \mathbb{R}^\ell \times \mathbb{R}^m \times \mathbb{R}^m$ and $v = (r, r_p, r_q)$. Finally, let Λ be a (possibly trivial) sub-lattice of $\mathbb{Z}^\ell$ and let ω denote the gradient $\partial_I h$. Assume that*

$$|\omega\cdot k| \ge \alpha \quad \forall\ I \in A_r\ ,\ \forall\ k \notin \Lambda\ ,\ |k| \le \bar K \tag{D.10}$$

$$E := \|f\|_{v,\bar s+s} < \frac{\alpha d}{2^7 c_m \bar K s}\ ,\ \text{where}\ \ d = \min\{r s,\ r_p r_q\}\ ,\ \ c_m := \frac{e(1+em)}{2}\ . \tag{D.11}$$

Then there exists a real-analytic, symplectic transformation

$$\Psi\colon (I',\varphi',p',q') \in W_{v/2,\bar s+s/6} \to (I,\varphi,p,q) \in W_{v,\bar s+s}$$

such that

$$H_* := H\circ\Psi = h + g + f_*\ ,$$

with g in normal form and f_* small:

$$g = \sum_{k\in\Lambda} g_k(I',p',q')\, e^{ik\cdot\varphi'}\ ,\quad \|g - \Pi_\Lambda T_{\bar K} f\|_{v/2,\bar s+s/6} \le \frac{12}{11}\,\frac{2^7 c_m E^2}{\alpha d} \le \frac{E}{4}\ , \tag{D.12}$$
$$\|f_*\|_{v/2,\bar s+s/6} \le e^{-\bar K s/6} E\ .$$

Moreover, denoting by $z = z(I',\varphi',p',q')$, the projection of $\Psi(I',\varphi',p',q')$ onto the z-variables ($z = I$, φ, p or q) one has

$$\max\{\alpha s|I-I'|,\ \alpha r|\varphi-\varphi'|,\ \alpha r_q|p-p'|,\ \alpha r_p|q-q'|\} \le 9E\,. \tag{D.13}$$

Proof of Proposition D.1. Assume (D.2). Pick two numbers C_0 and $C \ge 1$; let the numbers $c_\star$, $\epsilon_\star$ of the statement satisfy $c_\star \ge (2CC_0)^2$ and $\epsilon_\star \le \left(\frac{1}{(2CC_0)^2}\right)^{\frac{1}{1-4a(\tau+1)}}$. The proof will be based on the following inequalities (implied by (D.2)) for $\bar\epsilon$ and μ and the definition of $\bar\gamma = \bar\gamma(\bar\epsilon,\mu)$:

$$\begin{cases} \bar\epsilon < \left(\frac{1}{(2CC_0)^2}\right)^{\frac{1}{1-4a(\tau+1)}} \\ \left((2CC_0)^2\bar\epsilon\right)^{\frac{1}{2a(\tau+1)}} < \mu < \left(\frac{\bar\epsilon}{(2CC_0)^2}\right)^{\frac{1}{1-2a(\tau+1)}} \\ \bar\gamma := 2C\gamma_0 \max\left\{\sqrt{\frac{\mu}{\bar\epsilon}}\,\bar K^{\tau+1/2},\ \sqrt{\bar\epsilon}\,\bar K^{\tau+1/2}\right\} \quad \text{with} \quad \bar K := \frac{1}{\mu^a}\ . \end{cases} \tag{D.14}$$

The numbers C_0 and C will be chosen later, independently of γ_0, a and, obviously, on $\bar\epsilon$ and μ.

Step 1 (Averaging over the "fast angles" φ). Let (I_0,φ_0,p_0,q_0) denote the variables in (D.1). We can assume that $P_{\rm av}(p_0,q_0;I_0)$ is in Birkhoff normal form of order $2s$. The first step consists in removing, in H, the dependence on φ up an exponential order (namely, up to $O(e^{-1/\mu^a})$). Let ρ_0, ϵ_0, s_0 denote the analyticity radii of H in I_0, (p_0,q_0), φ_0, respectively and take $\bar\epsilon \le \epsilon_0$. We apply Lemma D.2, with equal scales, i.e., taking $\alpha_1 = \alpha_2 := \alpha$ (see below). Next, we take $\ell := \ell_1 + \ell_2 = n_1$, $m = n_2$ $h = H_0$, $B = B' = \{0\}$, $r_p = r_q = \epsilon_0$, $s = s_0$,

$\bar{s} = 0$, $\Lambda = \{0\} \in \mathbb{Z}^{n_1}$, $A = \bar{D}$, $\mathfrak{r} = \bar{\rho}$, where $\bar{D}$, $\bar{\rho}$ are defined as follows. Let $\tau > n_1$, $\bar{M} := \max_{i,j} \sup_{V_{\rho_0}} |\partial^2_{ij} H_0(I_0)|$, $c_0 := \frac{32\bar{M}}{C}$, $\hat{\rho} := \max\{\sqrt{\frac{\mu}{\bar{\epsilon}}},\ \sqrt{\bar{\epsilon}}\}\mu^{a/2}$. Take

$$\bar{D} := \bar{\omega}_0^{-1}\left(\mathscr{D}^{n_1}_{\bar{\gamma},\tau}\right) \cap V \quad \text{and} \quad \bar{\rho} := 32\frac{\gamma_0}{c_0}\hat{\rho} = \frac{\bar{\gamma}}{2\bar{M}\bar{K}^{\tau+1}} \leq \rho_0\ , \tag{D.15}$$

where $\mathscr{D}_{\bar{\gamma},\tau} \subseteq \mathbb{R}^{n_1}$ is the set of $(\bar{\gamma},\tau)$-diophantine numbers in $\mathbb{R}^{n_1}$, i.e.,

$$\mathscr{D}_{\bar{\gamma},\tau} := \left\{\omega \in \mathbb{R}^{n_1} :\ |\omega \cdot k| \geq \frac{\bar{\gamma}}{|k|^\tau} \quad \text{for all} \quad k \in \mathbb{Z}^{n_1}\ ,\ k \neq 0\right\}.$$

Let now $\rho_\star$, $V_\star$ be defined as

$$\rho_\star = \frac{\bar{\rho}}{32} = \frac{\gamma_0}{c_0}\hat{\rho}\ , \quad V_\star := \bar{D}_{\rho_\star}\ . \tag{D.16}$$

The following measure estimate is standard, since $\bar{\omega}_0 = \partial H_0$ is a diffeomorphism of V and $\tau > n-1$.

$$\text{meas}\left(V \smallsetminus V_\star\right) \leq \text{meas}\left(V \smallsetminus \bar{D}\right) \leq \bar{C}_0\bar{\gamma}\,\text{meas}\,(V)$$

where $\bar{C}_0$ is a suitable number depending only on V. Take in (D.14) $C_0 \geq \bar{C}_0$ and $C > 2^{-1}\sqrt{s_0\bar{M}2^9 c_{n_2}\|P\|_{(\rho_0,\epsilon_0,\epsilon_0),s_0}}$.

By a standard argument, for $I_0 \in \bar{D}_{\bar{\rho}}$, the unperturbed frequency map $\bar{\omega}_0 = \partial H_0$ verifies (D.10), with $\alpha_1 = \alpha_2 = \alpha := \frac{\bar{\gamma}}{2\bar{K}^\tau}$, r and A as above. The smallness condition (D.11) is easily checked by the choices (D.14): since $\bar{\epsilon}\bar{K} = \bar{\epsilon}\mu^{-a} < \epsilon^{1-1/(2a(\tau+1))} < 1$ and $C > 2^{-1}\sqrt{s_0\bar{M}2^9 c_{n_2}\|P\|_{(\rho_0,\epsilon_0,\epsilon_0),s_0}}$,

$$E = \mu\|P\|_{(\rho_0,\epsilon_0,\epsilon_0),s_0} < \frac{4C^2}{s_0\bar{M}2^9 c_{n_2}}\frac{\mu}{\bar{\epsilon}\bar{K}} \leq \frac{\bar{\gamma}^2}{s_0\bar{M}2^9 c_{n_2}\bar{K}^{2\tau+2}} \leq \frac{\alpha\bar{\rho}}{2^7 c_{n_2}\bar{K}s_0}\ .$$

Inequality $\bar{K}s_0 \geq 6$ is also trivially satisfied. Thus, by Proposition D.2, we find a real-analytic symplectomorphism

$$\bar{\phi}\colon (\bar{I},\bar{\varphi},\bar{p},\bar{q}) \in \bar{D}_{\rho_0/2} \times \mathbb{T}^{n_1}_{s_0/6} \times B^{n_2}_{\epsilon_0/2} \to (I_0,\varphi_0,p_0,q_0) \in \bar{D}_{\rho_0} \times \mathbb{T}^{n_1}_{s_0} \times B^{n_2}_{\epsilon_0}$$

and H is transformed into

$$\bar{H}(\bar{I},\bar{\varphi},\bar{p},\bar{q}) = H \circ \bar{\phi}(\bar{I},\bar{\varphi},\bar{p},\bar{q}) = H_0(\bar{I}) + \mu\bar{N}(\bar{I},\bar{p},\bar{q}) + \mu e^{-\bar{K}s/6}\bar{P}(\bar{I},\bar{\varphi},\bar{p},\bar{q})\ .$$

By (D.12), $\|\bar{P}\|_{\bar{v},\bar{s}} \leq \bar{C}$ and

$$\sup_{\bar{D}_{\bar{\rho}/2}} |\bar{N} - P_{\text{av}}| \leq \bar{C}\frac{\mu\bar{K}^{2\tau+1}}{\bar{\gamma}^2}\ . \tag{D.17}$$

Since $\bar{\epsilon} < \epsilon_0$, in particular, $\bar{\phi}$ is defined on the *smaller* set $W_{(\bar{\rho}/2,\bar{\epsilon}/2),\bar{s}}$, and the following inclusion holds

$$\bar{\phi}\colon \bar{D}_{\bar{\rho}/2} \times \mathbb{T}^{n_1}_{s_0/6} \times B^{n_2}_{\bar{\epsilon}/2} \to \bar{D}_{\bar{\rho}} \times \mathbb{T}^{n_1}_{s_0} \times B^{n_2}_{\bar{\epsilon}} \tag{D.18}$$

as it follows from the following inequalities

$$
\begin{aligned}
|I_0-\bar I| &\le \bar C\frac{\mu\bar K^\tau}{\bar\gamma}=\frac{\mu\bar C}{2C\gamma_0\bar K^{1/2}}\min\{\sqrt{\frac{\bar\epsilon}{\mu}},\frac{1}{\sqrt{\bar\epsilon}}\}\le\frac{2C\gamma_0\max\{\sqrt{\frac{\mu}{\bar\epsilon}},\sqrt{\bar\epsilon}\}}{128}=\frac{\bar\rho}{128}\\
|p_0-\bar p|,\ |q-\bar q| &\le \bar C\frac{\mu\bar K^\tau}{\bar\gamma}=\frac{\mu\bar C}{2C\gamma_0\bar K^{1/2}}\min\{\sqrt{\frac{\bar\epsilon}{\mu}},\frac{1}{\sqrt{\bar\epsilon}}\}\le\frac{3}{256\gamma_0}\bar\epsilon<\frac{\bar\epsilon}{2}\\
|\varphi_0-\bar\varphi| &\le \bar C\frac{\mu\bar K^{2\tau+1}}{\bar\gamma^2}=\frac{\bar C}{4C^2\gamma_0^2}\min\{\bar\epsilon,\frac{\mu}{\bar\epsilon}\}\le\frac{s_0}{192\gamma_0}\ .
\end{aligned}
\tag{D.19}
$$

Note that the former bounds in each line follow from (D.13); the latter ones follows from the definition of $\bar\rho$ in (D.15), from (D.14), Cauchy estimates and $\gamma_0\ge 1$.

Step 2 (Determination of the elliptic equilibrium for the "secular system"). In view of (D.17), $\bar N-P_{\rm av}$ is of order $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$. Using the Implicit-Function Theorem and standard Cauchy estimates for small values of this parameter, for any fixed $\bar I\in\bar D_{\bar\rho/2}$, $\bar N$ also has an equilibrium point $(p_{\rm e}(I),q_{\rm e}(I))$ which satisfies, by (D.14) and taking $C\ge\sqrt{64\tilde C/3}$ and using $\gamma_0\ge 1$

$$
|(p_{\rm e}(I),q_{\rm e}(I))|\le\tilde C\frac{\mu\bar K^{2\tau+1}}{\bar\gamma^2}=\frac{\tilde C}{4C^2\gamma_0^2}\min\{\bar\epsilon,\ \frac{\mu}{\bar\epsilon}\}\le\frac{3}{256\gamma_0^2}\min\{\bar\epsilon,\ \frac{\mu}{\bar\epsilon}\}<\frac{\bar\epsilon}{8}
\tag{D.20}
$$

Consider now a neighborhood of radius $3\bar\epsilon/8$ around $(p_{\rm e}(I),q_{\rm e}(I))$. We let

$$
\tilde\phi:\ (\tilde I,\tilde\varphi,\tilde p,\tilde q)\in\bar D_{\bar\rho/4}\times\mathbb{T}^{n_1}_{s_0/12}\times B^{n_2}_{3\bar\epsilon/8}\to(\bar I,\bar\varphi,\bar p,\bar q)\in\bar D_{\bar\rho/2}\times\mathbb{T}^{n_1}_{s_0/6}\times B^{n_2}_{\bar\epsilon/2}
\tag{D.21}
$$

the transformation which acts as

$$
\bar I=\tilde I\ ,\quad \bar p=p_{\rm e}(\tilde I)+\tilde p\ ,\quad \bar q=q_{\rm e}(\tilde I)+\tilde q\ ,\ \bar\varphi=\tilde\varphi-\partial_{\tilde I}\Big(\tilde p+p_{\rm e}(\tilde I)\Big)\cdot\Big(\bar q-q_{\rm e}(\tilde I)\Big)\ .
$$

Such transformation is easily seen to be symplectic, having

$$
\tilde s(\tilde I,\tilde p,\bar\varphi,\bar q)=\tilde I\cdot\bar\varphi+\Big(\tilde p+p_{\rm e}(\tilde I)\Big)\cdot\Big(\bar q-q_{\rm e}(\tilde I)\Big)
$$

as generating function. Note that $\tilde\phi$ is well defined, since, in view of (D.14), (D.20), Cauchy estimates, one has

$$
\begin{aligned}
&|\bar p-\tilde p|=|p_{\rm e}|\le\frac{3}{256\gamma_0^2}\min\{\bar\epsilon,\ \frac{\mu}{\bar\epsilon}\}\ ,\quad |\bar q-\tilde q|=|q_{\rm e}|\le\frac{3}{256\gamma_0^2}\min\{\bar\epsilon,\ \frac{\mu}{\bar\epsilon}\}\\
&|\bar\varphi-\tilde\varphi|\le\tilde C\max\Big\{\frac{\bar\epsilon^2\bar K^{\tau+1}}{\bar\gamma},\frac{\mu\bar\epsilon\bar K^{3\tau+2}}{\bar\gamma^3}\Big\}\le\frac{\tilde C}{2C\gamma_0}\mu^{a(6\tau+5)/2}\le\frac{s_0}{192\gamma_0}<\frac{s_0}{12}
\end{aligned}
\tag{D.22}
$$

where we have used $\mu<1$ and $C\ge\frac{192}{s_0}\tilde C$.

Finally, $\tilde\phi$ puts $\bar H$ into the form

$$
\tilde H:=\bar H\circ\tilde\phi=H_0(\tilde I)+\mu\tilde N(\tilde I,\tilde p,\tilde q)+\mu e^{-\bar K s/6}\tilde P(\tilde I,\tilde\varphi,\tilde p,\tilde q),
$$

with $\tilde N:=\bar N\circ\tilde\phi$, $\tilde P:=\bar P\circ\tilde\phi$. Observe that $\|\tilde P\|_{\tilde v,\tilde s}\le C$ and $\tilde N$ has an elliptic equilibrium point into the origin and, being $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$-close to $P_{\rm av}$ (see (D.17)), its quadratic part is $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$-close to be diagonal.

Step 3 (Symplectic diagonalization of the secular system). We now proceed to diagonalize the quadratic part (in $(\tilde p,\tilde q)$) of $\tilde N$. By (D.17), since $P_{\rm av}$ is in Birkhoff normal form, one has that $\tilde N$ is $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$-close to be diagonal. Therefore, one finds a symplectic transformation

$$\hat\phi:\ (\hat I,\hat\varphi,\hat p,\hat q)\in \bar D_{\bar\rho/8}\times\mathbb{T}^{n_1}_{s_0/24}\times B^{n_2}_{3\bar\epsilon/16}\to(\tilde I,\tilde\varphi,\tilde p,\tilde q)\in\bar D_{\bar\rho/4}\times\mathbb{T}^{n_1}_{s_0/12}\times B^{n_2}_{3\bar\epsilon/8} \tag{D.23}$$

which is estimated by

$$\begin{aligned}|\tilde p-\hat p|,\ |\tilde q-\hat q|&\le\hat C\frac{\mu\bar\epsilon\bar K^{2\tau+1}}{\bar\gamma^2}=\frac{\hat C}{4C^2\gamma_0^2}\min\{\bar\epsilon^2,\mu\}\le\frac{3}{256\gamma_0^2}\min\{\bar\epsilon^2,\mu\}<\frac{3\bar\epsilon}{16}\\ |\tilde\varphi-\hat\varphi|&\le\hat C\frac{\mu\bar\epsilon^2\bar K^{3\tau+2}}{\bar\gamma^3}\le\frac{\hat C}{2C\gamma_0}\mu^{a(6\tau+5)/2}\le\frac{s_0}{192\gamma_0}<\frac{s_0}{24}\end{aligned} \tag{D.24}$$

having used again Cauchy estimates, $\gamma_0\ge1$, $\bar\epsilon^2<\bar\epsilon<1$ and the second inequality in (D.22). By construction, the quadratic part of $\hat N$, where $\hat N$ is defined by the equality

$$\hat H:=\tilde H\circ\hat\phi=H_0(\hat I)+\mu\hat N(\hat I,\hat p,\hat q)+\mu e^{-\bar K s/6}\hat P(\hat I,\hat\varphi,\hat p,\hat q),\quad(\hat P:=\tilde P\circ\hat\phi)\ ,$$

is in diagonal form. Moreover, choosing a possibly bigger $c_\star$, one has that the first-order Birkhoff invariants $\hat\Omega$ of $\hat N$, being $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$-close to the corresponding ones of $P_{\rm av}$, are nonresonant of order $(2s)$. Note that, since $\hat N$ is $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$-close to $\bar N$, by (D.17), is also $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$-close to be in $(2s)$-Birkhoff normal form.

Step 4 (Birkhoff normal form of the secular part). We finally use Birkhoff theory to put $\hat N$ in Birkhoff normal form of order $2s$. This is possible since, as above remarked, the first-order Birkhoff invariants $\hat\Omega$ of $\hat N$ are nonresonant up to the order $(2s)$. Recalling that $\hat N$ is $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$-close to be in $(2s)$-Birkhoff normal form, we then find a real-analytic and symplectic transformation

$$\check\phi:\ (\check I,\check\varphi,\check p,\check q)\in\bar D_{\bar\rho/16}\times\mathbb{T}^{n_1}_{s_0/48}\times B^{n_2}_{3\bar\epsilon/32}\to(\hat I,\hat\varphi,\hat p,\hat q)\bar D_{\bar\rho/8}\times\mathbb{T}^{n_1}_{s_0/24}\times B^{n_2}_{3\bar\epsilon/16} \tag{D.25}$$

which acts as the identity on the $\check I$-variables and, on the other variables, is estimated by

$$\begin{aligned}|\hat p-\check p|,\ |\hat q-\check q|&\le\check C\frac{\mu\bar\epsilon^2\bar K^{2\tau+1}}{\bar\gamma^2}=\frac{\check C}{4C^2\gamma_0^2}\min\{\bar\epsilon^3,\mu\bar\epsilon\}\le\frac{3}{256\gamma_0^2}\bar\epsilon<\frac{3}{32}\bar\epsilon\\ |\hat\varphi-\check\varphi|&\le\check C\frac{\mu\bar\epsilon^3\bar K^{3\tau+2}}{\bar\gamma^3}\le\frac{\check C}{2C\gamma_0}\mu^{a(6\tau+5)/2}\le\frac{s_0}{192\gamma_0}<\frac{s_0}{48}\end{aligned} \tag{D.26}$$

by Cauchy estimates, $\mu<1$ and again by the second inequality in (D.22).

Moreover, $\check\phi$ puts $\hat H$ into the form

$$\check H:=\hat H\circ\check\phi:=H_0(\check I)+\mu\check N(\check I,\check r)+\mu\check P+\mu e^{-\bar K s/6}\check f \tag{D.27}$$

where $\check N$ is a polynomial of degree s in $\check r_i=\dfrac{\check p_i^2+\check q_i^2}{2}$ and $\check P$ has a zero of order $(2s+1)$ in $(\check p,\check q)=0$.

Step 5 (Conclusion). Take the transformation $\phi_\star$ in (D.3) as $\phi_\star := \check\phi \circ \tilde\phi \circ \hat\phi \circ \check\phi$ where $\bar\phi$, $\tilde\phi$, $\hat\phi$, $\check\phi$ are as above, $H_\star = \check H$, $N_\star = \check N$, $P_\star = \check P$ as in (D.27) and $f_\star$ by default. The transformation $\phi_\star$ is easily seen to be well defined by the definitions of $V_\star$ and of $\rho_\star$ in (D.16) and by the inclusions (D.18), (D.21), (D.23) and (D.25). Moreover, the bounds (D.19), (D.22), (D.24) and (D.26) and usual telescopic arguments easily imply (D.5) and (D.6). This completes the proof of the first part of the proposition.

The proof that (D.8) implies (D.9) in place of (D.4), (D.6) and (D.7) proceeds along the same lines above, replacing the "power low" choice of $\bar K$ and $\bar\gamma$ in (D.14) with the following "logarithmic" ones

$$\tilde K := \frac{6(2s+1)}{s_0}(\log(\epsilon^{-1}))^{-1}\ , \quad \tilde\gamma := 2C\gamma_0 \max\{\sqrt{\frac{\mu}{\bar\epsilon}},\ \sqrt{\bar\epsilon}\}\,\tilde K^{\tau+1}\ . \qquad \square$$

REFERENCES

[1] K. Abdullah and A. Albouy, *On a strange resonance noticed by M. Herman*, Regul. Chaotic Dyn., **6** (2001), 421–432.

[2] V. I. Arnol'd, *Small denominators and problems of stability of motion in classical and celestial mechanics*, Uspehi Mat. Nauk, **18** (1963), 91–192.

[3] L. Biasco, L. Chierchia and E. Valdinoci, *Elliptic two-dimensional invariant tori for the planetary three-body problem,* Arch. Rational Mech. Anal., **170** (2003), 91–135. See also Corrigendum, Arch. Ration. Mech. Anal., **180** (2006), 507–509.

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

[5] L. Chierchia and G. Pinzari, *Deprit's reduction of the nodes revisited*, Celestial Mech. Dynam. Astronom., **109** (2011), 285–301.

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

[7] L. Chierchia and F. Pusateri, *Analytic Lagrangian tori for the planetary many-body problem*, Ergodic Theory Dynam. Systems, **29** (2009), 849–873.

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

[9] J. Féjoz, *Quasiperiodic motions in the planar three-body problem*, J. Differential Equations, **183** (2002), 303–341.

[10] J. Féjoz, *Démonstration du 'théorème d'Arnold' sur la stabilité du système planétaire (d'après Herman)*, Ergodic Theory Dynam. Systems, **24** (2004), 1521–1582.

[11] M. R. Herman, *Torsion du problème planètaire*, ed. J. Fejóz, 'Archives Michel Herman', 2009. Available from: `http://www.college-de-france.fr/default/EN/all/equ_dif/archives_michel_herman.htm`.

[12] H. Hofer and E. Zehnder, "Symplectic Invariants and Hamiltonian Dynamics," Birkhäuser Verlag, Basel, 1994.

[13] N. N. Nehorošev, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems*, Uspehi Mat. Nauk, **32** (1977), 5–66, 287.

[14] L. Niederman, *Stability over exponentially long times in the planetary problem*, Nonlinearity, **9** (1996), 1703–1751.

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

[16] P. Robutel, *Stability of the planetary three-body problem. II. KAM theory and existence of quasiperiodic motions*, Celestial Mech. Dynam. Astronom., **62** (1995), 219–261. See also Erratum, *Celestial Mech. Dynam. Astronom.*, **84** (2002), 317.

LUIGI CHIERCHIA <luigi@mat.uniroma3.it>: Dipartimento di Matematica, Università "Roma Tre", Largo S. L. Murialdo 1, I-00146 Roma, Italy

GABRIELLA PINZARI <pinzari@mat.uniroma3.it>: Dipartimento di Matematica, Università "Roma Tre", Largo S. L. Murialdo 1, I-00146 Roma, Italy