ISSN 1560-3547, Regular and Chaotic Dynamics, 2013, Vol. 18, No. 6, pp. 860–906. © Pleiades Publishing, Ltd., 2013.

Aspects of the Planetary Birkhoff Normal Form

Gabriella Pinzari*

Dipartimento di Matematica ed Applicazioni "R. Caccioppoli",
Università di Napoli "Federico II",
Monte Sant'Angelo — Via Cinthia I-80126 Napoli, Italy

Received July 16, 2013; accepted December 4, 2013

Abstract—The discovery of the Birkhoff normal form around circular, co-planar motions for the planetary system opened new insights and hopes for the comprehension of the dynamics of this problem. Remarkably, it allowed to give a *direct* proof (after the proof in [18]) of the celebrated Arnold's Theorem [5] on the stability of planetary motions. In this paper, after reviewing the story of the proof of this theorem, we focus on technical aspects of this normal form. We develop an asymptotic formula for it that may turn to be useful in applications. Then we provide two simple applications to the three-body problem: we prove that the "density" of the *Kolmogorov set* of the spatial three-body problem does not depend on eccentricities and the mutual inclination but depends only on the planets' masses and the separation among semi-axes (going in the direction of an assertion by V. I. Arnold [5]) and, using Nehorošev Theory [33], we prove, in the planar case, stability of *all* planetary actions over exponentially-long times, provided mean-motion resonances are excluded. We also briefly discuss difficulties for full generalization of the results in the paper.

MSC2010 numbers: 34D10, 34C20, 70E55, 70F10, 70F15, 70F07, 37J10, 37J15, 37J25, 37J35, 37J40, 70K45

DOI: 10.1134/S1560354713060178

Keywords: averaging theory, Birkhoff normal form, Nehorošev theory, planetary many-body problem, Arnold's theorem on the stability of planetary motions, properly-degenerate KAM theory, steepness

This paper is dedicated to Professor Alain Chenciner on his 70th birthday

Contents

*E-mail: gabriella.pinzari@unina.it

1. INTRODUCTION AND RESULTS

1.1. The planetary many-body problem consists in determining the dynamics of $(1+n)$ masses undergoing Newtonian attraction. The term "planetary" is reserved to the case when one mass, the "sun", or "star", denoted with $\bar m_0$, is taken to be much greater than the others, $\mu\bar m_1$, $\cdots$, $\mu\bar m_n$, which are called "planets". Here $\mu \ll 1$ is a small number. After the "heliocentric[1] reduction" of invariance by translations, this dynamical system is governed by the $3n$ degrees of freedom Hamiltonian

$$\mathrm{H}_{\mathrm{plt}} = \sum_{i=1}^{n}\left(\frac{|y^{(i)}|^2}{2m_i} - \frac{m_iM_i}{|x^{(i)}|}\right) + \mu\sum_{1\leqslant i<j\leqslant n}\left(\frac{y^{(i)}\cdot y^{(j)}}{\bar m_0} - \frac{\bar m_i\bar m_j}{|x^{(i)}-x^{(j)}|}\right) \tag{1.1}$$

on the "collisionless" phase space

$$(y,x) = (y^{(1)},\cdots,y^{(n)},x^{(1)},\cdots,x^{(n)}) \in (\mathbb{R}^3)^{2n}: \quad x^{(i)}\neq 0, \quad x^{(i)}\neq x^{(j)}$$

endowed with the standard 2-form

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

where $y^{(i)} = (y_1^{(i)}, y_2^{(i)}, y_3^{(i)})$, $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, x_3^{(i)})$. Here, m_i, M_i are suitable auxiliary masses related to $\bar m_i$ and μ via

$$M_i = \bar m_0 + \mu\bar m_i, \qquad m_i = \frac{\bar m_0\bar m_i}{\bar m_0+\mu\bar m_i}.$$

A procedure commonly followed in the past [5, 18, 22, 33] to regard the system as a "close to integrable", was to use a symplectic set of variables, usually called "Poincaré variables". These variables, that we denote

$$(\Lambda_i, \lambda_i, \eta_i, \xi_i, \mathrm{p}_i, \mathrm{q}_i) \qquad 1\leqslant i\leqslant n,$$

are "six per planet". They were introduced by H. Poincaré by modifying another set of "action–angle" variables $(\Lambda_i, \Gamma_i, \Theta_i, \ell_i, \mathrm{g}_i, \theta_i) \in \mathbb{R}^3\times\mathbb{T}^3$ (where $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$, having the Λ_i's in common, called "Delaunay variables". Delaunay variables are "natural", "action–angle" variables related

[1] See, *e.g.*, [42].

to the "Cartesian variables" $(y^{(i)}, x^{(i)})$ in (1.1) via the integration of each of the "two-body" Hamiltonians

$$\frac{|y^{(i)}|^2}{2m_i} - \frac{m_i M_i}{|x^{(i)}|}$$

in a domain of phase space where such Hamiltonians are simultaneously negative. In this case, the "unperturbed motions" of $\mathcal{H}_{\rm plt}$ (i.e., neglecting the term proportional to μ in (1.1)) correspond to Keplerian ellipses $\mathcal{E}_i$ having their foci in common. Poincaré variables are in part "action–angle" (i.e., $(\Lambda_i, \lambda_i) \in \mathbb{R} \times \mathbb{T}$), in part "rectangular" (i.e., $(\eta_i, \xi_i, {\rm p}_i, {\rm q}_i) \in \mathbb{R}^4$). The couples (Λ_i, λ_i), often referred to as "fast variables", are related to semi-major axes and mean anomalies of $\mathcal{E}_i$'s, while the "secular variables" ${\rm z}_i = (\eta_i, \xi_i, {\rm p}_i, {\rm q}_i)$ are related to their eccentricities, perihelia (η_i, ξ_i); and to the directions of their planes $({\rm p}_i, {\rm q}_i)$. See, *e.g.*, [5, 15, 18] for precise definitions. In Delaunay–Poincaré variables, any of the two-body Hamiltonian above takes the "Kepler form"

$$h^{(i)}_{\rm Kep}(\Lambda_i) = -\frac{m_i^3 M_i^2}{2\Lambda_i^2}. \tag{1.2}$$

It is "properly degenerate": two degrees of freedom disappear, as it is well known. This proper degeneracy naturally reflects on the system (1.1), which in fact takes the form

$$\mathcal{H}_{\rm P}(\Lambda, \lambda, {\rm z}) = h_{\rm Kep}(\Lambda) + \mu f_{\rm P}(\Lambda, \lambda, {\rm z}) \tag{1.3}$$

where $h_{\rm Kep}(\Lambda)$ is the n degrees of freedom "Keplerian" part $-\sum_{i=1}^n \frac{M_i^2 m_i^3}{2\Lambda_i^2}$, while $f_{\rm P}(\Lambda, \lambda, {\rm z})$ is the $3n$ degrees of freedom "perturbation"

$$\sum_{1 \leqslant i < j \leqslant n} \left(\frac{y^{(i)} \cdot y^{(j)}}{\bar m_0} - \frac{\bar m_i \bar m_j}{|x^{(i)} - x^{(j)}|} \right) \tag{1.4}$$

in (1.1), expressed in Poincaré variables. Here, we have denoted as $(\Lambda, \lambda, {\rm z})$ the $3n$-dimensional collection of

$$\Lambda = (\Lambda_1, \cdots, \Lambda_n), \quad \lambda = (\lambda_1, \cdots, \lambda_n), \quad {\rm z} = ({\rm z}_1, \cdots, {\rm z}_n). \tag{1.5}$$

The phase space of the variables $(\Lambda, \lambda, {\rm z})$ is commonly taken to be $\mathcal{P} := \mathcal{P}_\epsilon = \mathcal{A} \times \mathbb{T}^n \times B$ where, typically, $\mathcal{A}$ is a set of "well spaced" semi-major axes $a_i = a_i(\Lambda_i)$

$$\mathcal{A}: \quad \underline{a}_i \leqslant a_i(\Lambda_i) \leqslant \overline{a}_i \qquad 1 \leqslant i \leqslant n \tag{1.6}$$

for some $\underline{a}_i < \overline{a}_i < \underline{a}_{i+1}$, while $B := B_\epsilon$ is a suitable ϵ-neighborhood (for example, a ball of radius ϵ) of ${\rm z} = 0 \in \mathbb{R}^{4n}$. This corresponds to consider small eccentricities and inclinations of the ellipses $\mathcal{E}_i$'s. Clearly, the radius of B and the parameters $\underline{a}_i$, $\overline{a}_i$ in the definition of $\mathcal{A}$ have to be chosen so as to exclude collisions among the $\mathcal{E}_i$'s.

The averaged ("secular") perturbing function

$$(f_{\rm P})_{\rm av} := \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f_{\rm P} d\lambda$$

of the system (1.3) turns out to have an elliptic equilibrium point[2)] in ${\rm z} = 0$ for any Λ [5]. This equilibrium is a consequence of the symmetries (which, expressed in Poincaré coordinates are often referred to as "D'Alambert rules") of $(f_{\rm P})_{\rm av}$ due to invariance of the system (1.1) by reflections with respect to the coordinate planes and rotations around the coordinate axes. Therefore, the "secular equilibrium $\{{\rm z} = 0\}$", i.e., the $(2n)$-dimensional manifold $\mathcal{A} \times \mathbb{T}^n \times \{0\} \subset \mathcal{A} \times \mathbb{T}^n \times \mathbb{R}^{4n}$, is invariant for the motions of the "secular system"

$$(\mathcal{H}_{\rm P})_{\rm av} := h_{\rm Kep} + \mu (f_{\rm P})_{\rm av}.$$

[2)] Namely, $(f_{\rm P})_{\rm av}$ has an expansion $(f_{\rm P})_{\rm av} = C_0(\Lambda) + \mathcal{Q}(\Lambda) \cdot {\rm z}^2 + {\rm O}({\rm z}^4; \Lambda)$, where $\mathcal{Q}(\Lambda) \cdot {\rm z}^2$ is a quadratic form and the $(4n) \times (4n)$ matrix $\mathcal{Q}(\Lambda)$ is such that $J\mathcal{Q}(\Lambda)$, with J the standard symplectic $(2n) \times (2n)$ matrix, has only purely imaginary eigenvalues $\pm {\rm i}\Omega_1, \cdots, \pm {\rm i}\Omega_{2n}$.

In general, the secular system is not integrable and its dynamics may be very complicated. Due to its elliptic equilibrium, it is natural to ask whether it is possible to find another neighborhood $\tilde{B}$ of the origin and a symplectic change of variables of the form

$$\phi^{\mathrm{P}}_{\mathrm{bnf}}: \quad (\tilde{\Lambda}, \tilde{\lambda}, \tilde{\mathrm{z}}) \in \tilde{\mathcal{P}} = \mathcal{A} \times \mathbb{T}^n \times \tilde{B} \to (\Lambda, \lambda, \mathrm{z}) \in \mathcal{P} = \mathcal{A} \times \mathbb{T}^n \times B$$

which sends $\{\tilde{\mathrm{z}} = 0\}$ to $\{\mathrm{z} = 0\}$ and conjugates $\mathcal{H}_{\mathrm{P}}$ to

$$\mathcal{H}_{\mathrm{P}} \circ \phi^{\mathrm{P}}_{\mathrm{bnf}} =: \mathcal{H}_{\mathrm{bnf}} = h_{\mathrm{Kep}} + \mu f^{\mathrm{P}}_{\mathrm{bnf}}, \tag{1.7}$$

so that $(f^{\mathrm{P}}_{\mathrm{bnf}})_{\mathrm{av}}$ is in "Birkhoff normal form" of some order $2s$ around this equilibrium ($s \in \mathbb{N}$). Namely, letting $\tilde{\mathrm{z}} = (\tilde{\mathrm{u}}, \tilde{\mathrm{v}})$, $(f^{\mathrm{P}}_{\mathrm{bnf}})_{\mathrm{av}}$ is a polynomial of degree s in $(\frac{\tilde{\mathrm{u}}_1^2+\tilde{\mathrm{v}}_1^2}{2}, \cdots, \frac{\tilde{\mathrm{u}}_{2n}^2+\tilde{\mathrm{v}}_{2n}^2}{2})$, plus a remainder of order $\mathrm{O}\big((\tilde{\mathrm{u}}, \tilde{\mathrm{v}})^{2s+1}\big)$. As for the possible transformations $\phi^{\mathrm{P}}_{\mathrm{bnf}}$ realizing (1.7) (which, when existing, are not not unique), it is common to look for those of the form

$$\phi^{\mathrm{P}}_{\mathrm{bnf}}: \quad \Lambda = \tilde{\Lambda}, \quad \lambda = \tilde{\lambda} + \varphi(\tilde{\Lambda}, \tilde{\mathrm{z}}), \quad \mathrm{z} = \mathrm{S}(\tilde{\Lambda})(\tilde{\mathrm{z}} + \mathrm{Z}(\tilde{\Lambda}, \tilde{\mathrm{z}})) \tag{1.8}$$

where $\varphi: \mathcal{A} \times \tilde{B} \to \mathbb{T}^n$, $\mathrm{Z}: \mathcal{A} \times \tilde{B} \to \mathbb{R}^{2n}$, with $\mathrm{Z}(\tilde{\Lambda}, \tilde{\mathrm{z}}) = \mathrm{O}(\tilde{\mathrm{z}}^2; \tilde{\Lambda})$ are suitable smooth function in $\tilde{B}$ and $\mathrm{S}(\tilde{\Lambda})$ is a real-valued symplectic[3)] matrix function on $\tilde{\Lambda}$ of order $(2n) \times (2n)$ such that the transformation $\mathrm{z} = \mathrm{S}(\Lambda)\hat{\mathrm{z}}$ puts the quadratic part of $(f_{\mathrm{P}})_{\mathrm{av}}$ into the form $\sum_{i=1}^{2n} \Omega_i(\Lambda) \frac{\hat{\mathrm{u}}_i^2+\hat{\mathrm{v}}_i^2}{2}$, where $\hat{\mathrm{z}} = (\hat{\mathrm{u}}, \hat{\mathrm{v}})$. Such definitions are standard at least for systems depending only on the z-variables [3, 23, 46] (i.e., neglecting the projection of $\phi^{\mathrm{P}}_{\mathrm{bnf}}$ over (Λ, λ) variables); the extension to properly-degenerate systems as given in (1.8) being quite straightforward, since it calls for a[4)] "natural" procedure in order to achieve (1.7). Actually, this definition has been implicitly used[5)] by several authors [5, 18, 22, 42], relatively to the same context considered here. Let us recall that, as soon as "a" $\phi^{\mathrm{P}}_{\mathrm{bnf}}$, as in (1.8) were existing, $(f^{\mathrm{P}}_{\mathrm{bnf}})_{\mathrm{av}}$ would be uniquely defined[6)]. Since we expect [39] that $(f^{\mathrm{P}}_{\mathrm{bnf}})_{\mathrm{av}}$ would retain many properties of the dynamics of the system (1.1) and due to the relevant physical meaning of the secular equilibrium $\{\mathrm{z} = 0\}$, we refer to systems $\mathcal{H} = h_{\mathrm{Kep}} + \mu f_{\mathrm{bnf}}$ as in (1.7)–(1.8) (or, simply, to $(f_{\mathrm{bnf}})_{\mathrm{av}}$), as "the" planetary Birkhoff normal form (clearly, this does not exclude that different normal forms around different invariant objects may be studied).

The problem of the existence of the systems (1.7) was settled by V. I. Arnold [5] and involved efforts by M. R. Herman, J. Laskar, P. Robutel, F. Malige [22, 27, 30, 42]. Its solution has been achieved in [16, 38]. Let us recall the historical background around this problem and facts that are necessary to explain the results of the paper. For more details, we refer to the review papers [10, 19].

1.2. In 1962 V. I. Arnold announced (International Congress for Mathematicians, Stockholm, [24]) and one year later published his more than celebrated "theorem on the stability of planetary motions"; or the "Planetary Theorem", for short.

Theorem 1 (V. I. Arnold, [5, p. 127]). *In the n-body problem there exists a set of initial conditions having positive Lebesgue measure and such that, if the initial positions and velocities belong to this set, the distances of the bodies from each other will remain perpetually bounded.*

---

3) It verifies $\mathrm{S}^{\mathrm{t}} J \mathrm{S} = J$, with "t" denoting transpose and J the standard symplectic matrix.

4) In general, having a properly-degenerate system $\mathcal{H} = h(\Lambda) + \mu f(\Lambda, \lambda, \mathrm{z})$, $(\Lambda, \lambda, \mathrm{z}) \in \mathcal{A} \times \mathbb{T}^n \times B$, $\mathcal{A} \subset \mathbb{R}^n$, $0 \in B \subset \mathbb{R}^{2m}$, n, $m \in \mathbb{N}$, with f_{av} having an elliptic equilibrium in $\mathrm{z} = (\mathrm{u}, \mathrm{v}) = 0$, one first considers [46] a transformation $\hat{\tau}: \ z = \mathrm{S}(\Lambda)\hat{\mathrm{z}}$ that preserves $d\mathrm{u} \wedge d\mathrm{v}$ for any Λ, so as to put the quadratic form of f_{av} in the form $\sum_i \Omega_i \frac{\hat{\mathrm{u}}_i^2+\hat{\mathrm{v}}_i^2}{2}$. Next, *provided the first order Birkhoff invariants Ω_i's do not verify resonances* (linear combinations with integer coefficients) *on $\mathcal{A}$ up to an order* $2s$, another smooth transformation $\tilde{\tau}: \ \hat{\mathrm{z}} = \tilde{\mathrm{z}} + \mathrm{Z}(\Lambda, \tilde{\mathrm{z}})$ preserving $d\hat{\mathrm{u}} \wedge d\hat{\mathrm{v}}$ may be found with $\mathrm{Z} = \mathrm{O}(\tilde{\mathrm{z}}^2; \Lambda)$ which puts the transformed averaged perturbation in Birkhoff normal form of order $2s$ [23]. The last (standard) step consists in producing a transformation of the form (1.8) which preserves $d\Lambda \wedge d\lambda + d\mathrm{u} \wedge d\mathrm{v}$.

5) In [18], with $\mathrm{Z} \equiv 0$.

6) Up to $\mathrm{O}\big((\tilde{\mathrm{u}}, \tilde{\mathrm{v}})^{2s+1}\big)$.

Arnold gave the details of the proof of the Planetary Theorem the case of three bodies constrained on a plane: the "first" non trivial case. The complete proof of this remarkable statement in the general case revealed to be much more difficult than expected for the strong degeneracies of the problem.

Arnold's ideas for proving Theorem 1 relied on two essential ingredients: (i) reduction of the system (1.3) to Birkhoff normal form of suitably high order[7)] to which to apply (ii) an abstract result of perturbation theory precisely suited to this situation. Arnold succeeded completely point (ii), leaving instead open, in the case of the problem in the space, the completion of (i). As for point (ii) he realized that the major technical difficulty to be solved was the proper degeneracy of the system (1.3). Arnold was aware that, because of this fact, previous results studied by Kolmogorov [26], Moser [31] or he himself [2] on the conservation of quasi-periodic motions with as many frequencies as the number of degrees of freedom could not be applied, since any non-degeneracy assumption required by such theorems would be dramatically violated[8)]. He then proved an abstract result (that he called the "Fundamental Theorem", see Appendix A; Theorem 8) stating the existence of a positive measure set of quasi-periodic motions ("Kolmogorov set") for systems in the form (1.7). He proved the existence of such motions under the condition that both the unperturbed term H_0 *and* the secular part $P_{\rm av}$ (respectively, $h_{\rm Kep}$ and $(f^{\rm P}_{\rm bnf})_{\rm av}$ in the application) of the perturbation should be non-degenerate in the sense of the Hessian. Namely, the matrix $\partial^2 H_0$ and the matrix β ("torsion") appearing in the expansion of $P_{\rm av}$ should be non-singular (see conditions (ii) and (iii) in Theorem 8).

Notice that the thesis of Theorem 8 establishes that the "density" of the invariant set $\mathcal{K}_{\mu,\epsilon}$ in phase space $\mathcal{P}_\epsilon$, i.e. the ratio $\operatorname{meas}\mathcal{K}_{\mu,\epsilon}/\operatorname{meas}\mathcal{P}_\epsilon$ depends on ϵ, the radius of $B = B^{n_2}_\epsilon$. This is a precise byproduct of the proper degeneracy of the system *and*, especially, of the fact that $P_{\rm av}$, in general, is non integrable, so that, at a certain point of the proof of Theorem 8, ϵ is to be used as a small parameter[9)]. Relatively to the application to the planetary problem, one should then expect that invariant tori accumulate more around the part of in phase space close to zero inclined and circular motions. See also the text Section 1.3 a) for more notices.

As mentioned above, Arnold checked the assumptions of the Fundamental Theorem in the case of the planar three-body problem. In the case the problem in the space, Arnold was aware that some extra-difficulty related to the "rotation invariance" of the system (1.1) was to be overcome. Namely, the invariance by the two-parameter group of (non-commuting) transformations

$$(y^{(i)}, x^{(i)}) \to (\mathcal{R}y^{(i)}, \mathcal{R}x^{(i)}), \quad \mathcal{R} \in \mathrm{SO}(3). \tag{1.9}$$

From a dynamical point of view, rotation invariance is caused by the conservation, along the $\mathrm{H}_{\rm plt}$-trajectories, of the three components, C_1, C_2 and C_3, of the "angular momentum"

$$\mathrm{C} = \sum_{i=1}^{n} x^{(i)} \times y^{(i)}, \tag{1.10}$$

where "$\times$" denotes skew-product.

Arnold realized [5, Chapter III, Section 5, n. 3, p. 141] that the two non-commuting integrals C_1 and C_2 cause another strong degeneracy (besides the proper degeneracy) in the system: one of the first order Birkhoff invariants associated to $(f_{\rm P})_{\rm av}$, $\Omega_{2n}(\Lambda)$, vanishes identically: there is a resonance, among the first order Birkhoff invariants Ω_j's, which is identically satisfied. And, moreover, another

7) Arnold proved the theorem with $2s = 6$, but indeed $2s = 4$ is enough [12].

8) Compare, for example, the condition studied in [2], where for a system $H(I,\varphi) = H_0(I) + \mu P(I,\varphi)$ in action–angle variables, he requires that one of the determinants $|\partial_I^2 H_0|$ or $\begin{pmatrix} \partial_I^2 H_0 & \partial H_0 \\ \partial H_0 & 0 \end{pmatrix}$ should be non-singular. Using, for example Delaunay coordinates, instead of Poincaré's, one would obtain a system in action–angle variables with $3n$ degrees of freedom, but H_0 depending only on n actions, causing the identically vanishing of the determinants above.

9) See [5, p.158], Eq. (4.1.15) and the bound for $\tilde{\tilde{H}}_1$ below, with the following correspondences: $\bar{\bar{H}}_1 + \tilde{\tilde{H}}_1$ is $P_{\rm av}$ of Theorem 8; $\bar{\bar{H}}_1$ is the (integrable) truncation of $P_{\rm av}$ of degree six; $\varepsilon = \epsilon^2$.

identically satisfied resonance appears (the origin of which turns to be much more mysterious than the rotational one; see however [1] for an investigation of this resonance) which, even though not mentioned in [5], was later pointed out by Michael Robert Herman: the sum of the remaining first invariants $\Omega_1(\Lambda), \cdots, \Omega_{2n-1}(\Lambda)$, vanishes identically. Such two resonances,

$$\Omega_{2n}(\Lambda) \equiv 0, \quad \sum_{i=1}^{2n-1} \Omega_i(\Lambda) \equiv 0,$$

which are proper of the problem in the space, are usually referred to, respectively, as "rotational", "Herman" resonance or, jointly, "secular resonances". Clearly, the secular resonances represent an obstruction to the construction of the Birkhoff normal form.

To overcome the problem of the secular resonances (or, at least, of the rotational one), Arnold proposed, in [5, Chapter III, Section 5, n. 4–5], a sketchy program of which he did not give the complete details. We anticipate that filling such details will reveal, in the next, to be not trivial at all, since indeed it will require new ideas, but at the end will be completely achieved [16, 27, 30, 38, 42]. Previously to [16, 38], a different, independent strategy of proof of Theorem 1 will be thought and successfully achieved by Michael Robert Herman and Jacques Féjoz [18, 22].

Arnold suggested two qualitatively different strategies to handle the case of two and the one of and more than two planets. For two planets, he suggested to use a classical tool known as "Jacobi reduction of the nodes", in order to reduce *all* the integrals of the system. Reducing the integrals corresponds to eliminate the cause of the vanishing eigenvalue and, as a byproduct, to lower the number of degrees of freedom of the system from six to four (this is in fact the minimum number of degrees of freedom in the spatial three-body problem, due to the fact that the integrals C_1, C_2 and C_3 do not mutually Poisson-commute). At a practical level, Jacobi reduction of the nodes corresponds to substituting, in the Hamiltonian, the Delaunay variables

$$\Theta_1 = \frac{\mathrm{G}}{2} + \frac{\Gamma_1^2 - \Gamma_2^2}{2\mathrm{G}}, \quad \Theta_2 = \frac{\mathrm{G}}{2} - \frac{\Gamma_1^2 - \Gamma_2^2}{2\mathrm{G}}, \quad \theta_1 = 0, \quad \theta_2 = \pi \tag{1.11}$$

and leave the remaining ones $(\Lambda_i, \Gamma_i, \ell_i, \mathrm{g}_i)$ $(i = 1,\ 2)$ unvaried. Here, $\mathrm{G} := |\mathrm{C}| = \sqrt{\mathrm{C}_1^2 + \mathrm{C}_2^2 + \mathrm{C}_3^2}$ is the Euclidean of C, which, being an integral of the system, is regarded as an "external parameter". Even though the substitution (1.11) is commonly attributed to Jacobi[10] [25], in a slightly different[11] setting, it was proved to leave the Hamilton equations unvaried with respect to the variables $(\Lambda, \Gamma, \ell, \mathrm{g})$ by Radau [41].

In the 90's Michael Robert Herman pointed out that the application of the formulae (1.11) to the spatial three-body problem as given in [5] contains a flaw. Since this flaw is related to one of the results of the paper, we shall provide more details about it in the next Section 1.3 a). As for the purposes of the present survey, we limit to mention that Arnold aimed to deduce conditions (ii) and (iii) of Theorem 8 of the *spatial* problem from that corresponding ones of the *planar* one that he had previously studied, arguing that the perturbing function of the *spatial* problem might be considered a small perturbation (for small inclinations) of the one of the *planar* one. The flaw was next repaired by Jacques Laskar and Philippe Robutel [27, 42] who, starting from Arnold's indications, constructed the normal form around the elliptic equilibrium of the reduced problem. As the authors pointed out, such equilibrium is not related to the equilibrium of the planar problem, since it corresponds to two inclined circular motions of the planets, whose mutual inclination cannot be taken arbitrarily close to zero (because Jacobi reduction is singular in this situation). See also [18, Proposition 81 and the comment below] for a technical discussion of this issue. A careful evaluation of the involved range of variations of all the relevant physical parameters (planets' masses ratio, semi-axes ratio and mutual inclination) is a remarkably nice aspect of the study in [42].

---

[10] The "original" reduction by Jacobi in [25] is not in the framework of Hamiltonian systems. It consists in a procedure for lowering the order (as number of equations) of the system of differential equations of the three-body problem.

[11] Instead of "reducing" Delaunay variables $(\Lambda, \Gamma, \Theta, \ell, \mathrm{g}, \theta)$, Radau wrote (the same) equations to reduce the variables $(\mathrm{R}, \Phi, \Theta, \mathrm{r}, \varphi, \theta)$, where (Θ, θ) are the same as in the Delaunay's set, while $(\mathrm{R}, \Phi, \mathrm{r}, \varphi)$, with $\Phi_i = \Gamma_i$, are related to $(\Lambda, \Gamma, \ell, \mathrm{g})$ via the "planar Delaunay map" described in (1.12) below.

In the case of more than two planets, where an analogue tool to Jacobi reduction was not available, Arnold conjectured [5, Chapter III, Section 5, n. 5] it were possible to reduce only two (out of three) non-commuting components of C (or functions of them). He believed this should let the system free of the vanishing eigenvalue. More precisely, Arnold imagined it were possible to construct a system of coordinates having analogue properties to Poincaré coordinates, but containing, among them, a couple of integrals (Φ_1, Φ_2) that should simultaneously play the rôle of Poincaré's coordinates $(\mathrm{p}_n, \mathrm{q}_n)$ and disappear from the Hamiltonian (being a couple of integrals), so as to eliminate the identically vanishing frequency Ω_{2n}. He briefly sketched the first step of a possible procedure for constructing (by series) such variables, but then stopped there. It was again Michael Robert Herman who realized this proof was not complete. Many years later, F. Malige, P. Robutel and J. Laskar [30] notice that, for the three-body case, Jacobi's reduction of the nodes may be decomposed in two steps, the first of which provides a set of variables on the "vertical angular momentum (invariant, ten-dimensional) manifold" $\mathrm{C}_1 = \mathrm{C}_2 = 0$ having analogue properties (apart for the fact they do not define the couple (Φ_1, Φ_2), because the "direction" of C is fixed) of Arnold's variables for this case. Analogously to Arnold's paper, for the case of more than two planets, the authors provide a an iterative procedure in order to obtain formal expansions of the coordinates.

Note that the two strategies that Arnold imagined for the two cases look very different: Jacobi reduction for $n = 2$ is *singular* (for planar motions) and reduces *completely* the integrals of the system, hence, *two* degrees of freedom; the procedure projected in [5, Chapter III, Section 5, n. 5] for $n \geqslant 3$ would, if existing, reduce only *one* degree of freedom and, apparently, seems *regular*. It will turn out [16, 38] that, actually, such strategies can be simultaneously achieved for any $n \geqslant 2$ and, moreover, are intimately related[12)].

The first complete proof of Arnold's Planetary Theorem in the general case appeared in [18], including efforts by Michael Robert Herman. As mentioned above, this important and beautiful result was reached with a different KAM technique, in particular, avoiding Birkhoff normal form: only the properties of the "frequency map" $(\Lambda_1, \cdots, \Lambda_n) \to (\partial h_{\mathrm{Kep}}, \Omega)$ associated to the system (1.3) are exploited in [18]. The underlying elegant, KAM Theory in [18] (for "smooth" systems) goes back to [43] (analytic) and exploits "non-planarity" or "gauchness" non-degeneracy conditions previously studied, since the late 60s and up to the 80s, by Pyartli, Arnold, Parasyuk, Sprinzuk and others; see [18] and references therein, for more information. Moreover, the problem of secular resonances is solved in [18] via arguments of abstract reductions, [4]. The fact of avoiding Birkhoff normal form was a ingredient precisely wanted by Michael Robert Herman, in order to simplify the proof avoiding annoying computations, [22]. However, he also investigated the aspect of planetary (at least, formal) torsion and expressed doubts on its non-singularity, since he claimed [22, p. 24] "*J' ignore si* $\det \beta$ *est identiquement nulle !* ".

With a similar proof, based on non-planarity of the frequency map, but using only real-analytic KAM theory in [43], Chierchia and Pusateri find real-analytic tori [14, 40]. Here, following an initial idea by Herman [22], secular degeneracies are reduced by adding an extra-integral to the Hamiltonian.

Another proof of Theorem 1, involving, as Arnold projected, the construction of the Birkhoff normal form for the system (1.3) was found in the PhD thesis [38]; the results of which were next published in the 2011's papers [12, 13, 16]. The starting point for this was a set of "action–angle" variables $(\Lambda, \Gamma, \Psi, \lambda, \gamma, \psi) \in \mathbb{R}^{3n} \times \mathbb{T}^{3n}$, qualitatively analogue to Delaunay variables, but better suited to the SO(3)-invariance of the Hamiltonian (1.1). Such variables exhibit, among their actions, the three components of the angular momentum, in a different form; precisely, $\Psi_{n-1} := |\mathrm{C}| = \mathrm{G}$, $\Psi_n := \mathrm{C}_3$ and $\psi_n := \zeta$, the longitude of the "node" (the intersection) of the plane orthogonal to C with a prefixed plane, when this is defined. Indeed, Ψ_{n-1}, Ψ_n and ψ_n, at contrast with C_1, C_2 and C_3, verify standard commutation rules[13)]. There follows that the three variables ψ_{n-1}, ψ_n and Ψ_n are *cyclic* in the system (1.1), which so is reduced to $(3n-2)$ degrees of freedom. In

[12)] Clearly, our remark deals with the explicit construction of symplectic coordinates. Indeed, reducing symmetries via the use of quotient spaces in Hamiltonian systems is always possible [4] and, for the planetary system, has been achieved in [18].

[13)] Ie, $\{\Psi_{n-1}, \Psi_n\} = \{\Psi_{n-1}, \psi_n\} = 0$, $\{\Psi_n, \psi_n\} = 1$.

particular, the couple (Ψ_n, ψ_n) disappears from from the Hamiltonian ("rotational degeneracy"). The use of the variables $(\Lambda, \Gamma, \Psi, \lambda, \gamma, \psi)$, substantially, reduces to the change in (1.11) in the case of $n = 2$ planets. Namely, when $n = 2$, the change of variables between the $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ and Delaunay variables $(\Lambda, \Gamma, \Theta, \ell, \mathrm{g}, \theta)$ *restricted to* $\Psi_n = \Psi_{n-1}$, $\psi_{n-1} = \psi_n = 0$ gives (1.11), besides $\gamma_i = \mathrm{g}_i$. This restriction has no influence on the Hamiltonian, and hence, in this setting, we recover, and extend to general $n \geqslant 2$, Radau's proof that Jacobi reduction of the nodes (1.11) preserves Hamilton equations, for the remaining variables.

The variables $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ may be obtained via a simple symplectic modification of analogue sets of variables $(\mathrm{R}, \Phi, \Psi, \mathrm{r}, \varphi, \psi)$, having the same (Ψ, ψ) that had been considered, in the 80's, by Francoise Boigey [8] for $n = 3$ and, in their full generality, by André Deprit [17] for $n \geqslant 4$. By their relation with the $(\mathrm{R}, \Phi, \Psi, \mathrm{r}, \varphi, \psi)$, we might call the $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ the "planetary" version of Boigey–Deprit variables or "Deprit's elements" [37]. Indeed, considering the symplectic "planar", "Delaunay" transformations that let the "two body Hamiltonians" $\frac{\mathrm{R}_i^2}{2m_i} + \frac{\Phi_i^2}{2m_i \mathrm{r}_i^2} - \frac{m_i M_i}{\mathrm{r}_i}$ into (1.2), one obtains the change

$$(\Lambda_i, \Gamma_i, \ell_i, \gamma_i) \to (\mathrm{R}_i, \Phi_i, \mathrm{r}_i, \varphi_i). \tag{1.12}$$

Precisely in this planetary form of $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$, Boigey–Deprit variables were rediscovered by the author, who was mainly stimulated by their application to this problem. Part[14)] of the proof of the symplectic character of the $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ found in [38] was later published in [13].

Differently from Delaunay variables, the elements $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ (as well as the "original" ones by Deprit) are not "six per planet", because of a complicated, hierarchical structure of planes and nodes. Incidentally, this structure led Deprit to formulate a very negative judgment and doubts on the possible practical usefulness of his variables [17, p. 194]. This complication might also be related to the fact that, except[15)] for the case $n = 2$, where they reduce to Jacobi's, Deprit's variables had never been applied before, in none of their two previous forms, to rotation invariant systems.

Analogously to what happens for Delaunay variables, also the variables $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ exhibit singularities when eccentricities or suitable mutual inclinations among the planes of the structure vanish. And, analogously to what Poincaré did for Delaunay variables, it is possible to describe also such singular situations switching to new symplectic variables, in part action–angle, in part rectangular, called "RPS variables" ("Regular", "Planetary" and "Symplectic"). They are denoted with analogue symbols as Poincaré variables[16)]

$$\begin{aligned} &\Lambda = (\Lambda_1, \cdots, \Lambda_n), \quad \lambda = (\lambda_1, \cdots, \lambda_n), \\ &z = (\eta_1, \cdots, \eta_n, \xi_1, \cdots, \xi_n, p_1, \cdots, p_n, q_1, \cdots, q_n), \end{aligned} \tag{1.13}$$

since they are qualitatively similar to them. Indeed, also in this set the couples (Λ_i, λ_i)'s are related to semi-major axes and (suitable) mean-longitudes on the instantaneous ellipses $\mathcal{E}_i$; the (η_i, ξ_i)'s to their eccentricities and perihelia; the (p_i, q_i)'s, to the directions of their planes. These latter couples are however very different from their analogue ones in Poincaré's set: the first $(n-1)$ couples of (p_i, q_i)'s are related to the *mutual* inclinations among planes; the last one (p_n, q_n) to the direction of C with respect to a prefixed frame. RPS variables are well fitted to rotation invariance of the problem, since (p_n, q_n) is a couple of integrals of the motion. Indeed, p_n and q_n are functions only of the integrals $\Psi_{n-1} = \mathrm{G}$, $\Psi_n = \mathrm{C}_3$ and $\psi_n = \zeta$:

$$\begin{cases} p_n = \sqrt{2(\mathrm{G} - \mathrm{C}_3)} \cos\zeta, \\ q_n = -\sqrt{2(\mathrm{G} - \mathrm{C}_3)} \sin\zeta. \end{cases}$$

---

14) The variables $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ were recovered in [38] in two steps. For $n = 2$, looking for a symplectic change of variables $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi) \to (\Lambda, \Gamma, \Theta, \ell, \mathrm{g}, \theta)$ from Delaunay's to the new ones, so as to fix the new actions (Λ, Γ, Ψ), which had previously checked to Poisson-commute. This part of the proof, containing elementary however lengthy computations, was never published, nor in [38]. After realizing the existence of Deprit's paper [17], it was substituted with a shorter proof, both in [38] and in [13]. The case of $n \geqslant 3$, which goes by induction, was interely published in [13].

15) See, *e.g.*, [20].

16) Clearly, the variables (1.13) were not discussed in [8]–[17], where the $(\Lambda, \Gamma, \Psi, \ell, \gamma, \psi)$ were not considered.

Another amusing common aspect of RPS variables with Poincaré's, besides regularity[17], is that the system retains D'Alembert rules, as it happens for the system (1.3). More details on the relation between Poincaré's and RPS variables may be found in [15].

Clearly, RPS variables play the rôle of the set that Arnold conjectured in [5, Chapter III, Section 5, n. 5] and (p_n, q_n) the one of (Φ_1, Φ_2).

At this point, the proof goes as follows. In place of the "Poincaré Hamiltonian" (1.3), we consider the "RPS Hamiltonian"

$$\mathcal{H}_{\mathrm{rps}} = h_{\mathrm{Kep}}(\Lambda) + \mu f_{\mathrm{rps}}(\Lambda, \lambda, \bar{z}) \tag{1.14}$$

with

$$\bar{z} = (\eta, \xi, \bar{p}, \bar{q}), \tag{1.15}$$

where $\eta = (\eta_1, \cdots, \eta_n)$, $\bar{p} = (p_1, \cdots, p_{n-1})$ and so on. The variables (p_n, q_n), which do not appear, are kept fixed once forever. The system (1.14) has $(3n-1)$ degrees of freedom and an extra-integral, $\mathrm{G} = |\mathrm{C}|$, which is a linear function of Λ_1, $\cdots$, Λ_n, $\frac{\eta_1^2+\xi_1^2}{2}$, $\cdots$, $\frac{p_{n-1}^2+q_{n-1}^2}{2}$:

$$\mathrm{G} = \sum_{i=1}^{n} \Lambda_i - \sum_{i=1}^{n} \frac{\eta_i^2 + \xi_i^2}{2} - \sum_{i=1}^{n-1} \frac{p_i^2 + q_i^2}{2}.$$

D'Alembert rules imply that $\bar{z} = 0$ (which now corresponds to zero eccentricities and inclinations with respect to C) is an elliptic equilibrium point for $(f_{\mathrm{rps}})_{\mathrm{av}}$. We then construct the Birkhoff normal form of order four associated to (1.14) around the elliptic equilibrium $\bar{z} = 0$. Namely, we conjugate the system (1.14) to a system $\mathcal{H}_{\mathrm{bnf}}^{\mathrm{rps}} = h_{\mathrm{Kep}} + \mu f_{\mathrm{bnf}}^{\mathrm{rps}}$ analogue to (1.7), via a symplectic transformation $\phi_{\mathrm{bnf}}^{\mathrm{rps}}$ analogue to (1.8), but with $(4n-2)$ replacing $4n$, for the dimension of the $\bar{z}$-variables. This construction is possible since the first order Birkhoff invariants $(\Omega_1, \cdots, \Omega_{2n-1})$ associated to $(f_{\mathrm{rps}})_{\mathrm{av}}$ can be explicitly computed, at least asymptotically. They turn out not to verify resonances of any arbitrary order $2s$ on a domain $\mathcal{A}$ (depending on s) of semi-axes as in (1.6), besides Herman's. Since it has been proved [15] that $(\Omega_1, \cdots, \Omega_{2n-1})$ coincide with the first $(2n-1)$ first order Birkhoff invariants associated to $(f_{\mathrm{P}})_{\mathrm{av}}$ in (1.3), in particular, the non-resonance (up to Herman's) of $(\Omega_1, \cdots, \Omega_{2n-1})$ on $\mathcal{A}$ refines an analogue statement in [18], where it is proved that Herman resonance is the only one *identically* satisfied by $(\Omega_1, \cdots, \Omega_{2n-1})$. Herman resonance, however, does not prevent the construction of Birkhoff normal form (which so turns out to exist *for any choice of s*) since, due to the integral G, $(f_{\mathrm{rps}})_{\mathrm{av}}$ exhibits a symmetry[18] for which only resonances $\sum_{i=1}^{2n-1} k_i \Omega_i = 0$ with $\sum_{i=1}^{2n-1} k_i = 0$ are relevant for the construction of the Birkhoff normal form, while[19] Herman resonance is not in this class. The final step is to compute (asymptotically) β and check non-trivial torsion.

In this setting, it is also possible, at the cost of introducing a singularity in correspondence of co-planarity, to reduce completely the integrals of the system, eliminating the integral G. This gives rise to quasi-periodic motions bifurcating from a different equilibrium point with $(3n-2)$ independent frequencies. Such "full reduced" quasi-periodic tori, not mentioned in [5] had been previously found in [18], via a different technique.

Let us conclude this survey with two considerations.

---

[17] Following [30], the reduction performed by the RPS variables might be called "partial reduction", at contrast with the "full reduction", also discussed in [16, 38], that reduces the number of degrees of freedom to the minimum, $(3n-2)$. Pay attention not to confuse, however, the *regular* "partial reduction" performed by RPS variables with the elementary (but *singular*) reduction that can be obtained reducing the integral C_3 in Poincaré variables. This latter one does not exhibit a cyclic couple and has nothing to do with the aforementioned Arnold's claim in [5, Ch. 3, Section 5, 5].

[18] The Hamiltonian flow $(\Lambda_i, \lambda_i, \eta_i + \mathrm{i}\xi_i, p_j + \mathrm{i}q_j) \to (\Lambda_i, \lambda_i + \tau, (\eta_i + \mathrm{i}\xi_i)e^{-\mathrm{i}\tau}, (p_j + \mathrm{i}q_j)e^{-\mathrm{i}\tau})$, with $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n-1$, generated by G, corresponding to rotations around the C-axis. Note that in Poincaré's variables we have an analogue symmetry, which is the flow generated by $\mathrm{C}_3 = (\Theta_1 + \cdots + \Theta_n)$. Here, we have to replace (η, ξ), $(\bar{p}, \bar{q})$ with, respectively, $(\mathrm{\eta}, \mathrm{\xi})$, (p, q).

[19] In a different context, this observation is already in [30].

Firstly, we might ask how the Birkhoff normal form associated to the system (1.14) is related to the system (1.3).

Let

$$\phi^{\rm P}_{\rm rps}: \quad (\Lambda, \lambda, z) \to (\Lambda, \lambda, {\rm z})$$

denote the change of variables between the two sets and $\bar\phi^{\rm rps}_{\rm bnf}$ the (trivial) lift of "a" $\phi^{\rm rps}_{\rm bnf}$ as above over (Λ, λ, z), i.e., including also (p_n, q_n) via the identity. It has been proved [15] that $\phi^{\rm P}_{\rm rps}$ has the form (1.8) (see Theorem 5 below), therefore also the composition $\bar\phi^{\rm P}_{\rm bnf} := \phi^{\rm P}_{\rm rps} \circ \bar\phi^{\rm rps}_{\rm bnf}$ has the same form. Then we find, in a indirect way, that also the system (1.3) admits, via $\bar\phi^{\rm P}_{\rm bnf}$, the form (1.7), with $(f^{\rm P}_{\rm bnf})_{\rm av} = (f^{\rm rps}_{\rm bnf})_{\rm av}$, *independent of* $\frac{p_n^2+q_n^2}{2}$. This fact may be interpreted, on one side, as a remarkable counterexample to Birkhoff theory (the system (1.3) — indeed, $(f_{\rm P})_{\rm av}$ — is an example of resonant system with non-resonant Birkhoff normal form) and, on the other side, shows that, without reduction, the unique Birkhoff normal form that can be obtained for the system (1.3) is *degenerate at any order*. Namely, besides having an identically vanishing frequency, any coefficient of $(f^{\rm P}_{\rm bnf})_{\rm av}$ with one of the indices equal to $2n$ vanishes. In particular, the torsion of $(f^{\rm P}_{\rm bnf})$ vanishes identically and hence this answers negatively to the question raised by Herman in [22, p. 24] mentioned above. In general, this degeneracy of the normal form of the unreduced system is another aspect of the rotational degeneracy remarked above.

The second comment is that, even if symmetry considerations and some extra work on the computation of the $\Omega_1, \cdots, \Omega_{2n-1}$ for the system in Poincaré variables (1.3) might lead to the observation that the secular resonances do not affect the direct construction of the system[20] (1.7) starting with (1.3), nevertheless, the rotational degeneracy discussed above shows that such normal form would be useless.

1.3. This paper is concerned with a more detailed study of the planetary Birkhoff normal form. Before describing it, we anticipate two applications.

a) A "uniform" theorem on quasi-periodic motions. The former result of this paper is a refinement of the proofs found in [16, 38, 42] about quasi-periodic motions in the (spatial) three-body problem.

As mentioned in the previous section, both the proofs of Theorem 1 given in [42] for $n = 2$ and [16, 38] for $n \geqslant 2$ are based on the direct application of Theorem 8 or refined versions of it [12]. This implies that one has to take smaller and smaller values of the range of the secular variables (eccentricities and inclinations), in order to find a major number of tori.

Arnold realized that in some particular cases we might assert something more.

For example, he knew that, already in the case of the *planar* three-body problem, the smallness condition on eccentricities can be relaxed. In Arnold's words:

[5, Chapter III, Section 1, n.6, p. 128]. *"In the case of three bodies* [on a plane] *we can obtain stronger results (...). It turns out that it is not necessary to require the eccentricities to be small; all that is necessary is that they should be small enough to exclude the possibility of collision."*

And in fact, he stated

Theorem 2 (V. I. Arnold, [5, p. 128]). *Consider the case of the planar three-body problem. Fix positive numbers* $\underline{a}_1 < \overline{a}_1 < \underline{a}_2 < \overline{a}_2$. *Let* $\mathcal{D}_{\epsilon_0}$ *be a domain*[21] *defined by*

$$\mathcal{D}_{\epsilon_0}: \quad \underline{a}_1 \leqslant a_1 \leqslant \overline{a}_1, \quad \underline{a}_2 \leqslant a_2 \leqslant \overline{a}_2, \quad \sum_{i=1}^{2} m_i\sqrt{M_i a_i}(1-\sqrt{1-e_i^2}) \leqslant \epsilon_0^2$$

[20] At least at low orders. The fact that only resonances $\sum_{i=1}^{2n} k_i\Omega_i = 0$ with $\sum_{i=1}^{2n} k_i = 0$ are relevant for the construction of the Birkhoff normal form holds also for the system (1.3) in Poincaré variables. The resonance with $k_1 = \cdots = k_{2n-1} = 1$ and $k_{2n} = -(2n-1)$ prevents taking $2s = 2(2n-1)$.

[21] The second inequality is just $\frac{1}{2}\|(\eta, \xi)\|_2^2 \leqslant \epsilon_0^2$.

(for a suitably chosen ϵ_0, depending on $\bar{m}_1$, $\bar{m}_2$, $\overline{a}_1$, $\underline{a}_2$ so that collisions[22] *are excluded). For any $\kappa > 0$ it is possible to find $\mu_* > 0$ such that if*

$$0 < \mu \leqslant \mu_* \tag{1.16}$$

an invariant set $\mathcal{K}_\mu \subset \mathcal{D}_{\epsilon_0}$, with

$$\operatorname{meas} \mathcal{K}_\mu \geqslant (1 - \kappa) \operatorname{meas} \mathcal{D}_{\epsilon_0}$$

formed by the union of invariant four-dimensional tori, on which the motion is analytically conjugated to linear Diophantine quasi-periodic motions.

The main point of this theorem is that κ does not depend on ϵ_0, which in turn related to the bound on eccentricities, but only on μ_* (incidentally, we note that, even if Arnold does not specify the relation between μ_* and κ, in the statement, following the proof, one finds $\mu_* \sim \kappa_*^{1/a_*}$, with some small number a_*). The key point [5, Chapter III, Section 5, n. 2, p. 139] of the proof of Theorem 2 is that, for this problem, $(f_{\mathrm{P}})_{\mathrm{av}}$ is *integrable*, since it has two degrees of freedom (the two eccentricities) and two commuting integrals: the third component C_3 of C and itself. Using this ingredient, he shows that, via a simple modification[23] of the proof of Theorem 8, he proves a less general result than Theorem 8 but however very useful in this case: Compare Theorem 9 in Appendix A. In the thesis of Theorem 9 ϵ_0, the radius of $B^{n_2}_{\epsilon_0}$, does not appear in the measures ratio $\operatorname{meas} \mathcal{K}_\mu / \operatorname{meas} \mathcal{P}_{\epsilon_0}$.

Arnold believed that Theorem 2 admitted a generalization to the case of the *spatial* three-body problem. His generalization should go as follows. Letting $\mathcal{D}'_\epsilon$ to be the set where semi-major axes (a_1, a_2), eccentricities (e_1, e_2) and inclinations (ι_1, ι_2) of the two planets verify inequalities

$$\mathcal{D}'_\epsilon : \quad (a_1, a_2, e_1, e_2) \in \mathcal{D}_{\epsilon_0}, \quad \sum_{i=1}^{2} m_i \sqrt{M_i} \iota_i^2 \leqslant \epsilon^2$$

he claimed

"Theorem" 1 ([5, Chapter III, Section 1, n. 7, p. 129]). *"An analogous theorem* [to Theorem 2] *is valid for the space three-body problem. In this case $\mathcal{D}'_\epsilon$ is defined* (...) [as above] *with sufficiently small ϵ."*

Arnold did not provide the exact statement of this "theorem". From the underlying context it might be argued that he believed it were possible replace condition (1.16) with

$$0 < \mu \leqslant \mu_*, \quad 0 < \epsilon \leqslant \epsilon_*$$

(with, possibly, μ_*, $\epsilon_* \sim \kappa^{1/a'_*}$), so as to have an invariant set $\mathcal{K}'_{\mu,\epsilon}$ with larger and larger measure *only letting masses μ and inclinations $\epsilon \to 0$*, while keeping eccentricities finite. In fact, he aimed to use Theorem 9 also for this case, since we read, as a "proof" of "Theorem" 1:

[5, Chapter III, Section 5, n. 4, p. 141]. *"The space three-body problem reduces to a certain plane problem which turns into the plane three-body problem when the inclinations tend to* 0. *By comparison with* 2. [the averaged plane three-body problem is integrable] *and using again Ch. I, Section* 8 [Theorem 9]*, we arrive at the results of Section* 1*, n.* 7 ["Theorem" 1]*."*

The first sentence of this "proof" contains a flaw. Later results for the spatial three-body problem [16, 38, 42], being based on Theorem 8 (rather than on Theorem 9) are actually weaker than "Theorem" 1, since they do not state that the measures ratio $\operatorname{meas} \mathcal{K}_{\mu,\epsilon} / \operatorname{meas} \mathcal{P}_\epsilon$ depends only on the planets' masses and on the inclinations and does not depend on the eccentricities.

---

[22] See the first equation after [5, Eq. (3.1.5)], with ϵ_0 replaced by ϵ_0^2.

[23] I. e., replacing in [5, Eq. (4.1.15), p. 158] $\bar{\bar{H}}_1$ with $\bar{\bar{H}}_1^* := \bar{\bar{H}}_1 + \tilde{\bar{H}}_1$, $\tilde{\bar{H}}_1$ with 0 and, in the Equation just below, neglecting the bound on $\tilde{\bar{H}}_1$ and replacing the bound on H_2 with $\mu^2 \delta^{-T}$, $T \gg 1$, as it follows from the "Lemma on averaging over rapid variables" [5, Chapter IV, Section 5, p. 147] with $M = \bar{M} = \mu$, to which we also refer for notations.

In this paper, we shall prove a result which goes in the direction of “Theorem” 1, but is not quite the same. In a sense, we replace ϵ, the bound on inclinations, with α, the bound on the semi-major axes ratio. We also extend such result to the planar, general problem with[24] $n \geqslant 3$ planets.

Theorem A. *In the spatial three-body problem, there exist positive numbers α_*, μ_*, ϵ_*, c_*, $1/C_*$ and $1/\beta_*$ smaller than one such that, if the numbers α and μ (where μ is the parameter in (1.1)) verify*

$$0 < \mu \leqslant \mu_*, \quad 0 < \alpha \leqslant \alpha_*, \quad \mu \leqslant c_* \log(\alpha^{-1})^{-4\beta_*}$$

for any choice of a_-, a_+, with $a_- < \alpha a_+$, in the domain $\mathcal{D}_\alpha$ where semi-axes a_1, a_2, eccentricities e_1, e_2 and mutual inclination ι verify

$$\mathcal{D}_\alpha: \quad a_- \leqslant a_1 \leqslant \alpha\, a_2 \leqslant \alpha a_+, \quad |(e_1, e_2, \iota)| < \epsilon_*$$

a set $\mathcal{K}_{\mu,\alpha} \subset \mathcal{D}_\alpha$ may be found, formed by the union of invariant 5-dimensional tori, on which the motion is analytically conjugated to linear Diophantine quasi-periodic motions. The set $\mathcal{K}_{\mu,\alpha}$ is of positive Liouville–Lebesgue measure and satisfies

$$\operatorname{meas} \mathcal{K}_{\mu,\alpha} \geqslant \Big(1 - C_*(\sqrt[4]{\mu}(\log \alpha^{-1})^{\beta_*} + \sqrt{\alpha})\Big) \operatorname{meas} \mathcal{D}_\alpha.$$

The same assertion holds for the planar $(1+n)$-body problem.

b) A “full” Nehorošev stability theorem The latter result of the paper is concerned with the stability for the planetary system. To introduce it, we recall the following fundamental result by N. N. Nehorošev[25], mainly motivated by its application to the Hamiltonian (1.3).

Theorem 3 (N. N. Nehorošev, 1977, [33, 34]). *Let*

$$H(I, \varphi, p, q) = H_0(I) + \mu P(I, \varphi, p, q), \quad (I, \varphi, p, q) \in \mathcal{P} \subset \mathbb{R}^{n_1} \times \mathbb{T}^{n_1} \times \mathbb{R}^{2n_2}$$

be of the form of (1.3), real-analytic. Assume that $H_0(I)$ is “steep”. Then, one can find a, $b > 0$, C and μ_0 such that, if $\mu < \mu_0$, any trajectory $t \to \gamma(t) = (I(t), \varphi(t), p(t), q(t))$ solution of H such that

$$(p(t), q(t)) \in \Pi_{(p,q)}\mathcal{P}, \quad \forall\ 0 \leqslant t \leqslant T_0 := \frac{1}{C\mu} e^{\frac{1}{C\mu^a}} \tag{1.17}$$

verifies

$$|I(t) - I(0)| \leqslant r_0 := \frac{C}{2}\mu^b \qquad \forall\ 0 \leqslant t \leqslant T_0.$$

As for the definition of “steepness”, we refer to the papers [32–34]; see also [36] for an equivalent definition. We aim to point out that, despite of the almost 150-pages length of the proof of Theorem 3 and the complication of notion of steepness, in [33] Nehorošev easily[26] applied Theorem 3 to the planetary Hamiltonian $\mathcal{H}_{\mathrm{P}}$ in (1.3) (with $I = \Lambda$, the actions related to the semi-axes, and $(p, q) = \mathrm{z}$ in (1.5), the secular variables), since the unperturbed term $H_0 = h_{\mathrm{Kep}}$ is *concave*, a special case of steepness. Nehorošev then obtained a spectacular result of stability for the planetary semi-axes (implying, in particular, absence of collisions) over exponentially-long times *for all initial data* in phase space (see also [35] for a different approach and improved estimates). Up no now, Nehorošev’s result is the only rigorous, global (i.e., valid on the whole phase space, or, possibly, on a very large open subset of it) stability result for the planetary problem. Indeed, there do exist

[24] The proof holds also for $n = 2$, but for this case Theorem 2 gives a stronger result.

[25] A more technical statement of Theorem 3 is given in Appendix D: Compare Theorem 11. Recall that other improved statements of Theorem 3 have later been found in particular cases: see, for example, [9, 39] and references therein.

[26] The only delicate point in the application of Theorem 3 to $\mathcal{H}_{\mathrm{P}}$ consisted in checking assumption (1.17), that Nehorošev accomplished using the conservation of the third component C_3 of the total angular momentum (1.10) along the $\mathcal{H}_{\mathrm{P}}$-trajectories. Note that, in the non-degenerate case, i.e., when the variables (p, q) do not appear, this assumption is void.

in literature results involving also strong numerical efforts for physical systems (see, *e.g.*, [21, 44] and references therein) true on Cantor sets (in general, they are obtained via KAM techniques).

A physically relevant and widely studied open problem is related to the study of the stability of the whole system; i.e., the study of the secular variation of eccentricities and inclinations of the planets' instantaneous orbits, besides the ones of semi-axes. See, for example, [28] and references therein. Partial rigorous results in this direction have been obtained in [15], where it has been proved that, if eccentricities and inclinations are initially suitably small, they remain confined with respect to their initial values over *polynomially* long times, up to exclude the so-called[27)] "mean-motion resonances". More precisely, the following result has been proved.

Theorem 4 ([15]). *Whatever is the number of planets, for any arbitrarily fixed $s \in \mathbb{N}$, with $s \geqslant 5$, one can find positive numbers C, $\underline{a}_j$, $\overline{a}_j$, $\underline{\epsilon}$, $\overline{\epsilon}$ with $\underline{a}_j < \overline{a}_j < \underline{a}_{j+1}$ and $\underline{\epsilon} < \overline{\epsilon}$ such that for any $\kappa > 0$, $\underline{\epsilon} < \epsilon \leqslant \overline{\epsilon}$, in the domain where semi-major axes a_i, eccentricities a_i and mutual inclinations ι_j verify*

$$\hat{\mathcal{D}}_{s,\epsilon}: \quad \underline{a}_j \leqslant a_j \leqslant \overline{a}_j \quad \underline{\epsilon} \leqslant \max_{i,j}\{e_i, \iota_j\} \leqslant \epsilon$$

under suitable relations between μ and ϵ, one can find an open set $\hat{\hat{\mathcal{D}}}_{s,\mu,\epsilon}$ such that, for all the motions starting in $\hat{\hat{\mathcal{D}}}_{s,\mu,\epsilon}$, the displacement of eccentricities and inclinations with respect to their initial values is bounded by $\kappa\underline{\epsilon}$, for all

$$|t| \leqslant \frac{\mathrm{C}\kappa}{\mu\underline{\epsilon}^s}.$$

The proof of Theorem 4 again relies with the Birkhoff normal form of the system discussed in section 1.2; the time of stability is related in fact to the remainder of this normal form. No analysis of resonance zones, trapping arguments... is used for its proof. An undesirable aspect of Theorem 4, is that the size of $\hat{\hat{\mathcal{D}}}_{\sigma,\epsilon}$ decreases with with the time of stability.

In this paper, we prove a stronger result, at least for the planar three-body problem.

Theorem B. *In the planar three-body problem, there exist numbers $\bar{a}_-$, $\bar{\alpha}$, $\bar{\epsilon}$, $\bar{a}$, $\bar{b}$, $\bar{c}$, $\bar{d}$ such that, for any $\bar{a}_-$, $\bar{a}_+$, with $\bar{a}_- < \bar{\alpha}\bar{a}_+$ and $\underline{\epsilon} < \epsilon \leqslant \bar{\epsilon}$ in the domain*

$$\bar{\mathcal{D}}_\epsilon: \quad \bar{a}_- \leqslant a_1 \leqslant \bar{\alpha}\, a_2 \leqslant \bar{\alpha}\bar{a}_+, \quad \underline{\epsilon} \leqslant |(e_1, e_2)| \leqslant \epsilon$$

under suitable relations between μ and ϵ, one can find an open set $\bar{\bar{\mathcal{D}}}_{\mu,\epsilon} \subset \bar{\mathcal{D}}_\epsilon$, defined by absence of mean-motion resonances up to a suitable order, such that, for all the motions with initial datum in $\bar{\bar{\mathcal{D}}}_{\mu,\epsilon}$, one has

$$|a_i(t) - a_i(0)|,\ |e_i(t) - e_i(0)| \leqslant \bar{r} := \max\{\delta^{\bar{b}}, \mu^{1/12},\ \epsilon\} \quad \forall\ 0 \leqslant t \leqslant \bar{T} = \frac{e^{\frac{1}{\bar{\delta}^{\bar{a}}}}}{\bar{\delta}}$$

where $\bar{\delta} := \frac{\mu^{\bar{d}}\epsilon}{\bar{c}}$.

1.4. Let us sketch the proofs of Theorems A and B and make some comment.

The proof of Theorem A is a remake of Arnold's ideas for the proof of Theorem 2 described in Section 1.3 a). Let us denote as $f_{3\mathrm{b}}$ the function f_{rps} in (1.14) for the three-body case; $(f_{3\mathrm{b}})_{\mathrm{av}}$, its averaged value. We shall see below that a suitable approximation $(f_{3\mathrm{b}})_{\mathrm{av}}^{(2)}$ defined in Eq. (1.20) below, is *integrable*. This fact has been already used, in different settings, in [29, 47] and [37]. Moreover, the same property of integrability is proved to hold for the *planar* many-body problem; see below for more details on this assertion. Then, we apply a suitable slight generalization of Theorem 9 (Theorem 6), which allows to work on $(f_{3\mathrm{b}})_{\mathrm{av}}^{(2)}$, $(f_{\mathrm{pl}})_{\mathrm{av}}^{(2)}$, respectively.

The proof of Theorem B is an application of the Nehorošev's Theorem. Essentially, it relies on checking "steepness" of the first terms of the Birkhoff normal form associated to $h_{\mathrm{Kep}} + \mu(f_{3\mathrm{b}})_{\mathrm{av}}$, in the planar case, in all of its degrees of freedom. Here the difficulty is that, at contrast with

[27)] I.e., resonances of the Keplerian frequencies $\omega_{\mathrm{Kep}} := \partial h_{\mathrm{Kep}}$.

the application in [33] (where only the concavity of $h_{\rm Kep}$ is exploited), the "full torsion" of the system, given by the Hessian of $h_{\rm Kep}$ and the matrix β of the second-order Birkhoff invariants, *is not convex*, nor quasi-convex. Its eigenvalues are alternating in sign. Therefore, it is necessary to consider higher orders of Birkhoff normal form and apply more refined conditions for steepness. It is not clear, in general, what is the right order of the Birkhoff series to be involved and, especially, how steepness can be checked for systems with many degrees of freedom (see [45] for a study in this direction and [36] and references therein for a different approach). For three-degrees of freedom systems Nehoroševproved that the "three-jet condition" (recalled in Appendix D) is "generic". But the *planar* three-body problem, after reducing completely rotations, has three degrees of freedom, so it is not surprising that this problem satisfies three-jet. We do this check in Section 4.4.

Before passing to describe technical aspects, we provide a few comments.

– Theorem B is stated for the planar three-body problem. As previously outlined, the secular problem associated to it is *integrable*: its Birkhoff normal form converges. And in fact this circumstance allowed Arnold to obtain refined results for this case (see Section 1.3, a)). One might ask if, analogously to the result of Theorem 2, the set $\bar{\bar{\mathcal{D}}}_{\mu,\epsilon}$ may be chosen to be independent of ϵ. However, with our proof we are not able[28] to refine the result in that direction. The reason is technical: instead of the (integrable) secular system $\mathcal{H}_{\rm pl3b} := h_{\rm Kep} + \mu(f_{\rm pl3b})_{\rm av}$ that would be more natural, during the proof we consider a *non integrable* system close[29] to it, by performing not only one but *many* steps of averaging with respect to fast (mean motion) frequencies. Therefore, we need to *truncate* the Birkhoff series associated to this closely to integrable system and this is the reason we have the dependence of ϵ. Indeed[30], the estimates of theory developed in [33], which we use in a quantitative way during the proof of Theorem B (see Theorem 11), do not allow us to perform just one step of averaging.

– In Section 4 we do more than we need for Theorem B. We compute the Birkhoff normal form of the *spatial* three-body problem, which is[31]

$$
\begin{aligned}
(f_{\rm bnf})_{\rm av} = &-\frac{\bar m_1 \bar m_2}{a_2} - \bar m_1 \bar m_2 \frac{a_1^2}{4a_2^3}\Bigg(\left(1 + 3\frac{t_1}{\Lambda_1} + 3\frac{t_2}{\Lambda_2} - 3\left(\frac{1}{\Lambda_1} + \frac{1}{\Lambda_2}\right)t_3\right) \\
&- \bar m_1 \bar m_2 \frac{a_1^2}{4a_2^3}\Bigg(-\frac{3}{2}\frac{t_1^2}{\Lambda_1^2} + 6\frac{t_2^2}{\Lambda_2^2} + \frac{3}{2}\frac{t_3^2}{\Lambda_1^2} + 9\frac{t_1 t_2}{\Lambda_1\Lambda_2} - 12\frac{t_1 t_3}{\Lambda_1^2} - 9\frac{t_2 t_3}{\Lambda_1\Lambda_2} \\
&+ 10\frac{t_2^3}{\Lambda_2^3} - \frac{3}{2}\frac{t_3^3}{\Lambda_1^2\Lambda_2} - \frac{9}{2}\frac{t_1^2 t_2}{\Lambda_1^2\Lambda_2} - \frac{105}{4}\frac{t_1^2 t_3}{\Lambda_1^3} - 18\frac{t_2^2 t_3}{\Lambda_1\Lambda_2^2} + 18\frac{t_1 t_2^2}{\Lambda_1\Lambda_2^2} \\
&+ \frac{105}{4}\frac{t_1 t_3^2}{\Lambda_1^3} + \frac{9}{2}\frac{t_2 t_3^2}{\Lambda_1^2\Lambda_2} - 36\frac{t_1 t_2 t_3}{\Lambda_1^2\Lambda_2}\Bigg)\left(1 + {\rm O}\left(\frac{\Lambda_1}{\Lambda_2}\right)\right) + \frac{a_1^2}{a_2^3}{\rm O}\left(|t|^{7/2}\right) + {\rm O}\left(\frac{a_1^3}{a_2^4}\right)\Bigg)
\end{aligned}
\tag{1.18}
$$

and then we reduce to the planar case setting $t_3 = 0$. However, we are not able to extend Theorem B to the spatial case, since we are not able to check steepness for this case. The three-jet condition might fail at least on manifolds of co-dimension one: see Remark 3.

– Besides the previous case, a possible extension of Theorem B to the general *planar* problem might be helped by the fact that, for this case *we know a good approximation of* $(f_{\rm bnf})_{\rm av}$, *at any order*. This result is a corollary of the analysis of Section 2. See also Section 1.5 below.

– In our strategy of proofs, the planetary Birkhoff normal form (hence, the system (1.14) in RPS variables) plays a central rôle. The author is not aware (and would be interesting to know) what kind of results could be obtained (and what would be the relative difficulty) via Herman–Féjoz's normal form [18].

---

[28] The dependence of $\bar{\bar{\mathcal{D}}}_{\mu,\epsilon}$ on ϵ may be read in inequality just before (4.3) and by the formula (4.3), that define this set.

[29] Compare the system $h_{\rm Kep} + \mu(\hat N + \hat N_*)$ in (4.14).

[30] Compare Lemma 5 in Section 4.3.

[31] In particular, truncating this formula to the fourth order we recover the formulae found in [16, 38].

1.5. The main novelty of this paper (with respect to our previous ones on this subject) is a technical lemma of geometrical nature (Proposition 1, which, in turn is a consequence of the more general Proposition 2. See also the second item in Remark 1) that helps in the analysis of the secular perturbing function of the system (1.14). This reflects on the computation of the Birkhoff invariants at higher orders.

Let us remark, in this respect that, in general, computing the Birkhoff invariants of the planetary problem is a huge work. See, for example the computations of the torsion in [5] ($n = 2$, planar), [42] ($n = 2$, spatial), [22] ($n \geqslant 2$, planar), [16, 38] ($n \geqslant 2$, spatial). So, our main progress relies on an improvement of the technique of computation of such invariants, which is particularly desirable if one wants to extend Theorem B to the general problem.

Let us introduce it briefly, referring to the following section for details.

Consider the system (1.14) and, in particular, its secular perturbing function $(f_{\rm rps})_{\rm av}$. Since the indirect[32)] part has zero λ-average, $(f_{\rm rps})_{\rm av}$ is given by

$$(f_{\rm rps})_{\rm av} = -\sum_{1\leqslant i<j\leqslant n} \frac{\bar m_i \bar m_j}{(2\pi)^2}\int_{\mathbb{T}^2}\frac{d\lambda_i d\lambda_j}{|x^{(i)}(\Lambda,\lambda_i,\bar z) - x^{(j)}(\Lambda,\lambda_j,\bar z)|}.$$

De-homogeneizating with respect to a_j, we expand each of the terms

$$(f^{(ij)}_{\rm rps})_{\rm av} := -\frac{\bar m_i \bar m_j}{(2\pi)^2}\int_{\mathbb{T}^2}\frac{d\lambda_i d\lambda_j}{|x^{(i)}(\Lambda,\lambda_i,\bar z) - x^{(j)}(\Lambda,\lambda_j,\bar z)|}$$

in powers of the ratio $\frac{a_i}{a_j}$, with a_j fixed:

$$(f^{(ij)}_{\rm rps})_{\rm av} = (f^{(ij)}_{\rm rps})^{(0)}_{\rm av} + (f^{(ij)}_{\rm rps})^{(1)}_{\rm av} + (f^{(ij)}_{\rm rps})^{(2)}_{\rm av} + \cdots . \tag{1.19}$$

Clearly, to this expansion there corresponds an analogue expansion of

$$(f_{\rm rps})_{\rm av} = (f_{\rm rps})^{(0)}_{\rm av} + (f_{\rm rps})^{(1)} + (f_{\rm rps})^{(2)}_{\rm av} + \cdots . \tag{1.20}$$

Analogously to what happens for the Poincaré Hamiltonian (1.3), one has that, in these expansions, the zeroth order terms $(f^{(ij)}_{\rm rps})^{(0)}_{\rm av}$ are independent[33)] of $\bar z$ by well known properties of the two-body potential and that the linear terms $(f^{(ij)}_{\rm rps})^{(1)}_{\rm av}$ vanish by Fubini's and Newton equation[34)]. The lowest order information on $(f_{\rm rps})_{\rm av}$ is then given by the second-order terms $(f_{\rm rps})^{(2)}_{\rm av}$.

By [16, 38] $(f_{\rm rps})^{(2)}_{\rm av}$ may be splitted into a sum

$$(f_{\rm rps})^{(2)}_{\rm av} = (f_{\rm pl})^{(2)}_{\rm av} + (f_{\rm vert})^{(2)}_{\rm av} \tag{1.21}$$

of a "planar" and[35)] a "vertical" part, where $(f_{\rm pl})^{(2)}_{\rm av}$ corresponds to the term that we would have for the problem in the plane, while $(f_{\rm vert})^{(2)}_{\rm av}$ vanishes for $(\bar p, \bar q) = 0$ and is even in $(\bar p, \bar q)$. In Section 2

[32)] The former term in (1.4) is of often referred to as "indirect part"; the latter as "direct part". As far as the author knows, this terminology has been introduced by the French school. The vanishing of the average of the indirect part, known Poincaré variables, holds also in RPS variables.

[33)] They are given by given by$-\frac{\bar m_i \bar m_j}{a_j}$.

[34)] I. e., by the vanishing of

$$\frac{1}{2\pi}\int_{\mathbb{T}}\frac{x^{(j)}(\Lambda,\lambda_j,\bar z)}{|x^{(j)}(\Lambda,\lambda_j,\bar z)|^3}d\lambda_j = \frac{1}{{\rm T}_j}\int_0^{{\rm T}_j}\frac{d}{dt}y^{(j)}(\Lambda,\omega_j t,\bar z)dt$$

with some ${\rm T}_j$ and $\omega_j = \frac{2\pi}{{\rm T}_j}$.

[35)] We follow the terminology in [18].

we prove that $(f_{\rm pl})^{(2)}_{\rm av}$, $(f_{\rm vert})^{(2)}_{\rm av}$ are given by, respectively,

$$
\begin{aligned}
(f_{\rm pl})^{(2)}_{\rm av} &= -\frac{1}{4}\sum_{1\leqslant i<j\leqslant n}\bar m_i\bar m_j\frac{a_i^2}{a_j^3}\frac{1+\frac{3}{2}e_i^2}{(1-\frac{\eta_j^2+\xi_j^2}{2\Lambda_j})^3}\\
(f_{\rm vert})^{(2)}_{\rm av} &= +\frac{3}{4}\sum_{1\leqslant i<j\leqslant n}\bar m_i\bar m_j\frac{a_i^2}{a_j^3}\frac{\frac{1}{2\pi}\int_{\mathbb T}(\hat x^{(i)}\cdot\hat{\rm C}^{(j)})^2d\lambda_i}{\left(1-\frac{\eta_j^2+\xi_j^2}{2\Lambda_j}\right)^3},
\end{aligned}
\tag{1.22}
$$

where e_i's are the eccentricities, expressed in terms of Λ_i and $\frac{\eta_i^2+\xi_i^2}{2}$; $\hat{\rm C}^{(j)}$ are the planets' normalized angular momenta $\frac{{\rm C}^{(j)}}{|{\rm C}^{(j)}|}$ and $\hat x^{(i)} := \frac{x^{(i)}(\Lambda,\lambda_i,z)}{a_i}$.

The author is not aware if the formulae (1.22) had been already noticed before (they hold also in the case of the Poincaré system (1.3)). Such formulae are the thesis of Proposition 1, that we prove using a new set of symplectic variables, defined in (2.10), and tools of normal form theory. The variables (2.10) in a sense resemble the well known Adoyer–Deprit variables of the rigid body, with the difference that have six degrees of freedom instead of three. Also the thesis of Proposition 1 resembles certain formulae for the rigid body, as outlined in Remark 1.

In particular, inspecting (1.22), it is to be remarked that $(f_{\rm pl})^{(2)}_{\rm av}$ not only is *integrable*, but *is in Birkhoff normal form*. This fact implies the validity of Theorem A for the planar general problem. Moreover, since this formula is of great help in the computation of its Birkhoff invariants *at any order*, applications to extension of Theorem B to this case are foreseen.

Secondly, formulae (1.22) imply that, in the three-body case ($n = 2$), $(f_{\rm 3b})^{(2)}_{\rm av} := (f_{\rm rps})^{(2)}_{\rm av}|_{n=2}$ is independent of the argument of (η_2,ξ_2), therefeore, it is *integrable* (compare [29], for an analogue assertion in a different setting and, *e.g.*, [20, 37, 47] for applications). More in general, for $n \geqslant 2$, $(f_{\rm rps})^{(2)}_{\rm av}$ is independent on the argument of (η_n,ξ_n). But while, for this general case, the expression of $(f_{\rm vert})^{(2)}_{\rm av}$ in terms of RPS variables is complicated, due to the factors $(\hat x^{(i)}\cdot\hat{\rm C}^{(j)})^2$, it is not so for three bodies, where there is only one of such factors ($i = 1$, $j = 2$). The aspect of the corresponding vertical term is nice:

$$
(f_{\rm 3bvert})^{(2)}_{\rm av} = \frac{3}{4}\bar m_1\bar m_2\frac{a_1^2}{a_2^3}\frac{1}{(1-\frac{\eta_2^2+\xi_2^2}{2\Lambda_2})^3}\Big((1+\frac{3}{2}e_1^2)({\rm i}vv^\star)+\frac{5}{2}\big((u_1^\star)^2v^2+(v^\star)^2u_1^2\big)\bar e_1^2\Big)\bar{\mathfrak s}^2 \tag{1.23}
$$

where u_i, u_i^* are the Birkhoff variables associated to (η_i,ξ_i); (v,v^*) to (p_1,q_1), $\bar e_1$ and $\bar{\mathfrak s}$ are suitable functions in normal form. Since the first non-normal terms in this formula appear from the fourth order on, the computation of the sixth orders Birkhoff invariants for the three-body case is quickly done: it takes less than two pages (see Section 4.1) and gives (1.18).

2. AN ASYMPTOTIC FORMULA FOR THE SECULAR PERTURBATION

Let, for fixed $1 \leqslant i < j \leqslant n$,

$$
f_{ij}(\Lambda,\bar z) := \frac{1}{(2\pi)^2}\int_{\mathbb T^2}\frac{d\lambda_i d\lambda_j}{|x^{(i)}(\Lambda,\lambda_i,\bar z)-x^{(j)}(\Lambda,\lambda_j,\bar z)|}
$$

so as to write

$$
(f_{\rm rps})_{\rm av}(\Lambda,\bar z) = -\sum_{1\leqslant i<j\leqslant n}\bar m_i\bar m_j f_{ij}(\Lambda,\bar z). \tag{2.1}
$$

Here[36] $(\Lambda,\lambda_i,z)\to x^{(i)}(\Lambda,\lambda_i,\bar z)$ denotes the $x^{(i)}$-projection of the map

$$
\phi_{\rm rps}^{-1}:\quad (\Lambda,\lambda,z)\to(y,x)\in\mathbb R^{3n}\times\mathbb R^{3n}. \tag{2.2}
$$

[36] Actually, the map (2.2) depends on z, rather than $\bar z$. However, by the independence of the Hamiltonian (1.14) of (p_n,q_n), we may arbitrarily fix such couple of variables to some value, *e.g.*, $(0,0)$. Abusively, *just in* (2.4) and similar formulae below, we denote again as $(\Lambda,\lambda,\bar z)\to(y(\Lambda,\lambda,\bar z),x(\Lambda,\lambda,\bar z))$ the map $\phi_{\rm rps}^{-1}|_{(p_n,q_n)=(0,0)}$.

Consider the formal expansions

$$f_{ij} = f_{ij}^{(0)} + f_{ij}^{(2)} + \cdots \tag{2.3}$$

in powers of the semi-major axes ratio $\alpha_{ij} := a_i/a_j$, with a_j fixed. Here,

$$f_{ij}^{(k)} := \frac{1}{k!}\frac{d^k}{d\varepsilon^k}\left[\frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}\frac{d\lambda_i d\lambda_j}{|\varepsilon x^{(i)}(\Lambda,\lambda_i,\bar z) - x^{(j)}(\Lambda,\lambda_j,\bar z)|}\right]_{\varepsilon=0}.$$

In particular, we focus on the second-order term of this expansion, given by

$$f_{ij}^{(2)} = \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2} d\lambda_i d\lambda_j \frac{3(x^{(i)}(\Lambda,\lambda_i,\bar z)\cdot x^{(j)}(\Lambda,\lambda_j,\bar z))^2 - |x^{(i)}(\Lambda,\lambda_i,\bar z)|^2|x^{(j)}(\Lambda,\lambda_j,\bar z)|^2}{2|x^{(j)}(\Lambda,\lambda_j,\bar z)|^5}. \tag{2.4}$$

Note that $(f_{\rm rps})_{\rm av}^{(2)}$ in (1.20) corresponds to

$$(f_{\rm rps})_{\rm av}^{(2)} = -\sum_{1\leqslant i<j\leqslant n}\bar m_i \bar m_j f_{ij}^{(2)}. \tag{2.5}$$

Let $\mathrm{C}^{(i)}(\Lambda,\bar z) := x^{(i)}(\Lambda,\lambda_i,\bar z)\times y^{(i)}(\Lambda,\lambda_i,\bar z)$ (by definition of the map (2.2), $\mathrm{C}^{(i)}(\Lambda,\bar z)$ is independent of λ_i). We have the following identity

Proposition 1.

$$f_{ij}^{(2)} = -\frac{M_j m_j^2}{4}\frac{\frac{1}{2\pi}\int_{\mathbb{T}}\left(3(\mathrm{C}^{(j)}\cdot x^{(i)})^2 - |x^{(i)}|^2|\mathrm{C}^{(j)}|^2\right)d\lambda_i}{|\mathrm{C}^{(j)}|^4}\left(\frac{1}{2\pi}\int_{\mathbb{T}}\frac{d\lambda_j}{|x^{(j)}|^2}\right). \tag{2.6}$$

Note that Eqs. (2.5), (2.6) and the formulae of $|\mathrm{C}^{(\mathrm{j})}|$, $|x^{(j)}|$ in terms of RPS variables (see [16, 38] and eventually Appendix B) imply (1.21)–(1.22).

We first discuss

2.1. The Three-body Case

Let

$$\mathrm{P}^{(2)} := \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2} d\lambda_1 d\lambda_2 \frac{3(x^{(1)}(\Lambda,\lambda_1,\mathrm{z})\cdot x^{(2)}(\Lambda,\lambda_2,\mathrm{z}))^2 - |x^{(1)}(\Lambda,\lambda_1,\mathrm{z})|^2|x^{(j)}(\Lambda,\lambda_2,\mathrm{z})|^2}{2|x^{(2)}(\Lambda,\lambda_2,\mathrm{z})|^5},$$

where, for $i = 1, 2$,

$$(\Lambda_1,\Lambda_2,\lambda_i,\mathrm{z})\in\mathcal{A}^2\times\mathbb{T}^1\times B^8 \to (y^{(i)}(\Lambda_1,\Lambda_2,\lambda_i,\mathrm{z}), x^{(i)}(\Lambda_1,\Lambda_2,\lambda_i,\mathrm{z}))\in\mathbb{R}^3\times\mathbb{R}^3$$

are two mappings such that

(A) The map $(\Lambda_1,\Lambda_2,\lambda_2,\mathrm{z})\to(y^{(2)}(\Lambda_1,\Lambda_2,\lambda_2,\mathrm{z}), x^{(2)}(\Lambda_1,\Lambda_2,\lambda_2,\mathrm{z}))$ verifies

$$\frac{|y^{(2)}(\Lambda_1,\Lambda_2,\lambda_2,\mathrm{z})|^2}{2m_2} - \frac{m_2 M_2}{|x^{(2)}(\Lambda_1,\Lambda_2,\lambda_2,\mathrm{z})|} = -\frac{m_2^3 M_2^2}{2\Lambda_2^2}\ ; \tag{2.7}$$

(B) The map

$$\bar\phi: \qquad (\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,\mathrm{z})\to(y^{(1)},y^{(2)},x^{(1)},x^{(2)}) \tag{2.8}$$

is symplectomorphism of $\mathcal{A}^2\times\mathbb{T}^2\times B^8$ into $\mathbb{R}^{12}$ (where $\mathcal{A}^2\subset\mathbb{R}^2$, $B^8\subset\mathbb{R}^8$ are open and connected).

Proposition 2. *Under assumptions* **(A)** *and* **(B)***, the following identity holds*

$$\mathrm{P}^{(2)} = -\frac{M_2 m_2^2}{4} \frac{\frac{1}{2\pi}\int_{\mathbb{T}}\big(3(\mathrm{C}^{(2)}\cdot x^{(1)})^2 - |x^{(1)}|^2|\mathrm{C}^{(2)}|^2\big)d\lambda_1}{|\mathrm{C}^{(2)}|^4}\left(\frac{1}{2\pi}\int_{\mathbb{T}}\frac{d\lambda_2}{|x^{(2)}|^2}\right) \tag{2.9}$$

where $\mathrm{C}^{(2)}(\Lambda_1, \Lambda_2, \mathrm{z}) := x^{(2)}(\Lambda_1, \Lambda_2, \lambda_2, \mathrm{z}) \times y^{(2)}(\Lambda_1, \Lambda_2, \lambda_2, \mathrm{z})$.

Remark 1.

– Note that, in the case $n = 2$, the map (2.2) satisfies assumptions (A) and (B), hence Proposition 2 is just Proposition 1 in this particular case.

– We shall prove more than (2.9): letting $\mathrm{P}^{(1)}(\Lambda, \lambda_1, \mathrm{z})$ as in (2.18) below, then $\mathrm{P}^{(1)}$ satisfies an analogue identity as in (2.9), but neglecting the first average $\frac{1}{2\pi}\int_{\mathbb{T}} d\lambda_1$ (compare the last sentences in the proof of Proposition 2).

– The formula (2.9) resembles the expression of the averaged quartic term in the spin-orbit problem, using Andoyer–Deprit coordinates: see [6, Eq. (24)], in turn based on the expansions in [11, Section 12].

In the next sections, we prove Proposition 2. Next (in Section 2.5), we discuss the general case.

2.2. A Six-degrees of Freedom Set of Symplectic Variables

The proof of Proposition 2 is based on the use of a "ad hoc" variables for the three-body problem. Let us introduce them.

Let $(k^{(1)}, k^{(2)}, k^{(3)})$ be a prefixed orthonormal frame in $\mathbb{R}^3$ and let

$$(y^{(1)}, y^{(2)}, x^{(1)}, x^{(2)}) \in (\mathbb{R}^3)^4, \quad (y^{(i)}, x^{(i)}) = (y_1^{(i)}, y_2^{(i)}, y_3^{(i)}, x_1^{(i)}, x_2^{(i)}, x_3^{(i)})$$

be a system of "Cartesian coordinates" in the configuration space $\mathbb{R}^3$, with respect to $(k^{(1)}, k^{(2)}, k^{(3)})$.

Denote as

$$\mathrm{C}^{(i)} := x^{(i)} \times y^{(i)}$$

(with "$\times$" denoting skew product) the i^{th} angular momentum, and let $\mathrm{C} := \mathrm{C}^{(1)} + \mathrm{C}^{(2)}$ the total angular momentum. 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"). Define the "nodes"

$$\nu_1 := k^{(3)} \times \mathrm{C}, \quad \nu_2 := \mathrm{C} \times x^{(1)}, \quad \nu_3 := x^{(1)} \times \mathrm{C}^{(2)}.$$

Let $\mathcal{P}^{12}_\star$ denote the subset of $(\mathbb{R}^3)^4$ where C, C_2, $x^{(1)}$, $x^{(2)}$, ν_1, ν_2 and ν_3 simultaneously do not vanish. On $\mathcal{P}^{12}_\star$ define a map

$$\phi^{-1}: \quad (y^{(1)}, y^{(2)}, x^{(1)}, x^{(2)}) \to (\mathrm{C}_3, \mathrm{G}, \mathrm{R}_1, \Theta, \mathrm{R}_2, \Phi_2, \zeta, \mathfrak{g}, \mathrm{r}_1, \vartheta, \mathrm{r}_2, \varphi_2)$$

via the following formulae

$$\phi^{-1}: \qquad \begin{cases} \mathrm{C}_3 := \mathrm{C}\cdot k^{(3)} \\ \mathrm{G} := |\mathrm{C}| \\ \mathrm{R}_1 := \dfrac{y^{(1)}\cdot x^{(1)}}{|x^{(1)}|} \\ \Theta := \dfrac{\mathrm{C}^{(2)}\cdot x^{(1)}}{|x^{(1)}|} \\ \mathrm{R}_2 := \dfrac{y^{(2)}\cdot x^{(2)}}{|x^{(2)}|} \\ \Phi_2 := |\mathrm{C}^{(2)}| \end{cases} \qquad \begin{cases} \zeta := \alpha_{k^{(3)}}(k^{(1)}, \nu_1) \\ \mathfrak{g} := \alpha_{\mathrm{C}}(\nu_1, \nu_2) \\ \mathrm{r}_1 := |x^{(1)}| \\ \vartheta := \alpha_{x^{(1)}}(\nu_2, \nu_3) \\ \mathrm{r}_2 := |x^{(2)}| \\ \varphi_2 := \alpha_{\mathrm{C}_2}(\nu_3, x^{(2)}) \end{cases} \tag{2.10}$$

Note that the variables (2.10) provide a reduction of the angular momentum which is *regular* for planar motions (when $\mathrm{C}^{(1)} \parallel \mathrm{C}^{(2)} \parallel \mathrm{C}$, $\Theta = 0$ and $\vartheta = \pi$).

Proposition 3. *The map ϕ^{-1} in (2.10) is invertible on $\mathcal{P}^{12}_*$ and preserves the standard Liouville 1-form $\lambda = \sum_{i=1}^6 \mathrm{P}_i d\mathrm{Q}_i$.*

We denote as

$$\mathcal{R}_1(i) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos i & -\sin i \\ 0 & \sin i & \cos i \end{pmatrix}, \qquad \mathcal{R}_3(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The invertibility is proven by exhibiting the inverse ϕ. Indeed, the definitions in (2.10) and elementary geometric considerations easily imply the following

Lemma 1. *On $\phi^{-1}(\mathcal{P}^{12}_*)$, the inverse map of ϕ^{-1} in (2.10) has the following analytical expression:*

$$\phi : \begin{cases} x^{(1)} = \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\begin{pmatrix} 0 \\ 0 \\ \mathrm{r}_1 \end{pmatrix} \\ y^{(1)} := \dfrac{\mathrm{R}_1}{\mathrm{r}_1} x^{(1)} + \dfrac{1}{\mathrm{r}_1^2}\mathrm{C}^{(1)} \times x^{(1)} \\ x^{(2)} = \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\mathcal{R}_3(\vartheta)\mathcal{R}_1(i_2)\begin{pmatrix} \mathrm{r}_2\cos\varphi_2 \\ \mathrm{r}_2\sin\varphi_2 \\ 0 \end{pmatrix} \\ y^{(2)} = \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\mathcal{R}_3(\vartheta)\mathcal{R}_1(i_2)\begin{pmatrix} \mathrm{R}_2\cos\varphi_2 - \frac{\Phi_2}{\mathrm{r}_2}\sin\varphi_2 \\ \mathrm{R}_2\sin\varphi_2 + \frac{\Phi_2}{\mathrm{r}_2}\cos\varphi_2 \\ 0 \end{pmatrix} \end{cases} \tag{2.11}$$

where, if i, i_1, $i_2 \in (0, \pi)$ are defined by

$$\cos i = \frac{\mathrm{C}_3}{\mathrm{G}}, \quad \cos i_1 = \frac{\Theta}{\mathrm{G}}, \quad \cos i_2 = \frac{\Theta}{\Phi_2} \tag{2.12}$$

and C, $\mathrm{C}^{(2)}$ *by*

$$\begin{aligned} \mathrm{C} &:= \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\begin{pmatrix} 0 \\ 0 \\ \mathrm{G} \end{pmatrix} \\ \mathrm{C}^{(2)} &:= \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\mathcal{R}_3(\vartheta)\mathcal{R}_1(i_2)\begin{pmatrix} 0 \\ 0 \\ \Phi_2 \end{pmatrix} \end{aligned} \tag{2.13}$$

then

$$\mathrm{C}^{(1)} := \mathrm{C} - \mathrm{C}^{(2)}. \tag{2.14}$$

To prove symplecticity we shall use the following easy

Lemma 2 ([13]). *Let*

$$x = \mathcal{R}_3(\theta)\mathcal{R}_1(i)\bar{x}, \quad y = \mathcal{R}_3(\theta)\mathcal{R}_1(i)\bar{y}, \quad \mathrm{C} := x \times y, \quad \bar{\mathrm{C}} := \bar{x} \times \bar{y},$$

with $x, \bar{x}, y, \bar{y} \in \mathbb{R}^3$. *Then,*

$$y \cdot dx = \mathrm{C} \cdot k^{(3)} d\theta + \bar{\mathrm{C}} \cdot k^{(1)} di + \bar{y} \cdot d\bar{x},$$

with $k^{(1)} := (1, 0, 0)$, $k^{(3)} := (0, 0, 1)$.

Proof of Proposition 3. Let us preliminarily verify that, if $\mathrm{C}^{(i)}$ are as in (2.13)–(2.14), and $y^{(i)}$, $x^{(i)}$ as in (2.11), then as expected,

$$x^{(i)} \times y^{(i)} = \mathrm{C}^{(i)}. \tag{2.15}$$

Indeed, for $i = 2$, this identity follows trivially from the definitions. To check that it holds also for $i = 1$, one can do as follows: firstly, to check that $x^{(1)} \cdot \mathrm{C}^{(1)} = 0$. This is an elementary consequence of (2.11) and, in particular, of (2.12). Next, using the rule of the double skew product, one has

$$\begin{aligned} x^{(1)} \times y^{(1)} &= x^{(1)} \times \left(\frac{\mathrm{R}_1}{\mathrm{r}_1} x^{(1)} + \frac{1}{\mathrm{r}_1^2} \mathrm{C}^{(1)} \times x^{(1)} \right) \\ &= 0 + \frac{1}{\mathrm{r}_1^2} \left(\mathrm{r}_1^2\, \mathrm{C}^{(1)} - (x^{(1)} \cdot \mathrm{C}^{(1)})\, x^{(1)} \right) = \mathrm{C}^{(1)}. \end{aligned}$$

Define now

$$\begin{aligned} \bar{\mathrm{C}}^{(1)} &:= \mathcal{R}_1(-i)\mathcal{R}_3(-\zeta)\mathrm{C}^{(1)} \\ \bar{\mathrm{C}}^{(2)} &:= \mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\mathcal{R}_3(\vartheta)\mathcal{R}_1(i_2) \begin{pmatrix} 0 \\ 0 \\ \Phi_2 \end{pmatrix} \\ \bar{\bar{\mathrm{C}}}^{(1)} &:= \mathcal{R}_1(-i_1)\mathcal{R}_3(-\mathfrak{g}) \begin{pmatrix} 0 \\ 0 \\ \mathrm{G} \end{pmatrix} - \mathcal{R}_3(\vartheta)\mathcal{R}_1(i_2) \begin{pmatrix} 0 \\ 0 \\ \Phi_2 \end{pmatrix} \\ \bar{\bar{\mathrm{C}}}^{(2)} &:= \mathcal{R}_3(\vartheta)\mathcal{R}_1(i_2) \begin{pmatrix} 0 \\ 0 \\ \Phi_2 \end{pmatrix} \\ \bar{\bar{\bar{\mathrm{C}}}}^{(2)} &:= \begin{pmatrix} 0 \\ 0 \\ \Phi_2 \end{pmatrix} \end{aligned}$$

and

$$\bar{\bar{y}}^{(1)} := \begin{pmatrix} 0 \\ 0 \\ \mathrm{R}_1 \end{pmatrix} + \frac{1}{\mathrm{r}_1^2} \bar{\bar{\mathrm{C}}}^{(1)} \times \begin{pmatrix} 0 \\ 0 \\ \mathrm{r}_1 \end{pmatrix}, \quad \bar{\bar{\mathrm{x}}}^{(1)} := \begin{pmatrix} 0 \\ 0 \\ \mathrm{r}_1 \end{pmatrix}$$

$$\bar{\bar{\bar{x}}}^{(2)} := \begin{pmatrix} \mathrm{r}_2 \cos\varphi_2 \\ \mathrm{r}_2 \sin\varphi_2 \\ 0 \end{pmatrix}, \qquad \bar{\bar{\bar{y}}}^{(2)} := \begin{pmatrix} \mathrm{R}_2 \cos\varphi_2 - \frac{\Phi_2}{\mathrm{r}_2} \sin\varphi_2 \\ \mathrm{R}_2 \sin\varphi_2 + \frac{\Phi_2}{\mathrm{r}_2} \cos\varphi_2 \\ 0 \end{pmatrix}$$

so as to write

$$\begin{aligned}
y^{(1)} &= \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\bar{\bar{y}}^{(1)}\\
x^{(1)} &= \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\bar{\bar{x}}^{(1)}\\
x^{(2)} &= \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\mathcal{R}_3(\vartheta)\mathcal{R}_1(i_2)\bar{\bar{\bar{x}}}^{(2)}\\
y^{(2)} &= \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)\mathcal{R}_3(\vartheta)\mathcal{R}_1(i_2)\bar{\bar{\bar{y}}}^{(2)}.
\end{aligned}$$

Applying repeatedly Lemma 2, Eq. (2.15) and the rule

$$\mathcal{R}x \times \mathcal{R}y = \mathcal{R}(x \times y) \qquad \text{for all} \quad \mathcal{R} \in \mathrm{SO}(3),\ x,\ y \in \mathbb{R}^3$$

gives

$$\begin{aligned}
y^{(1)} \cdot dx^{(1)} &= \mathrm{C}^{(1)} \cdot k^{(3)} d\zeta + \bar{\mathrm{C}}^{(1)} \cdot k^{(1)} di + \bar{\mathrm{C}}^{(1)} \cdot k^{(3)} d\mathfrak{g} + \bar{\bar{\mathrm{C}}}^{(1)} \cdot \mathrm{k}^{(1)} \mathrm{di}_1 + \mathrm{R}_1 \mathrm{dr}_1\\
y^{(2)} \cdot dx^{(2)} &= \mathrm{C}^{(2)} \cdot k^{(3)} d\zeta + \bar{\mathrm{C}}^{(2)} \cdot k^{(1)} di + \bar{\mathrm{C}}^{(2)} \cdot k^{(3)} d\mathfrak{g} + \bar{\bar{\mathrm{C}}}^{(2)} \cdot \mathrm{k}^{(1)} \mathrm{di}_1 + \bar{\bar{\mathrm{C}}}^{(2)} \cdot \mathrm{k}^{(3)} \mathrm{d}\vartheta\\
&\quad + \bar{\bar{\bar{\mathrm{C}}}}^{(2)} \cdot \mathrm{k}^{(1)} \mathrm{di}_2 + \mathrm{R}_2 \mathrm{dr}_2 + \Phi_2 \mathrm{d}\varphi_2.
\end{aligned}$$

Taking the sum of the two equations and recognizing that, if

$$e^{(i)} := \mathcal{R}_3(\zeta)\mathcal{R}_1(i)k^{(i)}, \quad f^{(i)} := \mathcal{R}_3(\zeta)\mathcal{R}_1(i)\mathcal{R}_3(\mathfrak{g})\mathcal{R}_1(i_1)k^{(i)}$$

then

$$\begin{aligned}
&(\mathrm{C}^{(1)} + \mathrm{C}^{(2)}) \cdot k^{(3)} = \mathrm{C} \cdot k^{(3)} = \mathrm{G}\cos i = \mathrm{C}_3\\
&(\bar{\mathrm{C}}^{(1)} + \bar{\mathrm{C}}^{(2)}) \cdot k^{(1)} = \mathrm{C} \cdot e^{(1)} = 0\\
&(\bar{\mathrm{C}}^{(1)} + \bar{\mathrm{C}}^{(2)}) \cdot k^{(3)} = \mathrm{C} \cdot e^{(3)} = \mathrm{G}\\
&(\bar{\bar{\mathrm{C}}}^{(1)} + \bar{\bar{\mathrm{C}}}^{(2)}) \cdot \mathrm{k}^{(1)} = \mathrm{C} \cdot \mathrm{f}^{(1)} = (\mathrm{G}\mathrm{k}^{(3)}) \cdot (\mathcal{R}_3(\mathfrak{g})\mathrm{k}^{(1)}) = 0\\
&\bar{\bar{\mathrm{C}}}^{(2)} \cdot \mathrm{k}^{(3)} = \Phi_2 \cos \mathrm{i}_2 = \Theta\\
&\bar{\bar{\bar{\mathrm{C}}}}^{(2)} \cdot \mathrm{k}^{(1)} = 0
\end{aligned}$$

we have the thesis:

$$y^{(1)} \cdot dx^{(1)} + y^{(2)} \cdot dx^{(2)} = \mathrm{C}_3 d\zeta + \mathrm{G} d\mathfrak{g} + \Theta d\vartheta + \mathrm{R}_1 d\mathrm{r}_1 + \mathrm{R}_2 d\mathrm{r}_2 + \Phi_2 d\varphi_2.$$

□

2.3. Two-steps Averaging for Properly-degenerate Systems

In this section we discuss a uniqueness argument for normal forms of degenerate systems.

Consider a real-analytic and properly-degenerate Hamiltonian

$$\mathrm{H}(I, \varphi, u, v) = \mathrm{H}_0(I) + \alpha \mathrm{P}(I, \varphi, u, v), \qquad 0 < \alpha < 1$$

defined on some phase $2(n_1 + n_2)$-dimensional phase space of the form $V \times \mathbb{T}^{n_1} \times B^{2n_2}$, where V, B^{2n_2} are an open, connected sets of $\mathbb{R}^{n_1}$, $\mathbb{R}^{2n_2}$. Perturbation theory (*e.g.*, [5, 7, 12, 33, 39]) tells us that, under suitable assumptions of non resonance of the unperturbed frequency map $\omega := \partial_I \mathrm{H}_0$ and of smallness of the perturbation $\alpha\mathrm{P}$, the system may be conjugated, at least formally, to a new system

$$\mathrm{H}_p(I, \varphi, u, v) = \mathrm{H}_0(I) + (\alpha \bar{\mathrm{P}}_1(I, u, v) + \cdots + \alpha^p \bar{\mathrm{P}}_p) + \alpha^{p+1} \mathrm{P}_{p+1}, \qquad (\mathrm{P}_1 \equiv \mathrm{P}) \tag{2.16}$$

where the term inside parentheses ("p-step normal form") is of degree p and is *independent of* φ. Quantitative versions of this fact are well known in the literature since [5] and have been more and more refining themselves (depending on needs) both in the non-degenerate [11, 39] and degenerate case [5, 7, 33, 35]. Moreover, we know that, when the system in non-degenerate, i.e., the variables (u, v) do not appear, the *p-step normal form* is uniquely determined (though the change of variables realizing it may be not). In general, when the system is degenerate, uniqueness does not hold. However, the following lemma is easily proved.

Lemma 3. *Let*[37] $n_1 = 1$ *and* H *be a properly-degenerate system, such that*

$$\mathrm{P}_{\rm av} := \frac{1}{2\pi}\int_{\mathbb{T}} \mathrm{P}(I, \varphi; u, v) d\varphi \equiv 0. \tag{2.17}$$

Then, the two-step normal form

$$\tilde{\mathrm{H}}(\tilde{I}, \tilde{\varphi}; \tilde{u}, \tilde{v}) = \mathrm{H}_0(\tilde{I}) + (\alpha\bar{\mathrm{P}}_1(\tilde{I}; \tilde{u}, \tilde{v}) + \alpha^2\bar{\mathrm{P}}_2(\tilde{I}; \tilde{u}, \tilde{v})) + \mathrm{O}(\alpha^3)$$

is uniquely determined, up to real-analytic and symplectic changes $(\tilde{I}, \tilde{\varphi}; \tilde{u}, \tilde{v}) \in \tilde{V} \times \mathbb{T} \times \tilde{B}^{2n_2} \to (I, \varphi; u, v) \in V \times \mathbb{T} \times B^{2n_2}$*, α-close to the identity.*

Proof. Let $p \geqslant 0$. Assuming to have reached the form in (2.16) (with the term inside parentheses identically vanishing for $p = 0$), the $(p+1)^{\rm th}$ Hamiltonian H_{p+1} is obtained applying to H_p any transformation in the class of *infinitesimal transformations* having as α^{p+1} germ the time-one flow of $\alpha^{p+1}\psi_{p+1}$, where

$$\psi_{p+1} := \sum_{k\neq 0} \frac{\mathrm{P}_k^{(p+1)}(I; u, v)}{\mathrm{i}k \cdot \omega(I)} e^{\mathrm{i}k\cdot\varphi} + \bar{\psi}_p$$

if P_{p+1} has the Fourier expansion

$$\mathrm{P}_{p+1} = \sum_{k\neq 0} \mathrm{P}_k^{(p+1)}(I; u, v) e^{\mathrm{i}k\cdot\varphi}$$

and $\bar{\psi}_p$ is any function independent of φ. Moreover, as it is known, $\bar{\mathrm{P}}_j$'s and P_j's are related by

$$\bar{\mathrm{P}}_{p+1} = (\mathrm{P}_{p+1})_{\rm av} = \frac{1}{2\pi}\int_{\mathbb{T}} \mathrm{P}_{p+1} d\varphi.$$

Therefore, if we perform two steps of the procedure, i.e. , with $p = 0$, 1, we find the *two-step normal form* is defined by $\bar{\mathrm{P}}_1 = \mathrm{P}_{\rm av} = 0$ and

$$\bar{\mathrm{P}}_2 = \frac{1}{2}\frac{1}{2\pi}\int_{\mathbb{T}} \{\psi_1, \mathrm{P}\} d\varphi,$$

where $\{\cdot, \cdot\}$ denotes Poisson parentheses with respect to all the variables. (The relative transformation will be given by $\phi_1 \circ \phi_2$, where ϕ_j is generated by $\alpha^j\psi_j$.) Therefore, to prove uniqueness, all we have to do is to check that, if we change $\psi_1 \to \psi_1 + \tilde{\psi}_1$, where $\tilde{\psi}_1$ is independent of φ, the function $\bar{\mathrm{P}}_2$ does not change. And in fact this term changes by adding

$$\frac{1}{2}\frac{1}{2\pi}\int_{\mathbb{T}} \{\tilde{\psi}_1, \mathrm{P}\} d\varphi.$$

Since $\tilde{\psi}_1$ is independent of φ, Poisson parentheses and the integral may be exchanged and we see that this term vanishes

$$\frac{1}{2}\frac{1}{2\pi}\int_{\mathbb{T}} \{\tilde{\psi}_1, \mathrm{P}\} d\varphi = \frac{1}{2}\frac{1}{2\pi}\Big\{\tilde{\psi}_1, \int_{\mathbb{T}} \mathrm{P} d\varphi\Big\} = \Big\{\tilde{\psi}_1, \frac{1}{2}\mathrm{P}_{\rm av}\Big\} = 0$$

because of (2.17). □

[37] We assume $n_1 = 1$ to avoid complications due to resonances of the frequency-map. This is enough for the purposes of the paper. Analogue statements for the case $n_1 \geqslant 1$ may be available.

2.4. Proof of Proposition 2

To prove Proposition 2, we write $\mathrm{P}^{(2)}$ as

$$\mathrm{P}^{(2)} = \frac{1}{2\pi}\int_{\mathbb{T}} \mathrm{P}^{(1)}(\Lambda, \lambda_1, \mathrm{z})d\lambda_1 \tag{2.18}$$

where

$$\mathrm{P}^{(1)}(\Lambda, \lambda_1, \mathrm{z}) := \frac{1}{2\pi}\int_{\mathbb{T}} d\lambda_2 \frac{3(x^{(1)}(\Lambda, \lambda_1, \mathrm{z}) \cdot x^{(2)}(\Lambda, \lambda_2, \mathrm{z}))^2 - |x^{(1)}(\Lambda, \lambda_1, \mathrm{z})|^2|x^{(2)}(\Lambda, \lambda_1, \mathrm{z})|^2}{2|x^{(2)}(\Lambda, \lambda_2, \mathrm{z})|^5}. \tag{2.19}$$

Then we consider the auxiliary Hamiltonian

$$\mathrm{H}_{\mathrm{Dip}}(y^{(1)}, x^{(1)}, y^{(2)}, x^{(2)}) := \frac{|y^{(2)}|^2}{2m_2} - \frac{m_2 M_2}{|x^{(2)}|} - \alpha m_2 M_2 \frac{x^{(1)} \cdot x^{(2)}}{|x^{(2)}|^3}$$

on the phase space

$$\{(y^{(1)}, x^{(1)}, y^{(2)}, x^{(2)}) \in (\mathbb{R}^3)^4 : \ x^{(2)} \neq 0\}$$

endowed with the standard symplectic form

$$\omega := dy^{(1)} \wedge dx^{(1)} + dy^{(2)} \wedge dx^{(2)}$$

and $\alpha \ll 1$ a small positive parameter.

By assumption (A), in the variables $(\Lambda, \lambda, \mathrm{z})$ in (2.8), $\mathrm{H}_{\mathrm{Dip}}$ takes the form

$$\begin{aligned} \mathrm{H}(\Lambda_1, \Lambda_2, \lambda_1, \lambda_2, \mathrm{z}) &= \mathrm{H}_{\mathrm{Kep}}(\Lambda_2) + \alpha \mathrm{P}(\Lambda_1, \Lambda_2, \lambda_1, \lambda_2, \mathrm{z}) \\ &= -\frac{M_2^2 m_2^3}{2\Lambda_2^3} - \alpha M_2 m_2 \frac{x^{(1)}(\Lambda, \lambda_1, \mathrm{z}) \cdot x^{(2)}(\Lambda, \lambda_2, \mathrm{z})}{|x^{(2)}(\Lambda, \lambda_2, \mathrm{z})|^3}. \end{aligned} \tag{2.20}$$

Lemma 4. *Under assumptions of Proposition 2, the Hamiltonian in* (2.20)*, endowed with the symplectic form*

$$\sum_{i=1}^{2} d\Lambda_i \wedge d\lambda_i + \sum_{i=1}^{4} du_i \wedge dv_i, \qquad \mathrm{z} = (u, v)$$

verifies the assumptions of Lemma 3, with the "variables" $(I, \varphi) := (\Lambda_2, \lambda_2)$ and the "parameters" $(\Lambda_1, \lambda_1, \mathrm{z})$. Its (unique) two-step normal form is

$$\tilde{\mathrm{H}}(\Lambda_1, \Lambda_2, \lambda_1, \mathrm{z}) = \mathrm{H}_{\mathrm{Kep}}(\Lambda_2) + \alpha^2 M_2 m_2 \mathrm{P}^{(1)}(\Lambda_1, \Lambda_2, \lambda_1, \mathrm{z}) + \mathrm{O}(\alpha^3)$$

with $\mathrm{P}^{(1)}$ as in (2.19).

Proof. Consider the auxiliary Hamiltonian

$$\mathrm{H}^\star(\Lambda_1, \Lambda_2, \lambda_1, \lambda_2, \mathrm{z}) = \mathrm{H}(\Lambda_1, \Lambda_2, \lambda_1, \lambda_2, \mathrm{z}) + \alpha^2 M_2 m_2 \mathrm{Q}(\Lambda_1, \Lambda_2, \lambda_1, \lambda_2, \mathrm{z}), \tag{2.21}$$

where H is as in (2.20) and

$$\begin{aligned} &\mathrm{Q}(\Lambda_1, \Lambda_2, \lambda_1, \lambda_2, \mathrm{z}) := \\ &\quad -\frac{3(x^{(1)}(\Lambda_1, \Lambda_2, \lambda_1, \mathrm{z}) \cdot x^{(2)}(\Lambda_1, \Lambda_2, \lambda_2, \mathrm{z}))^2 - |x^{(1)}(\Lambda_1, \Lambda_2, \lambda_1, \mathrm{z})|^2|x^{(2)}(\Lambda_1, \Lambda_2, \lambda_1, \mathrm{z})|^2}{2|x^{(2)}(\Lambda_1, \Lambda_2, \lambda_2, \mathrm{z})|^5}. \end{aligned}$$

Let us apply Lemma 3 to H, with (I, φ) corresponding to (Λ_2, λ_2) and (u, v) to $(\Lambda_1, \lambda_1, \mathrm{z})$. The assumption (2.7) implies that the zero-averaging (with respect to λ_2) assumption for P is satisfied:

$$\begin{aligned} \int_{\mathbb{T}} \mathrm{P} d\lambda_2 &= \int_{\mathbb{T}} (-m_2 M_2 \frac{x^{(1)} \cdot x^{(2)}}{|x^{(2)}|^3}) d\lambda_2 = -m_2 M_2 x^{(1)} \cdot \int_{\mathbb{T}} \frac{x^{(2)}}{|x^{(2)}|^3} d\lambda_2 \\ &= m_2 M_2 \omega_{\mathrm{Kep}}^{(2)} x^{(1)} \cdot \int_{\mathbb{T}} \partial_{\lambda_2} y^{(2)} = 0. \end{aligned}$$

Denote as

$$\tilde{\mathrm{H}}(\tilde{\Lambda}_1,\tilde{\Lambda}_2,\tilde{\lambda}_1,\tilde{\mathrm{z}}) := \mathrm{H}\circ\psi(\tilde{\Lambda}_1,\tilde{\Lambda}_2,\tilde{\lambda}_1,\tilde{\mathrm{z}}) = \mathrm{H}_{\mathrm{Kep}}(\tilde{\Lambda}_2) + \alpha^2 M_2 m_2 \mathrm{P}^{(1)}(\tilde{\Lambda}_1,\tilde{\Lambda}_2,\tilde{\lambda}_1,\tilde{\mathrm{z}}) + \mathrm{O}(\alpha^3)$$

the two-step normal form which H is put in via Lemma 3, where denotes the symplectic, α-close-to-the identity transformation realizing it. Then, since ψ is α-close to the identity, $\mathrm{H}^\star$ is transformed into

$$\tilde{\mathrm{H}}^\star = \tilde{\mathrm{H}}(\tilde{\Lambda}_1,\tilde{\Lambda}_2,\tilde{\lambda}_1,\tilde{\mathrm{z}}) + \alpha^2 M_2 m_2 \mathrm{Q}(\tilde{\Lambda}_1,\tilde{\Lambda}_2,\tilde{\lambda}_1,\tilde{\lambda}_2,\tilde{\mathrm{z}}) + \mathrm{O}(\alpha^3).$$

Hence, at expenses of a further $\tilde{\lambda}_2$-averaging (α^2-close to the identity), $\tilde{\mathrm{H}}^\star$ is transformed into

$$\begin{aligned}\hat{\mathrm{H}}(\hat{\Lambda}_1,\hat{\Lambda}_2,\hat{\lambda}_1,\hat{\mathrm{z}}) &= \tilde{\mathrm{H}}(\hat{\Lambda}_1,\hat{\Lambda}_2,\hat{\lambda}_1,\hat{\mathrm{z}}) + \alpha^2 M_2 m_2 \mathrm{Q}^{(1)}(\hat{\Lambda}_1,\hat{\Lambda}_2,\hat{\lambda}_1,\hat{\mathrm{z}}) + \mathrm{O}(\alpha^3)\\ &= \mathrm{H}_{\mathrm{Kep}}(\hat{\Lambda}_2) + \alpha^2 M_2 m_2 \mathrm{P}^{(1)}(\hat{\Lambda}_1,\hat{\Lambda}_2,\hat{\lambda}_1,\hat{\mathrm{z}}) + \alpha^2 M_2 m_2 \mathrm{Q}^{(1)}(\hat{\Lambda}_1,\hat{\Lambda}_2,\hat{\lambda}_1,\hat{\mathrm{z}})\\ &\quad + \mathrm{O}(\alpha^3),\end{aligned}$$

with

$$\mathrm{Q}^{(1)}(\hat{\Lambda}_1,\hat{\Lambda}_2,\hat{\lambda}_1,\hat{\mathrm{z}}) := \frac{1}{2\pi}\int_{\mathbb{T}} \mathrm{Q}(\hat{\Lambda}_1,\hat{\Lambda}_2,\hat{\lambda}_1,\hat{\lambda}_2,\hat{\mathrm{z}})d\hat{\lambda}_2.$$

On the other hand, $\mathrm{H}^\star$ in (2.21) may be written as

$$\mathrm{H}^\star(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,\mathrm{z}) = \mathrm{H}^\star_{2\mathrm{B}}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,\mathrm{z}) + \mathrm{O}(\alpha^3)$$

with

$$\mathrm{H}^\star_{2\mathrm{B}}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,\mathrm{z}) := \frac{|y^{(2)}(\Lambda_1,\Lambda_2,\lambda_2,\mathrm{z})|^2}{2m_2} - \frac{m_2 M_2}{|x^{(2)}(\Lambda_1,\Lambda_2,\lambda_2,\mathrm{z}) - \alpha x^{(1)}(\Lambda_1,\Lambda_2,\lambda_1,\mathrm{z})|}.$$

Due to assumption (B), we find a real-analytic symplectomorphsm

$$(\check{\Lambda}_1,\check{\Lambda}_2,\check{\lambda}_1,\check{\lambda}_2,\check{\mathrm{z}}) \to (\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,\mathrm{z})$$

α-close to the identity, which conjugates $\mathrm{H}^\star_{2\mathrm{B}}$ to $\mathrm{H}_{\mathrm{Kep}}(\check{\Lambda}_2) = -\dfrac{M_2^2 m_2^3}{2\check{\Lambda}_2^2}$ and hence $\mathrm{H}^\star$ is conjugated to

$$\check{\mathrm{H}}* = -\frac{M_2^2 m_2^3}{2\check{\Lambda}_2^2} + \mathrm{O}(\alpha^3).$$

By comparison with $\hat{\mathrm{H}}$ above, uniqueness claimed by Lemma 3 implies

$$\mathrm{P}^{(1)} + \mathrm{Q}^{(1)} = \mathrm{O}(\alpha^3)$$

which is the thesis. □

We are now ready for the

Proof of Proposition 2. For the purposes of this proof, if $f : x \in \mathbb{T} \to f(x) \in \mathbb{R}$ is continuous, we denote as $\langle f\rangle_x := \frac{1}{2\pi}\int_{\mathbb{T}} f(x)dx$.

Consider the Hamiltonian H in (2.20); let $\bar{\phi}$ as in (2.8) and ϕ as in (2.11). Denote as $\mathrm{H}_{\mathrm{red}} := \mathrm{H}\circ\bar{\phi}^{-1}\circ\phi$ the expression of H in the variables (2.10). This is

$$\mathrm{H}_{\mathrm{red}} = \mathrm{H}\circ\bar{\phi}^{-1}\circ\phi = \frac{\mathrm{R}_2^2}{2m_2} - \frac{M_2 m_2}{\mathrm{r}_2} + \frac{\Phi_2^2}{2m_2\mathrm{r}_2^2} - M_2 m_2 \alpha \frac{\mathrm{r}_1}{\mathrm{r}_2^2}\sqrt{1 - (\frac{\Theta}{\Phi_2})^2}\sin\varphi_2. \tag{2.22}$$

Let us split $\mathrm{H}_{\mathrm{red}}$ into two parts, a "radial" and a "tangential" one:

$$\mathrm{H}_{\mathrm{rad}} := \frac{\mathrm{R}_2^2}{2m_2} - \frac{M_2 m_2}{\mathrm{r}_2}$$

and

$$\mathrm{H}_{\tan} := \frac{\Phi_2^2}{2m_2\mathrm{r}_2^2} - M_2 m_2 \alpha \frac{\mathrm{r}_1}{\mathrm{r}_2^2}\sqrt{1-(\frac{\Theta}{\Phi_2})^2}\sin\varphi_2$$

and focus on $\mathrm{H}_{\tan}$. We shall eliminate the dependence from the angle φ_2 up to order α^3. To this end, define h_0, P_0 via $\mathrm{H}_{\tan} =: \mathrm{h}_0 + \alpha\mathrm{P}_0$ and denote $\varpi := \partial_{\Phi_2}\mathrm{h}_0 = \frac{\Phi_2}{m_2\mathrm{r}_2^2}$. Since $\langle \mathrm{P}_0\rangle_{\varphi_2} = 0$, a Hamiltonian vector field the time-one flow of which eliminates the dependence on φ_2 up to $\mathrm{O}(\alpha^2)$ has as Hamiltonian the function ψ_0 defined as a primitive

$$\psi_0 = \frac{1}{\varpi}\int^{\varphi_2}\alpha\mathrm{P}_0 = M_2 m_2^2\alpha\frac{\mathrm{r}_1}{\Phi_2}\sqrt{1-(\frac{\Theta}{\Phi_2})^2}\cos\varphi_2$$

with $\langle\phi_0\rangle_{\varphi_2} = 0$. It is a remarkable fact that r_2 is cancelled. Since ϕ_0 is also independent of R_1, R_2 and ϑ, this implies that its time-one flow, that we denote

$$\phi_0 : \quad (\tilde{\mathrm{R}}_1, \tilde{\mathrm{R}}_2, \tilde{\Phi}_2, \tilde{\Theta}, \tilde{\mathrm{r}}_2, \tilde{\mathrm{r}}_2, \tilde{\varphi}_2, \tilde{\vartheta}) \to (\mathrm{R}_1, \mathrm{R}_2, \Phi_2, \Theta, \mathrm{r}_2, \mathrm{r}_2, \varphi_2, \vartheta),$$

leaves $(\mathrm{R}_2, \mathrm{r}_2, \Theta, \mathrm{r}_1)$ unvaried. Using again $\langle\mathrm{P}_0\rangle_{\varphi_2} = 0$, we then have that H_0 is conjugated to

$$\mathrm{H}_1 = \mathrm{H}_{\tan}\circ\phi_0 = \mathrm{h}_0 + \alpha^2\mathrm{P}_1 + \mathrm{O}(\alpha^3),$$

where

$$\mathrm{P}_1 = \frac{1}{2}\{\psi_0, \mathrm{P}_0\} = -\frac{M_2^2 m_2^3}{2}\frac{\tilde{\mathrm{r}}_1^2}{\tilde{\mathrm{r}}_2^2\tilde{\Phi}_2^4}\Big(\tilde{\Theta}^2 - \frac{1}{2}(\tilde{\Phi}_2^2 - \tilde{\Theta}^2)(1+\cos 2\tilde{\varphi}_2)\Big).$$

A further step of averaging defined by the time-one flow

$$\phi_1 : \quad (\hat{\mathrm{R}}_1, \hat{\mathrm{R}}_2, \hat{\Phi}_2, \hat{\Theta}, \hat{\mathrm{r}}_2, \hat{\mathrm{r}}_2, \hat{\varphi}_2, \hat{\vartheta}) \to (\tilde{\mathrm{R}}_1, \tilde{\mathrm{R}}_2, \tilde{\Phi}_2, \tilde{\Theta}, \tilde{\mathrm{r}}_2, \tilde{\mathrm{r}}_2, \tilde{\varphi}_2, \tilde{\vartheta})$$

of

$$\begin{aligned}\psi_1 &= \frac{1}{\varpi}\int^{\varphi_2}\alpha^2(\mathrm{P}_1 - \langle\mathrm{P}_1\rangle)\\ &= +\frac{\alpha^2}{\varpi}\int^{\varphi_2}\frac{M_2^2 m_2^3}{4}\frac{\mathrm{r}_1^2}{\mathrm{r}_2^2\Phi_2^4}(\Phi_2^2-\Theta^2)\cos 2\varphi\\ &= +\alpha^2\frac{M_2^2 m_2^4}{8}\frac{\mathrm{r}_1^2}{\Phi_2^5}(\Phi_2^2-\Theta^2)\sin 2\varphi_2\end{aligned}$$

with $\langle\psi_1\rangle_{\varphi_2} = 0$. As in the previous step, ψ_1 is independent of $(\mathrm{R}_1, \mathrm{R}_2, \vartheta)$ and, again r_2, hence, ϕ_1 leaves $(\mathrm{R}_2, \mathrm{r}_2, \Theta, \mathrm{r}_1)$ unvaried. Then H_1 is let into the form

$$\mathrm{H}_2 = \mathrm{H}_1\circ\phi_1 = \mathrm{h}_0 + \alpha^2\mathrm{P}_2 + \mathrm{O}(\alpha^3),$$

where

$$\mathrm{P}_2 = \langle\mathrm{P}_1\rangle_{\varphi_2} = -\frac{M_2^2 m_2^3}{4}\frac{\hat{\mathrm{r}}_1^2}{\hat{\mathrm{r}}_2^2\hat{\Phi}_2^4}(3\hat{\Theta}^2 - \hat{\Phi}_2^2).$$

Including also the term $\mathrm{H}_{\mathrm{rad}}$ (left unvaried by this sequence of transformations) we finally have that the Hamiltonian $\mathrm{H}_{\mathrm{red}}$ in (2.22) is transformed into

$$\hat{\mathrm{H}} := \mathrm{H}_{\mathrm{red}}\circ\phi_0\circ\phi_1 = \frac{\hat{\mathrm{R}}_2^2}{2m_2} - \frac{M_2 m_2}{\hat{\mathrm{r}}_2} + \frac{1}{2m_2\hat{\mathrm{r}}_2^2}\Big(\hat{\Phi}_2^2 - \alpha^2\frac{M_2^2 m_2^4}{2}\frac{\hat{\mathrm{r}}_1^2}{\hat{\Phi}_2^4}(3\hat{\Theta}^2 - \hat{\Phi}_2^2)\Big) + \mathrm{O}(\alpha^3). \quad (2.23)$$

Let now

$$(\hat{y}^{(1)}, \hat{y}^{(2)}, \hat{x}^{(1)}, \hat{x}^{(2)}) \in \mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3$$

be related to $(\mathrm{C}_3, \mathrm{G}, \hat{\mathrm{R}}_1, \hat{\mathrm{R}}_2, \hat{\Phi}_2, \hat{\Theta}, \zeta, \mathfrak{g}, \hat{\mathrm{r}}_1, \hat{\mathrm{r}}_2, \hat{\varphi}_2, \hat{\vartheta})$ via relations analogue to (2.11)-(2.12), i.e., $(\hat{y}, \hat{x}) = \phi^{-1}(\mathrm{C}_3, \mathrm{G}, \hat{\mathrm{R}}_1, \hat{\mathrm{R}}_2, \hat{\Phi}_2, \hat{\Theta}, \zeta, \mathfrak{g}, \hat{\mathrm{r}}_1, \hat{\mathrm{r}}_2, \hat{\varphi}_2, \hat{\vartheta})$, with ϕ^{-1} as in (2.10) and let $(\hat{\Lambda}, \hat{\lambda}, \hat{\mathrm{z}})$ be defined via $(\hat y, \hat x) = \bar\phi(\hat\Lambda, \hat\lambda, \hat z)$, with $\bar\phi$ as in (2.8). In the variables $(\hat\Lambda, \hat\lambda, \hat z)$, the he Hamiltonian (2.23) takes the form

$$\tilde{\mathrm{H}} := \hat{\mathrm{H}} \circ \phi^{-1} \circ \bar\phi = -\frac{M_2^2 m_2^3}{2\hat\Lambda_2^2} - \frac{\alpha^2}{2m_2 \hat{\mathrm{r}}_2^2} \frac{M_2^2 m_2^4}{2} \frac{\hat{\mathrm{r}}_1^2}{\hat\Phi_2^4}(3\hat\Theta^2 - \hat\Phi_2^2),$$

where $\hat\Theta = \hat{\mathrm{C}}^{(2)} \cdot \hat x^{(1)}$, $\hat{\mathrm{r}}_i = |\hat x^{(i)}|$, $\hat\Phi_2 = |\hat{\mathrm{C}}^{(2)}|$, with $\hat{\mathrm{C}}^{(2)} = \hat x^{(2)} \times \hat y^{(2)}$ have to be regarded as functions of $(\Lambda, \hat\lambda, \hat z)$. A further $\hat\lambda_2$-averaging, α^2-close to the identity

$$\widehat\phi: \qquad (\widehat\Lambda, \widehat\lambda, \widehat{\mathrm{z}}) \to (\hat\Lambda, \hat\lambda, \hat{\mathrm{z}})$$

transforms $\tilde{\mathrm{H}}$ into

$$\widehat{\mathrm{H}} := \tilde{\mathrm{H}} \circ \widehat\phi = -\frac{M_2^2 m_2^3}{2\widehat\Lambda_2^2} - \alpha^2 \widehat{\mathrm{P}}(\widehat\Lambda, \widehat\lambda_1, \widehat{\mathrm{z}}), \quad \widehat{\mathrm{P}}(\widehat\Lambda, \widehat\lambda_1, \widehat{\mathrm{z}}) := \frac{M_2^2 m_2^3}{4} \frac{\widehat{\mathrm{r}}_1^2}{\widehat\Phi_2^4}(3\widehat\Theta^2 - \widehat\Phi_2^2) \frac{1}{2\pi} \int_{\mathbb{T}} \frac{d\widehat\lambda_2}{\widehat{\mathrm{r}}_2^2}, \quad (2.24)$$

where $\widehat\Theta = \widehat{\mathrm{C}}^{(2)} \cdot \widehat x^{(1)}$, $\widehat{\mathrm{r}}_i = |\widehat x^{(i)}|$, $\widehat\Phi_2 = |\widehat{\mathrm{C}}^{(2)}|$ have to be regarded as functions of $(\widehat\Lambda, \widehat\lambda, \widehat z)$. Note that we have used that $\widehat\Theta = \widehat{\mathrm{C}}^{(2)} \cdot \widehat x^{(1)}$, $\widehat{\mathrm{r}}_1 = |\widehat x^{(1)}|$, $\widehat\Phi_2 = |\widehat{\mathrm{C}}^{(2)}|$ are independent of $\widehat\lambda_2$. By construction, the overall change

$$\bar\phi^{-1} \circ \phi \circ \phi_0 \circ \phi_1 \circ \phi^{-1} \circ \bar\phi: \qquad (\widehat\Lambda, \widehat\lambda, \widehat{\mathrm{z}}) \to (\Lambda, \lambda, \mathrm{z})$$

is symplectic, α-close to the identity and puts the Hamiltonian H in (2.20) into the form claimed in Lemma 3. By the uniqueness claimed by this theorem, in comparison with the result of Lemma 4, we have that $\widehat{\mathrm{P}}(\widehat\Lambda, \widehat\lambda_1, \widehat{\mathrm{z}})$ in (2.24) satisfies

$$\alpha^2 \widehat{\mathrm{P}}(\widehat\Lambda, \widehat\lambda_1, \widehat{\mathrm{z}}) = -\alpha^2 \frac{M_2^2 m_2^3}{4} \frac{\widehat{\mathrm{r}}_1^2}{\widehat\Phi_2^4}(3\widehat\Theta^2 - \widehat\Phi_2^2) \frac{1}{2\pi} \int_{\mathbb{T}} \frac{d\widehat\lambda_2}{\widehat{\mathrm{r}}_2^2} \equiv \alpha^2 M_2 m_2 \mathrm{P}^{(1)}(\widehat\Lambda, \widehat\lambda_1, \widehat{\mathrm{z}}) + \mathrm{O}(\alpha^3),$$

where $\mathrm{P}^{(1)}$ is as in (2.19) (and, as above, $\widehat\Theta$, $\widehat\Phi_2$, $\widehat{\mathrm{r}}_1$ and $\widehat{\mathrm{r}}_1$ are regarded as functions of $(\widehat\Lambda_1, \widehat\Lambda_2, \widehat\lambda_1, \widehat\lambda_2, \widehat{\mathrm{z}})$). This formula is (2.9), neglecting $\frac{1}{2\pi}\int_{\mathbb{T}} d\lambda_1$. In particular, it proves the second item in Remark 1 and implies (2.9). □

2.5. Proof of Proposition 1

We shall need definitions and a result from [15], to which paper we refer for notations and details.

Let, as in [15], $\mathcal{P}^{6n}_{\mathrm{P}}$, $\mathcal{P}^{6n}_{\mathrm{rps}} \subset \mathbb{R}^{3n} \times \mathbb{R}^{3n}$ denote the respective domains of the maps

$$\phi_{\mathrm{P}}: \ (y, x) \in \mathcal{P}^{6n}_{\mathrm{P}} \to (\Lambda, \lambda, \mathrm{z}) \in \mathbb{R}^n \times \mathbb{T}^n \times \mathbb{R}^{4n}, \quad \phi_{\mathrm{rps}}: \ (y, x) \in \mathcal{P}^{6n}_{\mathrm{rps}} \to (\Lambda, \lambda, z) \in \mathbb{R}^n \times \mathbb{T}^n \times \mathbb{R}^{4n}$$

between "Cartesian" and, respectively, Poincaré, RPS variables. Consider the common domain of ϕ_{P} and ϕ_{rps}, i.e. the set $\mathcal{P}^{6n}_{\mathrm{rps}} \cap \mathcal{P}^{6n}_{\mathrm{P}}$. On the ϕ_{rps}-image of such domain consider the symplectic map

$$\phi^{\mathrm{rps}}_{\mathrm{P}}: \quad (\Lambda, \lambda, z) \to (\Lambda, \lambda, \mathrm{z}) := \phi_{\mathrm{P}} \circ \phi^{-1}_{\mathrm{rps}} \qquad (2.25)$$

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

Theorem 5 ([15]). *The symplectic map $\phi^{\mathrm{rps}}_{\mathrm{P}}$ in (2.25) has the form*

$$\lambda = \lambda + \varphi(\Lambda, z) \qquad \mathrm{z} = \mathcal{Z}(\Lambda, z), \qquad (2.26)$$

where $\varphi(\Lambda, 0) = 0$ and, for any fixed Λ, the map $\mathcal{Z}(\Lambda, \cdot)$ is 1:1, symplectic[38) and its projections verify

$$\Pi_{\eta}\mathcal{Z} = \eta + \mathrm{O}(|z|^3), \ \Pi_{\xi}\mathcal{Z} = \xi + \mathrm{O}(|z|^3), \ \Pi_{\mathrm{p}}\mathcal{Z} = \mathcal{V}p + \mathrm{O}(|z|^3), \ \Pi_{\mathrm{q}}\mathcal{Z} = \mathcal{V}q + \mathrm{O}(|z|^3)$$

for some $\mathcal{V} = \mathcal{V}(\Lambda) \in \mathrm{SO}(n)$.

38) I.e., it preserves the two form $d\eta \wedge d\xi + dp \wedge dq$.

Now we proceed to prove Proposition 1. Consider the inverse maps

$$\phi_{\rm rps}^{-1}: \qquad (\Lambda,\lambda,z)\in\mathcal{M}^{6n}_{\rm rps}\to\big(y_{\rm rps}(\Lambda,\lambda,z),x_{\rm rps}(\Lambda,\lambda,z)\big)$$

$$\phi_{\rm P}^{-1}: \qquad (\Lambda,\mathrm{\lambda},\mathrm{z})\in\mathcal{M}^{6n}_{\rm P}\to\big(y_{\rm P}(\Lambda,\mathrm{\lambda},\mathrm{z}),x_{\rm P}(\Lambda,\mathrm{\lambda},\mathrm{z})\big)$$

with $\mathcal{M}^{6n}_{\rm rps}:=\phi_{\rm rps}(\mathcal{P}^{6n}_{\rm rps})$, $\mathcal{M}^{6n}_{\rm P}:=\phi_{\rm rps}(\mathcal{P}^{6n}_{\rm P})$. Let $y^{(i)}_{\rm rps}\in\mathbb{R}^3$, $\cdots$ be the $i^{\rm th}$ projection of $y_{\rm rps}$, $\cdots$; i.e., to be defined by

$$y_{\rm rps}=\big(y^{(1)}_{\rm rps},\cdots y^{(n)}_{\rm rps}\big),\quad \cdots$$

Let, finally,

$$\alpha_{ij}^2(f_{ij}^{(2)})_{\rm P}:=\frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}d\mathrm{\lambda}_i d\mathrm{\lambda}_j\frac{3(x_{\rm P}^{(i)}(\Lambda,\mathrm{\lambda}_i,\mathrm{z})\cdot x_{\rm P}^{(j)}(\Lambda,\mathrm{\lambda}_j,\mathrm{z}))^2-|x_{\rm P}^{(i)}(\Lambda,\mathrm{\lambda}_i,\mathrm{z})|^2|x_{\rm P}^{(j)}(\Lambda,\mathrm{\lambda}_j,\mathrm{z})|^2}{2|x_{\rm P}^{(j)}(\Lambda,\mathrm{\lambda}_j,\mathrm{z})|^5}$$

and[39)]

$$\alpha_{ij}^2(f_{ij}^{(2)})_{\rm rps}:=\frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}d\lambda_i d\lambda_j\frac{3(x_{\rm rps}^{(i)}(\Lambda,\lambda_i,z)\cdot x_{\rm rps}^{(j)}(\Lambda,\lambda_j,z))^2-|x_{\rm rps}^{(i)}(\Lambda,\lambda_i,z)|^2|x_{\rm rps}^{(j)}(\Lambda,\lambda_j,z)|^2}{2|x_{\rm rps}^{(j)}(\Lambda,\lambda_j,z)|^5}.$$

We shall use the following properties, easily deducible from [15]:

(i) For $1\leqslant i\leqslant n$, $y^{(i)}_{\rm rps}$, $x^{(i)}_{\rm rps}$ depend on λ only via λ_i. Analogously, $y^{(i)}_{\rm P}$, $x^{(i)}_{\rm P}$ depend on $\mathrm{\lambda}$ only via $\mathrm{\lambda}_i$. In particular $y^{(i)}_{\rm P}$, $x^{(i)}_{\rm P}$ depend on Λ only via Λ_i and depend on z only via z_i, but this will not be used.

(ii) For any $1\leqslant i<j\leqslant n$, the map

$$(\Lambda_i,\Lambda_j,\mathrm{\lambda}_i,\mathrm{\lambda}_j,\mathrm{z}_i,\mathrm{z}_j)\to\big(y_{\rm P}^{(i)}(\Lambda,\mathrm{\lambda}_i,\mathrm{z}),y_{\rm P}^{(j)}(\Lambda,\mathrm{\lambda}_j,\mathrm{z}),x_{\rm P}^{(i)}(\Lambda,\mathrm{\lambda}_i,\mathrm{z}),x_{\rm P}^{(j)}(\Lambda,\mathrm{\lambda}_j,\mathrm{z})\big) \tag{2.27}$$

satisfies assumptions (A) and (B) of Proposition 2. Note that, unless we are in the case $n=2$, this is not true for the map

$$(\Lambda,\lambda_i,\lambda_j,z)\to\big(y_{\rm rps}^{(i)}(\Lambda,\lambda_i,z),y_{\rm rps}^{(j)}(\Lambda,\lambda_j,z),x_{\rm rps}^{(i)}(\Lambda,\lambda_i,z),x_{\rm rps}^{(j)}(\Lambda,\lambda_j,z)\big). \tag{2.28}$$

In particular, both (2.27) and (2.28) satisfy assumption (A) (for any n and any $1\leqslant i<j\leqslant n$), but assumption (B) fails for (2.28) (when $n>2$).

(iii) Letting $\mathrm{C}^{(i)}_{\rm rps}:=x^{(i)}_{\rm rps}\times y^{(i)}_{\rm rps}$ and, analogously, $\mathrm{C}^{(i)}_{\rm P}:=x^{(i)}_{\rm P}\times y^{(i)}_{\rm P}$, then, for any $1\leqslant i\leqslant n$, $\mathrm{C}^{(i)}_{\rm rps}$ does not depend on λ_i and, analogously, $\mathrm{C}^{(i)}_{\rm P}$ does not depend on $\mathrm{\lambda}_i$. This is because, as remarked in (ii), both (2.27) and (2.28) satisfy (A).

By the previous items, may apply Proposition 2 to the map (2.27). We find

$$\begin{aligned}\alpha_{ij}^2(f_{ij}^{(2)})_{\rm P}(\Lambda,\mathrm{z})=&-\frac{M_jm_j^2}{4}\\ &\times\frac{\frac{1}{2\pi}\int_{\mathbb{T}}\big(3(\mathrm{C}_{\rm P}^{(j)}(\Lambda,\mathrm{z})\cdot x_{\rm P}^{(i)}(\Lambda,\mathrm{\lambda}_i,\mathrm{z}))^2-|x_{\rm P}^{(i)}(\Lambda,\mathrm{\lambda}_i,\mathrm{z})|^2|\mathrm{C}_{\rm P}^{(j)}(\Lambda,\mathrm{\lambda}_j,\mathrm{z})|^2\big)d\mathrm{\lambda}_i}{|\mathrm{C}_{\rm P}^{(j)}(\Lambda,\mathrm{\lambda}_j,\mathrm{z})|^4}\\ &\times\frac{1}{2\pi}\int_{\mathbb{T}}\frac{d\mathrm{\lambda}_j}{|x_{\rm P}^{(j)}(\Lambda,\mathrm{\lambda}_j,\mathrm{z})|^2}.\end{aligned}$$

Letting now $\mathrm{z}=\mathcal{Z}(\Lambda,z)$ and changing the integration variables $\mathrm{\lambda}_i=\lambda_i+\varphi_i(\Lambda,z)$ with $\mathcal{Z}$, φ as in (2.26) we have the thesis. $\square$

[39)] As observed in footnote 36, the map $\phi_{\rm rps}^{-1}$ depends explicitly on (p_n,q_n), while SO(3)-invariant expressions, such as the right hand side of the formula below, do not.

3. PROOF OF THEOREM A

In this section, we aim to prove Theorem A. The preliminaries are as follows.

(i) In the case of spatial three-body problem, let

$$\mathcal{H}_{3b} := h_{\rm Kep}(\Lambda) + \mu f_{3b}(\Lambda, \lambda, \bar{z}) \tag{3.1}$$

the Hamiltonian $\mathcal{H}_{\rm rps}$ in (1.14) for $n = 2$. Here, a_j have to be regarded as functions of Λ_j, via

$$a_j(\Lambda_j) = \frac{1}{M_j}\left(\frac{\Lambda_j}{m_j}\right)^2. \tag{3.2}$$

Let

$$\mathcal{A}: \quad a_- \leqslant a_1 \leqslant \alpha a_2 \leqslant a_+. \tag{3.3}$$

Let $f_{ij}^{(k)}$ be as in (2.3); then split

$$(f_{3b}(\Lambda, \lambda, \bar{z}))_{\rm av} = N + \widetilde{N} \tag{3.4}$$

with

$$N := (f_{3b}(\Lambda, \lambda, \bar{z}))_{\rm av}^{(0)} + (f_{3b}(\Lambda, \lambda, \bar{z}))_{\rm av}^{(2)}, \quad \widetilde{N} := -\bar{m}_1\bar{m}_2 \sum_{j=3}^{\infty} f_{12}^{(j)} \tag{3.5}$$

where $(f_{3b}(\Lambda, \lambda, \bar{z}))_{\rm av}^{(0)}$ does not depend on $\bar{z}$, $(f_{3b}(\Lambda, \lambda, \bar{z}))_{\rm av}^{(2)}$ corresponds to $(f_{\rm av})_{\rm av}^{(2)}$ in (1.19) for this case. As we shall discuss in Claim 4, N is *integrable* and, due to the choice of $\mathcal{A}$, $\tilde{N}$ verifies $|\tilde{N}| \leqslant$ const α^3.

(ii) In the case of the planar $(1 + n)$-body problem, let

$$\mathcal{H}_{\rm pl} := h_{\rm Kep}(\Lambda) + \mu f_{\rm pl}(\Lambda, \lambda, z_{\rm pl})$$

the Hamiltonian $\mathcal{H}_{\rm rps}$ in (1.14), with $(\bar{p}, \bar{q}) = 0$. Assume the following asymptotics for semi-axes. Fix three numbers $0 < \underline{a} < \overline{a}$, $\alpha < \frac{\underline{a}}{\overline{a}}$. Then take

$$\mathcal{A}_{\rm pl}: \qquad \underline{a}_j \leqslant a_j \leqslant \bar{a}_j \quad 1 \leqslant j \leqslant n \tag{3.6}$$

(again with $a_j = a_j(\Lambda_i)$ as in (3.2)) where

$$\underline{a}_j := \alpha^{2[(\frac{3}{2})^{n-j}-1]}\underline{a}, \quad \overline{a}_j := \alpha^{2[(\frac{3}{2})^{n-j}-1]}\overline{a}, \quad \underline{a}_n := \underline{a}, \quad \bar{a}_n := \bar{a}$$

with $1 \leqslant j \leqslant n - 1$. Notice that this asymptotics requires that the a_j's are closer and closer as j increases.

By (2.1) and (2.3),

$$(f_{\rm pl})_{\rm av} = N_{\rm pl} + \widetilde{N}_{\rm pl}$$

with

$$N_{\rm pl} := (f_{\rm pl})_{\rm av}^{(0)} + (f_{\rm pl})_{\rm av}^{(2)}, \quad \widetilde{N}_{\rm pl} := -\sum_{1\leqslant i<j\leqslant n} \bar{m}_i\bar{m}_j \sum_{k=3}^{\infty} f_{ij}^{(k)}|_{\rm pl}$$

where $(f_{\rm pl})_{\rm av}^{(0)}$ does not depend on $z_{\rm pl} = (\eta, \xi)$, $(f_{\rm pl})_{\rm av}^{(2)}$ is as in (1.22) and $f_{ij}^{(k)}|_{\rm pl}$ are as in (2.3), in the planar case. As discussed in (1.22), $(f_{\rm pl})_{\rm av}^{(2)}$ is *integrable* and, moreover, due to the choice of the $\underline{a}_j$, $\overline{a}_j$,

$$\begin{cases} |(f_{\rm pl})_{\rm av}^{(2)}| \leqslant \text{const} \displaystyle\max_{1\leqslant i<j\leqslant n} \sup_{\mathcal{A}} \left(\frac{a_i^2}{a_j^3}\right) \leqslant \text{const}\,\frac{1}{\underline{a}}\left(\alpha\frac{\overline{a}}{\underline{a}}\right)^2 \\ |\widetilde{N}_{\rm pl}| \leqslant \text{const} \displaystyle\max_{1\leqslant i<j\leqslant n} \sup_{\mathcal{A}} \left(\frac{a_i^3}{a_j^4}\right) \leqslant \text{const}\,\frac{1}{\underline{a}}\left(\alpha\frac{\overline{a}}{\underline{a}}\right)^3. \end{cases}$$

As for the proof for the three-body problem, we shall need some details from [16] and [15] (to which papers we refer for proofs) that we recall in the following section.

3.1. Symmetries of the Partially Reduced System

The Hamiltonian (1.1) remains unvaried by reflections with respect to coordinate planes $\{x_1 = x_2\}$, $\{x_3 = 0\}$ or rotations, for example, around the $k^{(3)}$-axis. These transformations are, respectively,

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

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

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

Note, in particular, that $\mathcal{R}_3^-$ and $\mathcal{R}_g$ are symplectic transformations, while $\mathcal{R}_{1\leftrightarrow 2}$ is an involution. The expressions of $\mathcal{R}_{1\leftrightarrow 2}$, $\mathcal{R}_3^-$ and $\mathcal{R}_g$ in terms of the variables (1.13) turn out to be the same[40] as in Poincaré variables. They are

$$\begin{aligned}
&\mathcal{R}_{1\leftrightarrow 2}\Big(\Lambda,\ \lambda,\ z\Big) := \Big(\Lambda,\ \tfrac{\pi}{2} - \lambda,\ \mathcal{S}_{1\leftrightarrow 2} z\Big); \quad \mathcal{R}_3^-\Big(\Lambda,\ \lambda,\ z\Big) = \Big(\Lambda,\ \lambda,\ \mathcal{S}_{34}^- z\Big)\\
&\mathcal{R}_g\Big(\Lambda,\ \lambda,\ z\Big) = \Big(\Lambda,\ \lambda + g,\ \mathcal{S}_g z\Big)
\end{aligned} \tag{3.7}$$

where

$$\begin{cases}
\mathcal{S}_{1\leftrightarrow 2}(\eta, \xi, p, q) := (\xi, \eta, q, p)\\[2ex]
\mathcal{S}_{34}^-(\eta, \xi, p, q) := (\eta, \xi, -p, -q)\\[2ex]
\mathcal{S}_g: \ \Big(\eta_j + \mathrm{i}\xi_j, p_j + \mathrm{i}q_j\Big) \to \Big(e^{-\mathrm{i}g}(\eta_j + \mathrm{i}\xi_j),\ e^{-\mathrm{i}g}(p_j + \mathrm{i}q_j)\Big)
\end{cases}$$

with $\mathrm{i} := \sqrt{-1}$.

Since the Hamiltonian $\mathcal{H}_{\mathrm{rps}}$ (1.14) is independent of (p_n, q_n), in the above transformations, we may neglect this latter couple of variables and replace[41] z with $\bar z$ in (3.7). In particular, the one-parameter group $\{\bar{\mathcal{R}}_g\}_{g\in\mathbb{T}}$ defined by

$$\bar{\mathcal{R}}_g: \ (\Lambda,\ \lambda,\ \bar z,\ p_n,\ q_n) \to (\Lambda,\ \lambda + g,\ \mathcal{S}_g \bar z,\ p_n,\ q_n) \qquad g \in \mathbb{T} \tag{3.8}$$

leaves $\mathcal{H}_{\mathrm{rps}}$ unvaried. This group of transformations corresponds to be the time-g flow of

$$G = \sum_{i=1}^{n} \Lambda_i - \sum_{i=1}^{n} \frac{\eta_i^2 + \xi_i^2}{2} - \sum_{i=1}^{n-1} \frac{p_i^2 + q_i^2}{2} \tag{3.9}$$

which is the Euclidean length of the angular momentum (1.10): $G = |\mathrm{C}|$, expressed in the variables (1.13). Therefore, $\bar{\mathcal{R}}_g$ may be identified to be the group g-rotations about the C-axis.

[40] See, for example [22].

[41] Recall the definitions in (1.13)–(1.15).

In view of such relations, amusing symmetries (discussed[42] in [16]) appear among the Taylor coefficients of the expansion of the perturbation $f_{\rm rps}$ and hence also of its averaged value $(f_{\rm rps})_{\rm av}$. These symmetries are often referred to (for the classical Poincaré system (1.3)) as *D' Alembert rules*. To describe such relations, we switch[43] to "Birkhoff coordinates"

$$w_i = \frac{\eta_i - {\rm i}\xi_i}{\sqrt{2}}, \quad w_{n+j} = \frac{p_j - {\rm i}q_j}{\sqrt{2}}, \quad w_i^\star = \frac{\eta_i + {\rm i}\xi_i}{{\rm i}\sqrt{2}}, \quad w_{n+j}^\star = \frac{p_j + {\rm i}q_j}{{\rm i}\sqrt{2}} \tag{3.10}$$

with $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant n-1$ and we regard (abusively) $f_{\rm rps}$ and $(f_{\rm rps})_{\rm av}$ as functions of $(\Lambda, \lambda, w, w^\star)$.

Claim 3 ([15, 16]).

(i) $\mathcal{R}_3$-invariance implies that $f_{\rm rps}$ is even in $(w_{n+1}, \cdots w_{2n-1}, w^\star_{n+1}, \cdots w^\star_{2n-1})$ (equivalently, it is even in $(\bar p, \bar q)$);

(ii) $\bar{\mathcal{R}}_g$-invariance implies that, the only non-vanishing monomials appearing in the Taylor expansion of $(f_{\rm rps})_{\rm av}$ in powers $\{w_i,\ w_i^*\}_{1\leqslant i\leqslant 2n-1}$ are those with literal part $w^\alpha {w^*}^{\alpha^*}$ for which

$$\sum_{i=1}^{2n-1} (\alpha_i - \alpha_i^*) = 0. \tag{3.11}$$

Claim 3 and the independence of $f_{12}^{(2)}$ on the argument of (η_2, ξ_2) (see the Introduction) have the following corollary. Let $\mathcal{A}$ be as in (3.3) and let $\mathcal{M}^{10}_{\epsilon_0} := \mathcal{A} \times \mathbb{T}^2 \times B^6_{\epsilon_0}$.

Claim 4. N (namely[44], $f_{12}^{(2)}$) is integrable. More precisely: (i) it depends on (η_2, ξ_2) only via $\frac{\eta_2^2+\xi_2^2}{2}$; (ii) one can find $\epsilon_0 > 0$ and a symplectic change of variables

$$(\Lambda, \breve\lambda, \breve z) \to (\Lambda, \lambda, \bar z)$$

defined on the phase space $\mathcal{M}^{10}_{\epsilon_0} := \mathcal{A} \times \mathbb{T}^2 \times B^6_{\epsilon_0}$ of the form

$$\breve\phi: \quad \Lambda = \Lambda, \quad \lambda = \breve\lambda + \varphi(\Lambda, \breve z), \quad \bar z = \breve{\mathcal{Z}}(\Lambda, \breve z) \tag{3.12}$$

defined for $|\breve z| < \epsilon_0$ which transforms N into a new function $\breve N(\Lambda, \breve z)$ depending only on $\dfrac{\breve\eta_1^2 + \breve\xi_1^2}{2}$, $\dfrac{\breve\eta_2^2 + \breve\xi_2^2}{2}$ and $\dfrac{\breve p_1^2 + \breve q_1^2}{2}$. In particular, $\breve\psi$ preserves $\dfrac{\breve\eta_2^2 + \breve\xi_2^2}{2}$ and $\dfrac{\breve\eta_1^2 + \breve\xi_1^2}{2} + \dfrac{\breve p_1^2 + \breve q_1^2}{2}$.

Proof. Since $f_{12}^{(2)}$ is even in $(\bar p, \bar q) = (p_1, q_1)$ and has only monomials with $\alpha_2 = \alpha_2^\star$, Eq. (3.11) with $n = 2$ implies that $f_{12}^{(2)}$ is even in (η_1, ξ_1), (η_2, ξ_2) and (p_1, q_1) separately. Moreover, $f_{12}^{(2)}$ is integrable[45]. Let $\bar z = \breve{\mathcal{Z}}(\Lambda, \breve z)$ the transformation (parametrized by Λ) verifying

$$\sum_{i=1}^{2} d\eta_i \wedge d\xi_i + dp_1 \wedge dq_1 = \sum_{i=1}^{2} d\breve\eta_i \wedge d\breve\xi_i + d\breve p_1 \wedge d\breve q_1$$

---

42) In [16], $\bar{\mathcal{R}}_g$-invariance is called "rotation invariance". Here, to avoid confusions, we reserve this name only to the transformations (1.9).

43) $d\eta_i \wedge d\xi_i = dw_i \wedge dw_i^*$ and $dp_j \wedge dq_j = dw_{j+n} \wedge dw^*_{j+n}$

44) Recall that $f_{12}^{(0)}$ is independent of $\bar z$.

45) To integrate $f_{12}^{(2)}$, one can first reduce the integral $G_0 := \frac{\breve\eta_1^2 + \breve\xi_1^2}{2} + \frac{p_1^2+q_1^2}{2}$ via the change of variables

$$\eta_1 + {\rm i}\xi_1 = (\breve\eta_1 + \breve{\rm i}\xi_1) e^{{\rm i}g_0}, \quad p_1 + {\rm i}q_1 = \sqrt{2(G_0 - \frac{\breve\eta_1^2 + \breve\xi_1^2}{2})} e^{{\rm i}g_0}$$

with g_0 cyclic in $f_{12}^{(2)}$ (but not in $f_{3\rm b}$). Note that this reduction does not cause singularities in $f_{3\rm b}$, since $f_{3\rm b}$ is even in (p_1, q_1). Next, once $f_{12}^{(2)}$ is reduced to one degree of freedom, its integration is trivial.

such that $\breve{N}(\Lambda, \breve{z}) := \bar{N} \circ \breve{\mathcal{Z}}$ has the claimed properties. Then, it is standard to prove that $\breve{z} \to \breve{\mathcal{Z}}(\Lambda, \breve{z})$ may be lifted to a transformation as in (3.12) (compare, for example, [16, Proposition 7.3]). □

3.2. KAM *Theory*

In this section we complete the proof of Theorem A.

Let ϵ_0, $\breve{\phi}$ is as in Claim 4. For $(\Lambda, \breve{\lambda}, \breve{z}) \in \mathcal{M}^{10} := \mathcal{A} \times \mathbb{T}^2 \times B^6_{\epsilon_0}$, define

$$\begin{aligned}\breve{\mathcal{H}}_{3b}(\Lambda, \breve{\lambda}, \breve{z}) &:= \mathcal{H}_{3b} \circ \breve{\phi}(\Lambda, \breve{\lambda}, \breve{z}) \\ &= h_{\rm Kep}(\Lambda) + \mu \breve{f}_{3b}(\Lambda, \breve{\lambda}, \breve{z})\end{aligned} \tag{3.13}$$

where $\breve{\phi}$ is as in Claim 4. By Claim 4

$$(\breve{f}_{3b})_{\rm av} = \breve{N} + \tilde{N} \tag{3.14}$$

where $\breve{N}$ depends only on $\frac{\breve{\eta}_1^2+\breve{\xi}_1^2}{2}$, $\frac{\breve{\eta}_2^2+\breve{\xi}_2^2}{2}$, $\frac{\breve{p}_1^2+\breve{q}_1^2}{2}$ and

$$|\tilde{N}| \leqslant \text{const}\, \alpha^3. \tag{3.15}$$

To the system (3.13) and to the system $\mathcal{H}_{\rm pl}$ defined at the beginning of Section 3 (compare item (ii)) we shall apply an abstract result (Theorem 6 below) that refines and generalizes Theorem 9; see Remark 2. This is as follows.

Let n_1, $n_2 \in \mathbb{N}$, $B^{2n_2}_\epsilon = \{y \in \mathbb{R}^{2n_2} : |y| < \epsilon\}$ denote the $2n_2$-ball of radius ϵ and let

$$\mathcal{P}_{\epsilon_0} := V \times \mathbb{T}^{n_1} \times B^{2n_2}_{\epsilon_0} \tag{3.16}$$

where V is a open, connected set of $\mathbb{R}^{n_1}$. Let

$$H(I, \varphi, p, q; \mu) := H_0(I) + \mu P(I, \varphi, p, q; \mu) \tag{3.17}$$

be real-analytic on $\mathcal{P}_{\epsilon_0}$ and such that

(i) $\omega_0 := \partial H_0$ is a real-analytic diffeomorphism of V;

(ii) the average $P_{\rm av}(I, p, q; \mu) = \dfrac{1}{(2\pi)^{n_1}} \displaystyle\int_{\mathbb{T}^{n_1}} P(I, \varphi, p, q; \mu) d\varphi$ has the form

$P_{\rm av}(I, p, q; \mu, \alpha) = N(I, J; \mu) + \tilde{N}(I, p, q; \mu)$, where

$J = (\frac{p_1^2+q_1^2}{2}, \cdots, \frac{p_{n_2}^2+q_{n_2}^2}{2})$ and $\sup_{V \times B^{2n_2}} |\tilde{N}| \leqslant \kappa$;

(iii) the Hessians $\partial^2_I H_0$ $\partial^2_{I,J} N(I, J; \mu)$ do not vanish, respectively, on V, $V \times B^{2n_2}_{\epsilon_0}$.

Theorem 6. *Under the previous assumptions, one can find positive numbers C_*, μ_*, $\kappa_\star$, $\epsilon_1 < \epsilon_0$ depending only on H and ϵ_0 and an integer β depending only on n_1, n_2, such that, for*

$$|\mu| < \mu_*, \quad |\kappa| < \kappa_*, \quad |\mu| < (\log \kappa^{-1})^{-2\beta} \tag{3.18}$$

a set $\mathcal{K} \subset \mathcal{P}_{\epsilon_1}$ exists, formed by the union of H-invariant n-dimensional tori, on which the H-motion is analytically conjugated to linear Diophantine quasi-periodic motions. The set $\mathcal{K}$ is of positive Liouville–Lebesgue measure and satisfies

$$\text{meas}\, \mathcal{K} > \Big(1 - C_*(\sqrt[4]{\mu}(\log \kappa^{-1})^\beta + \sqrt{\kappa})\Big)\, \text{meas}\, \mathcal{P}_{\epsilon_1}. \tag{3.19}$$

Remark 2. Theorem 6 generalizes and refines Theorem 9: to obtain Theorem 9 from Theorem 6 it is sufficient to take $\kappa = \mu$. In this case condition (3.18) becomes just a smallness condition on μ (as inTheorem 9) and, by (3.19), $\mathcal{K}$ fills $\mathcal{P}_{\epsilon_1}$ up to a set of density $(1 - \tilde{C}\mu^a)$ with any $0 < a < \frac{1}{4}$. This should be compared with the measure estimate given in Theorem 9, where $a \sim \frac{1}{n}$.

The proof of Theorem 6 is completely analogous to the proof of [12, Theorem 1.4] and hence is only sketched in Appendix C.

We can now conclude the
Proof of Theorem A.

(i) In the case of the spatial three-body problem, apply Theorem 6 to $\breve{\mathcal{H}}_{3b} := \mathcal{H}_{3b} \circ \breve{\phi}$ (where $\breve{\phi}$ is as in (3.12)), hence, with

$$n_1 = 2, \quad n_2 = 3, \quad V = \mathcal{A}, \quad \kappa = \text{const}\,\alpha^3, \quad N = \breve{N}$$

where $\mathcal{A}$ is as in (3.3), $\breve{N}$ as in (3.14) and ϵ_0 as in Claim 4.

(i) In the case of the planar $(1+n)$-body problem with $n \geqslant 3$, apply Theorem 6 with

$$n_1 = n, \quad n_2 = n, \quad V = \mathcal{A}_{\text{pl}}, \quad \kappa = \text{const}\,\bar{\alpha}^3, \quad N = N_{\text{pl}},$$

where $\mathcal{A}_{\text{pl}}$ is as in (3.6), $\bar{\alpha} := \alpha\frac{\overline{a}}{\underline{a}}$ and ϵ_0 so small to avoid collisions. □

4. PROOF OF THEOREM B

In this section, we shall prove the following theorem, which is a more detailed statement of Theorem B. Let

$$\mathcal{H}_{\text{pl3b}} = h_{\text{Kep}} + \mu f_{\text{pl3b}} := \mathcal{H}_{3b}|_{p_1=q_1=0} \tag{4.1}$$

and denote as

$$\mathcal{M}^8_{\text{pl3b}}: \quad \overline{a}_- \leqslant a_1(\Lambda_1) \leqslant \alpha a_2(\Lambda_2) \leqslant \alpha\overline{a}_+, \quad \underline{\epsilon} \leqslant |z_{\text{pl}}| \leqslant \epsilon \leqslant \bar{\epsilon}, \quad \lambda_1,\ \lambda_2 \in \mathbb{T} \tag{4.2}$$

its eight-dimensional phase space, where $a_j(\Lambda_j)$ are as in (3.2). Here, $\mathcal{H}_{3b}$ is as in (3.1) and $z_{\text{pl}} := (\eta_1, \eta_2, \xi_1, \xi_2)$.

Theorem 7. *There exists positive numbers $\bar{\epsilon}$, $\bar{\alpha}$, $\bar{\mu}$, $\bar{\beta}$, τ, $\bar{K}_\star$, $\bar{a}$, $\bar{b}$, $\bar{c}$, $\bar{d}$ such that, if*

$$0 < \alpha \leqslant \bar{\alpha}, \quad 0 < \mu \leqslant \bar{\mu}, \quad \mu \leqslant \bar{c}(\log \epsilon^{-1})^{-\bar{\beta}}$$

one can find a an open set $\bar{\mathcal{M}}^8_{\text{pl3b}} \subset \mathcal{M}^8_{\text{pl3b}}$ defined by the following inequalities for the Keplerian frequencies $\omega_{\text{Kep}} := \partial_{\overline{\Lambda}} h_{\text{Kep}}$

$$|\omega_{\text{Kep}} \cdot k| \geqslant \frac{\sqrt[4]{\mu}}{\bar{c}\bar{K}} \quad \forall k:\ 0 < |k|_1 \leqslant \bar{K}$$

with

$$\bar{K} = \bar{K}_\star \log(\epsilon^{-1}) \tag{4.3}$$

such that for the $\mathcal{H}_{\text{pl3b}}$-flow starting from $\bar{\mathcal{M}}^8_{\text{pl3b}}$ the following holds. This flow is symplectically conjugated, via a $\{\mu^{1/12}, \epsilon^2\}$-close to the identity transformation ϕ to a flow

$$t \to (\tilde{\Lambda}_1(t), \tilde{\Lambda}_2(t), \tilde{\eta}_1(t), \tilde{\eta}_2(t), \tilde{\xi}_1(t), \tilde{\xi}_2(t))$$

such that, letting $\tilde{t}_i(t) := \frac{\tilde{\eta}_i^2 + \tilde{\xi}_i^2}{2}$, then, for $i = 1$, 2,

$$|\tilde{\Lambda}_i(t) - \tilde{\Lambda}_i(0)| \leqslant \delta^{\bar{b}},\ |\tilde{t}_i(t) - \tilde{t}_i(0)| \leqslant \delta^{\bar{b}} \quad \forall\ 0 \leqslant t \leqslant \frac{e^{\frac{1}{\delta^{\bar{a}}}}}{\delta},$$

with $\delta := \mu^{\bar{d}}\epsilon$.

For part of the proof, we shall deal with the *spatial* system $\mathcal{H}_{3\mathrm{b}}$ in (3.1). Next (in Section 4.4), letting $p_1 = q_1 = 0$, we shall reduce to the *planar* system $\mathcal{H}_{\mathrm{pl3b}}$ in (4.1).

Proof Step 0. Let us denote again as $\breve\phi$ a suitable symplectic transformation, whose existence is guaranteed by [16, 38], that conjugates $\mathcal{H}_{3\mathrm{b}}$ to a Hamiltonian $\breve{\mathcal{H}}_{3\mathrm{b}}$ having the same form as the one in (3.13)–(3.15), but with $\breve N + \tilde N$ in Birkhoff normal form up to order $2m$, with possibly smaller $\mathcal{A}$ of the form of (3.3), ϵ_0. In the domain (4.2), $\breve\phi$ is ϵ^2-close to the identity.

4.1. Step 1: The Birkhoff Normal Form of Order Six

In this section, we aim to compute the Birkhoff normal form of order six if the three-body problem (planar and spatial).

Let

$$\breve u_i := \frac{\breve\eta_i - \mathrm{i}\breve\xi_i}{\sqrt{2}}, \quad \breve u_i^\star := \frac{\breve\eta_i + \mathrm{i}\breve\xi_i}{\sqrt{2}\mathrm{i}}, \quad \breve v := \frac{\breve p_1 - \mathrm{i}\breve q_1}{\sqrt{2}}, \quad \breve v^\star := \frac{\breve p_1 + \breve{\mathrm{i}} q_1}{\sqrt{2}\mathrm{i}}. \tag{4.4}$$

We shall show that, if $t_1 := \mathrm{i}\breve u_1\breve u_1^\star$, $t_2 := \mathrm{i}\breve u_2\breve u_2^\star$, $t_3 := \mathrm{i}\breve v\breve v^\star$,

Claim 5. The Birkhoff normal form of order six of $(f_{3\mathrm{b}})_{\mathrm{av}}$ is given by (1.18).

Note that the $\big(1 + \mathrm{O}(\frac{\Lambda_1}{\Lambda_2})\big)$-factor in (1.18) has not been written for simplicity (it is available from below).

Proof. By Claim 4, the proof of (1.18) amounts to compute the Birkhoff normal form of order six of N in (3.5), up to an error of order $\frac{a_1^3}{a_2^4}$. The constant term $f_{12}^{(0)}$ in (3.5) contributes with $-\frac{\bar m_1\bar m_2}{a_2}$ to (1.18). We check that the Birkhoff normal form of $f_{12}^{(2)}$ is corresponds to what remains in (1.18). Recalling the definition of $f_{12}^{(2)}$ in (2.5) and the formulae in (1.21), (1.22) and (1.23), we have that the explicit formula of (2.6) in terms of RPS variables is

$$\begin{aligned} f_{12}^{(2)} = \frac{a_1^2}{4a_2^3}\Big(&1 + 3\mathrm{i}u_1u_1^\star\bar e_1^2 - 3\mathrm{i}vv^\star\bar{\mathfrak{s}}^2 - 9(\mathrm{i}u_1u_1^\star)(\mathrm{i}vv^\star)\bar{\mathfrak{s}}^2\bar e_1^2 \\ &- \frac{15}{2}\big((u_1^\star)^2v^2 + (v^\star)^2u_1^2\big)\bar e_1^2\bar{\mathfrak{s}}^2\Big)\mathrm{f}, \end{aligned} \tag{4.5}$$

where $\bar e_1$, $\bar{\mathfrak{s}}$, f are suitable functions of $\mathrm{i}u_1u_1^\star$, $\mathrm{i}u_2u_2^\star$ and $\mathrm{i}vv^\star$ (see Appendix B for more details). Here we shall need only the first terms of their respective Taylor expansions, which are

$$\begin{aligned} \bar e_1^2 &= \frac{1}{\Lambda_1} - \frac{\mathrm{i}u_1u_1^\star}{2\Lambda_1^2} \\ \bar{\mathfrak{s}}^2 &= \frac{1}{\Lambda_1} + \frac{1}{\Lambda_2} + \frac{\mathrm{i}u_1u_1^\star}{\Lambda_1^2} + \frac{\mathrm{i}u_2u_2^\star}{\Lambda_2^2} - \left(\frac{1}{4\Lambda_1^2} + \frac{1}{4\Lambda_2^2} + \frac{1}{\Lambda_1\Lambda_2}\right)\mathrm{i}vv^\star \\ &\quad + \frac{1}{\Lambda_1^3}(\mathrm{i}u_1u_1^\star)^2 + \frac{1}{\Lambda_2^3}(\mathrm{i}u_2u_2^\star)^2 - \left(\frac{1}{\Lambda_1^2\Lambda_2} + \frac{1}{2\Lambda_1^3}\right)(\mathrm{i}u_1u_1^\star)(\mathrm{i}vv^\star) \\ &\quad - \left(\frac{1}{\Lambda_1\Lambda_2^2} + \frac{1}{2\Lambda_2^3}\right)(\mathrm{i}u_2u_2^\star)(\mathrm{i}vv^\star) + \left(\frac{1}{4\Lambda_1\Lambda_2^2} + \frac{1}{4\Lambda_1^2\Lambda_2}\right)(\mathrm{i}vv^\star)^2 + \cdots \\ \mathrm{f} &= 1 + 3\frac{\mathrm{i}u_2u_2^\star}{\Lambda_2} + 6\left(\frac{\mathrm{i}u_2u_2^\star}{\Lambda_2}\right)^2 + 10\left(\frac{\mathrm{i}u_2u_2^\star}{\Lambda_2}\right)^3 + \cdots \end{aligned} \tag{4.6}$$

Since $f_{12}^{(2)}$ depends on $(u_2, u_2^\star)$ only via $\mathrm{i}u_2u_2^\star$, this "action" (besides being preserved by the transformation $\breve\psi$ in (3.12)) is also preserved at any step of Birkhoff normalization. Since the factor f in (4.5) depends only on $\mathrm{i}u_2u_2^\star$ (see Appendix B), we may leave such factor aside and look separately at the term inside parentheses

$$\mathrm{F} := 1 + 3\mathrm{i}u_1u_1^\star\bar{e}_1^2 - 3\mathrm{i}vv^\star\bar{\mathfrak{s}}^2 - 9(\mathrm{i}u_1u_1^\star)(\mathrm{i}vv^\star)\bar{\mathfrak{s}}^2\bar{e}_1^2 - \frac{15}{2}\big((u_1^\star)^2v^2 + (v^\star)^2u_1^2\big)\bar{e}_1^2\bar{\mathfrak{s}}^2.$$

Using this expression and (4.6), we see that the coefficients of $\mathrm{i}u_1u_1^\star$ and $\mathrm{i}vv^\star$ ("first order Birkhoff invariants"), are, respectively, given by[46)]

$$\Omega_{u_1} = \frac{3}{\Lambda_1}, \quad \Omega_v = -3\left(\frac{1}{\Lambda_1} + \frac{1}{\Lambda_2}\right).$$

Letting

$$f := -\frac{15}{2}\big((u_1^\star)^2v^2 + (v^\star)^2u_1^2\big)\bar{e}_1^2\bar{\mathfrak{s}}^2, \quad \phi := -\frac{15}{2}\frac{1}{2\mathrm{i}(\Omega_{u_1} - \Omega_v)}\big((u_1^\star)^2v^2 - (v^\star)^2u_1^2\big)\bar{e}_1^2\bar{\mathfrak{s}}^2,$$

one sees that the first step of Birkhoff normalization is obtained transforming F with the time-one flow of ϕ. Then F is transformed into

$$\mathrm{F}_1 := 1 + 3\mathrm{i}u_1u_1^\star\bar{e}_1^2 - 3\mathrm{i}vv^\star\bar{\mathfrak{s}}^2 - 9(\mathrm{i}u_1u_1^\star)(\mathrm{i}vv^\star)\bar{\mathfrak{s}}^2\bar{e}_1^2 + \frac{1}{2}\{\phi, f\} + \mathrm{o}(6).$$

where o(6) stands for an expression starting with degree seven in $(u_1, v, u_1^\star, v^\star)$. The Birkhoff normal form of order six of F, obtained with a further step of Birkhoff normalization, is then

$$\mathrm{F}_2 := 1 + 3\mathrm{i}u_1u_1^\star\bar{e}_1^2 - 3\mathrm{i}vv^\star\bar{\mathfrak{s}}^2 - 9(\mathrm{i}u_1u_1^\star)(\mathrm{i}vv^\star)\bar{\mathfrak{s}}^2\bar{e}_1^2 + \frac{1}{2}\Pi\{\phi, f\} + \mathrm{o}(6). \tag{4.7}$$

where $\frac{1}{2}\Pi\{\phi, f\}$ is obtained picking up normal terms[47)] of $\frac{1}{2}\{\phi, f\}$. But,

$$\frac{1}{2}\Pi\{\phi, f\} = \frac{225}{2}\frac{1}{(\Omega_{u_1} - \Omega_v)}((\mathrm{i}u_1u_1^\star)(\mathrm{i}vv)^2 - (\mathrm{i}u_1u_1^\star)^2(\mathrm{i}vv))\bar{\mathfrak{s}}^4\bar{e}_1^4 \tag{4.8}$$

where it is enough to replace $\bar{\mathfrak{s}}$, $\bar{e}_1$ with their respective lowest order terms in (4.6).

In view of (4.5), (4.6), (4.7) and (4.8), we have that (1.18) follows. □

4.2. *Step 2: Full Reduction of the SO(3)-symmetry*

The next step is to reduce completely the SO(3)-symmetry from the system $\breve{\mathcal{H}}_{3\mathrm{b}}$. Recall the definition of $\mathcal{A}$ in (3.3), ϵ_0 as in Claim 4.

Since the procedure we follow is analogue[48)] to the one in [16, Section 9], we shall skip some detail and refer to [16, Section 9] for complete information. We switch to a new set of symplectic variables $(\Lambda_1, \Lambda_2, G, \hat{u}_2, \hat{u}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{g}, \hat{u}_2^\star, \hat{u}_3^\star)$ defined via[49)]

$$\hat{\phi}: \quad \begin{cases} \Lambda_i = \Lambda_i \\ \breve{\lambda}_i = \hat{\lambda}_i + \hat{g} \end{cases} \quad \begin{cases} \breve{u}_2 = \hat{u}_2 e^{\mathrm{i}\hat{g}} \\ \breve{u}_2^\star = \hat{u}_2^\star e^{-\mathrm{i}\hat{g}} \end{cases} \quad \begin{cases} \breve{v}_2 = \hat{v}_2 e^{\mathrm{i}\hat{g}} \\ \breve{v}_2^\star = \hat{v}_2^\star e^{-\mathrm{i}\hat{g}} \end{cases}$$

---

46) Note that we do not need to assume non-resonance of (Ω_{u_1}, Ω_v) since N in (3.14) is integrable.

47) Ie, monomials of the form $(\mathrm{i}u_1u_1^\star)^\alpha(\mathrm{i}vv^\star)^\beta$.

48) The formulae in [16, Section 9] are a bit different from (4.9), since in [16, Section 9] we reduce the last couple of variables, denoted as [16, $(\breve{p}_{n-1}, \breve{q}_{n-1})$] (corresponding to $(\breve{p}_1, \breve{q}_1)$ in our case), while in (4.9), we reduce the first couple. This different choice has two reasons: (i) it provides simultaneously reduction in the planar and the spatial problem and (ii) formulae are a bit simpler, since the term t_1^3 does not appear in (1.18).

49) Analogue transformations were considered in [30].

$$\begin{cases} \breve{u}_1 = \sqrt{\varrho^2/2 - \hat{t}_2 - \hat{t}_3} e^{\mathrm{i}\hat{g}} \\ \\ \breve{u}_1^\star = -\mathrm{i}\sqrt{\varrho^2/2 - \hat{t}_2 - \hat{t}_3} e^{-\mathrm{i}\hat{g}} \end{cases} \tag{4.9}$$

with $\breve{u}_1$, $\breve{u}_2$, $\breve{v}$, $\breve{u}_1^\star$, $\breve{u}_2^\star$, $\breve{v}^\star$ defined as in (4.4), $\varrho^2/2 := \Lambda_1 + \Lambda_2 - G$, $\hat{t}_2 := \mathrm{i}\hat{u}_2\hat{u}_2^\star$, $\hat{t}_3 := \mathrm{i}\hat{v}\hat{v}^\star$. From the last couple of definitions, one sees that G is just the function in[50] (3.9) (with $n = 2$) and hence its conjugated angle, $\hat{g}$, is cyclic in the system. Let $(\hat{\eta}_2, \hat{\xi}_2)$, $(\hat{p}_1, \hat{q}_1)$ the real variables associated, respectively, to $(u_2, u_2^\star)$, $(v, v^\star)$ via (3.10) and $\hat{z} := (\hat{\eta}_2, \hat{p}_1, \hat{\xi}_2, \hat{q}_1)$. Fix $\varrho_\star < \epsilon_0$. There follows from [16, Remark 9.1-(iv)] that $\hat{\phi}$ is well defined and symplectic in the domain defined by $(\lambda_1, \lambda_2, \hat{g}) \in \mathbb{T}^3$ and

$$G \in \mathbb{R},\ (\Lambda_1, \Lambda_2) \in \mathcal{A}_G := \{(\Lambda_1, \Lambda_2) \in \mathcal{A}:\ 0 < \varrho_\star \leqslant \varrho(\Lambda, G) < \epsilon_0\}, \quad |\hat{z}| < \varrho_\star.$$

As usual, being $\hat{g}$ cyclic, we regard G as an external fixed parameter so as to have a reduced (four-dimensional) phase space for the variables $(\Lambda, \hat{\lambda}, \hat{z})$.

Let

$$\hat{\mathcal{H}}_G := \breve{\mathcal{H}}_{3\mathrm{b}} \circ \hat{\phi} = h_{\mathrm{Kep}} + \mu \hat{f}_G(\Lambda, \hat{\lambda}, \hat{z}) \tag{4.10}$$

denote the fully reduced system (where $\breve{\mathcal{H}}_{3\mathrm{b}}$ is as in Claim 4) on the phase space

$$\hat{\mathcal{M}}^8_G := \mathcal{A}_G \times \mathbb{T}^2 \times B^4_{\varrho_\star}. \tag{4.11}$$

We may assume that the function $\hat{N} + \check{N}$, where $\hat{N} := \breve{N} \circ \hat{\phi}$ and $\check{N} := \tilde{N} \circ \hat{\phi}$, is again in Birkhoff normal form of order $2m$. If not, proceeding as in [15, Proof of Proposition 5.1], one can find a ϵ^{2m+1}-close to the identity symplectic transformation $\check{\phi}$ such that $\hat{N}' + \check{N}$“ $:= (\hat{N} + \check{N}) \circ \check{\phi}$ is so. In the following statement, replace eventually $\hat{\phi}$, $\hat{N}$ and $\check{N}$ with, respectively, $\hat{\phi}' := \hat{\phi} \circ \check{\phi}$, $\hat{N}'$, $\check{N}'$.

Proposition 4. *The system* (4.10)–(4.11) *verifies*

$$(\hat{f}_G)_{\mathrm{av}} = \hat{N} + \check{N}$$

where $\hat{N} + \check{N}$ is in Birkhoff normal form of order $2m$, $|\check{N}| \leqslant \mathrm{const}\,\alpha^3$. Moreover, the first three orders of $\hat{N}$ are given by

$$\begin{aligned} \hat{N} := &-\frac{\bar{m}_1\bar{m}_2}{a_2} - \bar{m}_1\bar{m}_2\frac{a_1^2}{4a_2^3}\Bigg(\left(1 - 3\left(\frac{1}{\Lambda_1} - \frac{1}{\Lambda_2}\right)\hat{t}_2 - 3\left(\frac{2}{\Lambda_1} + \frac{1}{\Lambda_2}\right)\hat{t}_3\right) \\ &- \bar{m}_1\bar{m}_2\frac{a_1^2}{4a_2^3}\Bigg(-\frac{3}{2}\frac{\hat{t}_2^2}{\Lambda_1^2} + 9\frac{\hat{t}_2\hat{t}_3}{\Lambda_1^2} + 12\frac{\hat{t}_3^2}{\Lambda_1^2} \\ &-\frac{9}{2}\frac{\hat{t}_2^3}{\Lambda_1^2\Lambda_2} - \frac{105}{4}\frac{\hat{t}_2^3\hat{t}_3}{\Lambda_1^3} - \frac{315}{4}\frac{\hat{t}_2\hat{t}_3^2}{\Lambda_1^3} - \frac{105}{2}\frac{\hat{t}_3^3}{\Lambda_1^3}\Bigg)\left(1 + \mathrm{O}\left(\frac{\Lambda_1}{\Lambda_2}\right) + \mathrm{O}(\varrho^2)\right) \\ &+ \mathrm{O}(|t|^{7/2})\Bigg). \end{aligned} \tag{4.12}$$

Proof. The term $\hat{N}$ is easily computed from (1.18) and (4.9), which amounts to replace, in (1.18)

$$t_1 := \frac{\varrho^2}{2} - \hat{t}_2 - \hat{t}_3, \quad t_2 = \hat{t}_2, \quad t_3 = \hat{t}_3.$$

We then find (4.12). □

[50] As discussed in [16, Proposition 7.3] any step of Birkhoff normalization commutes with $\bar{\mathcal{R}}_g$ in (3.8), the the time-g flow of G in (3.9); equivalently, it preserves G.

4.3. Step 3: Averaging Fast Angles

In the next step we introduce, on a suitable phase space

$$\overline{\mathcal{M}}^8_G := \bar{D} \times \mathbb{T}^2 \times B^4_{\epsilon_1/4} \subset \hat{\mathcal{M}}^8_G, \tag{4.13}$$

(where $\hat{\mathcal{M}}^8_G$ is as in (4.11); $\epsilon_1 \leqslant \varrho_\star$ will be arbitrary) a new system

$$\overline{\mathcal{H}}_G := h_{\rm Kep}(\overline{\Lambda}) + \mu(\hat{N}(\overline{\Lambda}, \overline{z}) + \hat{N}_\star(\overline{\Lambda}, \overline{z})) + \mu \overline{f}_G(\overline{\Lambda}, \overline{\lambda}, \overline{z}) \tag{4.14}$$

where $\hat{N}$ is as in the previous sections, $\hat{N}_\star$ (as well as $\hat{N}$) depends only on $\overline{t}_1 = {\rm i}\overline{u}_1\overline{u}_1^\star$, $\overline{t}_2 = {\rm i}\overline{u}_2\overline{u}_2^\star$, $\overline{t}_3 = {\rm i}\overline{v}\overline{v}^\star$ and is suitably small and $\overline{f}_G$ is suitably small.

Lemma 5. *There exist positive numbers $\bar{M}$ ρ_0, s_0, depending only of $h_{\rm Kep}$ and $f_{\rm 3b}$ in (1.14) such that, for any given $m \in \mathbb{N}$, one can find $\gamma_\star$, α_*, $\mu_\star$, C (depending only on m, ϵ_0, s_0) such that for any μ, α, $\bar{\gamma} > 0$, $\tau > 2$, $\bar{K} > \frac{6}{s_0}$, verifying $0 < \alpha < \alpha_*$, $0 < \mu < \mu_\star$,*

$$\bar{\gamma} \geqslant \gamma_\star \max\{\sqrt{\mu}\bar{K}^{\tau+1},\ \sqrt[3]{\mu\epsilon_1}\bar{K}^{\tau+1}\}, \quad \bar{\rho} := \frac{\bar{\gamma}}{2\bar{M}\bar{K}^{\tau+1}} \leqslant \rho_0, \tag{4.15}$$

an open set $\bar{D} \subset \mathcal{A}_G$ with

$$\mathrm{meas}\left(\mathcal{A}_G \setminus \bar{D}\right) \leqslant C\bar{\gamma}\,\mathrm{meas}\,\mathcal{A}_G$$

defined by the following inequalities for the Keplerian frequencies $\omega_{\rm Kep} := \partial_{\overline{\Lambda}} h_{\rm Kep}$

$$|\omega_{\rm Kep} \cdot k| \geqslant \frac{\bar{\gamma}}{\bar{M}\bar{K}^\tau} \quad \forall k:\ 0 < |k|_1 \leqslant \bar{K}$$

such that for any positive number $\epsilon_1 \leqslant \varrho_\star$ a real-analytic transformation[51]

$$\overline{\phi}: \quad (\overline{\Lambda}, \overline{\lambda}, \overline{z}) \in \bar{D}_{\bar{\rho}/16} \times \mathbb{T}^2_{s_0/48} \times B^4_{\epsilon_1/4} \to (\Lambda, \hat{\lambda}, \hat{z}) \in (\mathcal{A}_G)_{\rho_0} \times \mathbb{T}^2_{s_0} \times B^4_{\varrho_\star}$$

exists, which is $\left\{\frac{\mu\bar{K}^{2(\tau+1)}}{\bar{\gamma}^2},\ \frac{\mu\epsilon_1\bar{K}^{3(\tau+1)}}{\bar{\gamma}^3}\right\}$ -close to the identity and lets the Hamiltonian (4.10)–(4.11) into $\overline{\mathcal{H}}_G := \hat{\mathcal{H}}_G \circ \overline{\phi}$ as in (4.14) with $\hat{N}$ as in Proposition 4, $\hat{N}_\star$ in Birkhoff normal form of order m, with Birkhoff invariants $\frac{\mu\bar{K}^{2\tau+1}}{\bar{\gamma}^2}$-close to 0 and

$$|\overline{f}_G| \leqslant C\mu\max\{e^{-\bar{K}s_0/6},\ \epsilon_1^{2m+1}\}. \tag{4.16}$$

The proof of Lemma 5 uses analogue techniques as the ones in [12, Theorem 1.4], therefore we shall only sketch it briefly, referring the reader to [12] for more details. It relies on Normal Form (Averaging[52]) Theory for properly-degenerate systems and the classical Birkhoff theory (see, *e.g.*, [23]). As for Normal form theory, we refer to the theory developed in [7] (see also [12]), which, in turn, generalizes ideas and techniques of [39] to the degenerate case. For information on Normal Form theory, see [5, 7, 12, 33, 39] and references therein.

Sketch of proof of Lemma 5. We shall describe only how to change the proof of [12, Theorem 1.4] in order to obtain the proof of Lemma 5. We refer, in particular, to [12, Steps 1–4 in the proof of Theorem 1.4]. First of all, choice, in [12, Steps 1–4 in the proof of Theorem 1.4],

$$\begin{aligned} &n_1 = 2, \quad n_2 = 2, \quad V = \mathcal{A}_G, \quad \kappa = \alpha^3, \quad \epsilon_0 := \varrho_\star, \quad H := \hat{\mathcal{H}}_G \\ &h = h_{\rm Kep}, \quad P_{\rm av} = \hat{N} + \check{N}, \quad P := \mu\hat{f}_G, \end{aligned}$$

[51] We refer to [39] for (now, standard) notations of the kind $\mathcal{A}_\rho$, or $\mathbb{T}^n_s$, where $\mathcal{A}$ is a subset of the reals and ρ, s are positive numbers.

[52] Sometimes distinction between "Normal Form" and "Averaging" Theory is made, depending on the strength of the remainder. For an exponentially small remainder, as in [7, 33, 39], "Normal Form" Theory is often used (after [39]); for a quadratically-small remainder, "Averaging" Theory is used, after [5]. Normal form Theory is obtained with suitably many steps of averaging.

$$I = (\Lambda_1, \Lambda_2), \quad \varphi := (\hat\lambda_1, \hat\lambda_2), \quad p := (\hat\eta_2, \hat p_1), \quad q := (\hat\xi_2, \hat q_1)$$
$$\Omega := \frac{3}{4}\bar m_1 \bar m_2 \frac{a_1^2}{a_2^3 \Lambda_1}\left(\frac{1}{\Lambda_1} - \frac{1}{\Lambda_2}, \frac{2}{\Lambda_1} + \frac{1}{\Lambda_2}\right) + \mathrm{O}\left(\frac{a_1^3}{a_2^4}\right).$$

Next, modify [12, Steps 1–4 in the proof of Theorem 1.4] as follows.

In [12, Step 1], neglect [12, Eq. (36)], so as to "leave $\bar K$ free" and hence replace $\log\epsilon^{-1}$ with $\frac{s_0}{30}\bar K$ wherever it appears (i.e., [12, Eqs. (41), (42), (43)]). Neglect the second line in [12, Eq. (40)]. At the end of [12, Step 1, 2, 3, 4], in the definition of $\bar H$, $\tilde H$, $\check H$, $\breve H$, respectively, replace ϵ^5 with $e^{-\bar K s_0/6}$. At the beginning of [12, Step 2, 3, 4], in the definition of, respectively, $\tilde v$, $\hat v$, $\check v$, replace ϵ with $\epsilon_1 \leqslant \epsilon_0$. In [12, Step 2] replace "$\bar N$ also has a $\mu(\log\epsilon^{-1})^{2\tau+1}\bar\gamma^{-2}$-close-to-0 elliptic equilibrium point" with "$\bar N$ also has a $\mu\bar K^{2\tau+1}\bar\gamma^{-2}$-close-to-0 elliptic equilibrium point". Replace[53] [12, Eqs. (43), (44), (45), (46)] with, respectively: (43)': $|\bar p - \tilde p|$, $|\bar q - \tilde q| \leqslant C\frac{\mu\bar K^{2\tau+1}}{\bar\gamma^2}$, $|\bar\varphi - \tilde\varphi| \leqslant C\max\left\{\frac{\epsilon_1^2\bar K^{\tau+1}}{\bar\gamma}, \frac{\mu\epsilon_1\bar K^{3\tau+2}}{\bar\gamma^3},\right\}$; (44)': $|\tilde p - \hat p|$, $|\tilde q - \hat q| \leqslant C\max\{\frac{\mu\epsilon_1\bar K^{2\tau+1}}{\bar\gamma^2}\}$, $|\tilde\varphi - \hat\varphi| \leqslant C\max\{\frac{\mu\epsilon_1^2\bar K^{3\tau+2}}{\bar\gamma^3}\}$; (45)': $|\hat\Omega - \Omega|$, $|\hat R| \leqslant C\frac{\mu\bar K^{2\tau+1}}{\bar\gamma^2}$ and (46)': $|\hat p - \check p|$, $|\hat q - \check q| \leqslant C\frac{\mu\epsilon_1^2\bar K^{2\tau+1}}{\bar\gamma^2}$, $|\hat\varphi - \check\varphi| \leqslant C\frac{\mu\epsilon_1^3\bar K^{3\tau+2}}{\bar\gamma^3}$, by suitably modifying the proofs below. Moreover replace Equation just before [12, Eq. 45] with[54] $\hat{\mathcal N}(I, p, q) := \tilde N \circ \hat\phi = \hat N + \hat R$ (where $\hat N$ is as in (4.12)) and replace the last line in [12, Step 4] with "where $\breve N(\breve I, \breve r)$ is a polynomial of degree m in $(I_{n_1+1}, \cdots, I_{n_2})$". Lemma 5 follows, with $\hat N_\star := \breve N - \hat N$ and $\hat f_G := \mu(e^{-\bar K s_0/6}\breve P + \mathrm{O}(\epsilon_1^{2m+1}))$. □

We then apply Lemma 5 to the system (4.10)–(4.11) with $\bar K$ as in (4.3), with ϵ, α replaced by ϵ_1, α_* and

$$\bar\gamma = \gamma_\star \sqrt[4]{\mu}\bar K^{\tau+1}, \tag{4.17}$$

where $\gamma_\star$ is as in (4.15). By the thesis of Lemma 5, we conjugate $\hat{\mathcal H}_G$ in (4.10)–(4.11) to $\overline{\mathcal H}_G$ in (4.13)–(4.14), with $\overline f_G$ satisfying (4.16), via a symplectic transformation which, by the choice of $\bar\gamma$ in (4.17), is $\mu^{1/12}$-close to the identity.

4.4. Step 4: Nehorošev Theory

We apply Nehorošev Theory (i.e., Theorem 11) to the system $\overline{\mathcal H}_G$ in (4.13)–(4.14), in the *planar* case, i.e., with $\hat t_4 = 0$.

For information on the tools that are used, compare [32–34] and Appendix D.

In applying Theorem 11, we shall take

$$n_1 = 3,\ n_2 = 0,\ V = \bar D,\ B^4 := B^4_{\epsilon_1/8},\ \rho := \min\{\bar\rho/16, s_0/48, \epsilon_1/8\}$$
$$H_0(\overline\Lambda_1, \overline\Lambda_2, \overline t_1) := h_{\rm Kep}(\overline\Lambda_1, \overline\Lambda_2) + \mu(\hat N + \hat N_\star)(\overline\Lambda_1, \overline\Lambda_2, \overline t_1),\ P := \mu\overline f_G \tag{4.18}$$

where $\bar\rho$, s_0 and ϵ_1 are as in Lemma 5

We have to check[55] steepness of $H_0(\overline\Lambda_1, \overline\Lambda_2, \overline t_1)$ and the smallness condition (D.1) of P. The first check is provided by the following claim.

Claim 6. The function H_0 in (4.18) is $(g, m, C_1, C_2, \mathfrak a_1, \mathfrak a_2, \delta_1, \delta_2)$-*steep*, with

$$g = \hat g, \quad m = \hat m, \quad \mathfrak a_i = \hat{\mathfrak a}_i, \quad \delta_i = \min\{\sqrt{\alpha_*}, \epsilon_1^2\}\hat\delta_i, \quad C_i = \mu\alpha_*^2\hat C_i \tag{4.19}$$

where $(\hat g, \hat m, \hat C_1, \hat C_2, \hat{\mathfrak a}_1, \hat{\mathfrak a}_2, \hat\delta_1, \hat\delta_2)$ suitable numbers independent of α_*, μ, ϵ_1.

---

[53] In [12, Eq (45)] $\mu\epsilon$ should be replaced by μ. This does not affect the thesis of [12, Theorem 1.4].

[54] The symbol $\hat N$ used in [12] is here replaced with $\hat{\mathcal N}$, to avoid confusions with (4.12).

[55] Recall that, for $n_2 = 0$, as it is in our case, condition (D.2) is void; see Appendix D.

Proof. We take, in (4.12), $\hat{t}_3 = 0$. The system has three degrees of freedom. We firstly prove *steepness* for a suitable "rescaled" system associated to F. That is, if $\hat{N}_0 := -\frac{\bar{m}_1\bar{m}_2}{a_2}$ is as in (4.12) and $\hat{N}_1 := \hat{N} - \hat{N}_0$ we consider the system

$$\begin{aligned}\mathrm{F}_{\mathrm{resc}}(\hat{\Lambda}_1, \hat{\Lambda}_2, \hat{t}_2) := \bar{m}_1^2\bar{m}_0\alpha_*\Big(&h^{(1)}_{\mathrm{Kep}}(\bar{m}_1\sqrt{\bar{m}_0\alpha_*}\hat{\Lambda}_1) + \beta_2 h^{(2)}_{\mathrm{Kep}}(\hat{\Lambda}_2) + \mu\hat{N}_0(\Lambda_2)\\ &+ \mu\beta_3(\hat{N}_1 + \hat{N}_\star)(\bar{m}_1\sqrt{\bar{m}_0\alpha_*}\hat{\Lambda}_1, \hat{\Lambda}_2, \epsilon_1^2\hat{t}_2)\Big)\end{aligned} \tag{4.20}$$

with α_*, ϵ_1 as in Lemma 5

$$\beta_2 := \alpha_*^{-3/2}, \quad \beta_3 := \mu^{-1}\alpha_*^{-3}\epsilon_1^{-2}. \tag{4.21}$$

We check that $\mathrm{F}_{\mathrm{resc}}$ is *steep* by verifying the three-jet condition: See Appendix D. The three-jet condition (D.5) for the system (4.20) is (neglecting mixed terms, which are smaller)

$$\begin{cases} \eta_1 + \beta_2\alpha_*^{3/2}(\dfrac{\hat{a}_1}{\hat{a}_2})^{3/2}\eta_2 + \beta_3\alpha_*^3\epsilon_1^2\mu\dfrac{3}{4}\dfrac{\bar{m}_2}{\bar{m}_0}(\dfrac{\hat{a}_1}{\hat{a}_2})^3\eta_3 = 0\\ \eta_1^2 + \dfrac{\bar{m}_1}{\bar{m}_2}\beta_2\alpha_*^2(\dfrac{\hat{a}_1}{\hat{a}_2})^2\eta_2^2 - \beta_3\alpha_*^3\mu\epsilon_1^4\dfrac{1}{4}\dfrac{\bar{m}_2}{\bar{m}_0}(\dfrac{\hat{a}_1}{\hat{a}_2})^3\eta_3^2 = 0\\ \eta_1^3 + (\dfrac{\bar{m}_1}{\bar{m}_2})^2\beta_2\alpha_*^{5/2}(\dfrac{\hat{a}_1}{\hat{a}_2})^{5/2}\eta_2^3 + \beta_3\alpha_*^{7/2}\mu\epsilon_1^6\dfrac{9}{16}\dfrac{\bar{m}_1}{\bar{m}_0}(\dfrac{\hat{a}_1}{\hat{a}_2})^{7/2}\eta_3^3 = 0 \end{cases} \tag{4.22}$$

where we have used $m_i = \bar{m}_i + \mathrm{O}(\mu)$, $M_i = \bar{m}_0 + \mathrm{O}(\mu)$ and neglected higher order terms going to zero with μ, ϵ_1, α_*. If we eliminate η_1 from the first and the second equation and from the first and the third equation, we obtain a homogeneous system of two equations in (η_2, η_3) that, in view of (4.21), generically, the has only solution $\eta_2 = \eta_3 = 0$, implying that also $\eta_1 = 0$. This implies that the function $\mathrm{F}_{\mathrm{resc}}$ (4.20) is $(2\hat{g}, \hat{m}/2, \hat{C}_1, \hat{C}_2, \hat{\mathfrak{a}}_1, \hat{\mathfrak{a}}_2, \hat{\delta}_1, \hat{\delta}_2)$-*steep* with suitable values of $(\hat{g}, \hat{m}, \hat{C}_1, \hat{C}_2, \hat{\mathfrak{a}}_1, \hat{\mathfrak{a}}_2, \hat{\delta}_1, \hat{\delta}_2)$ which are of order 1 in μ, α_*, ϵ_1. This readily implies that F in (4.18) is $(g, m, C_1, C_2, \mathfrak{a}_1, \mathfrak{a}_2, \delta_1, \delta_2)$-*steep*, with $(g, m, C_1, C_2, \mathfrak{a}_1, \mathfrak{a}_2, \delta_1, \delta_2)$ as in (4.19). □

Remark 3. In the case of the *spatial* three-body problem, instead of (4.22), we would have

$$\begin{cases} \eta_1 + \beta_2\alpha_*^{3/2}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^{3/2}\eta_2 + \beta_3\epsilon_1^2\alpha_*^3\mu\dfrac{3}{4}\dfrac{\bar{m}_2}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^3\eta_3 + \beta_3\epsilon_1^2\alpha_*^3\mu\dfrac{3}{2}\dfrac{\bar{m}_2}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^3\eta_4 = 0\\ \eta_1^2 + \dfrac{\bar{m}_1}{\bar{m}_2}\beta_2\alpha_*^2\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^2\eta_2^2 - \beta_3\epsilon_1^4\alpha_*^3\mu\dfrac{1}{4}\dfrac{\bar{m}_2}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^3\eta_3^2 + \beta_3\epsilon_1^4\alpha_*^3\mu\dfrac{3}{2}\dfrac{\bar{m}_2}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^3\eta_3\eta_4\\ \qquad + 2\beta_3\epsilon_1^4\alpha_*^3\mu\dfrac{\bar{m}_2}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^3\eta_4^2 = 0\\ \eta_1^3 + \left(\dfrac{\bar{m}_1}{\bar{m}_2}\right)^2\beta_2\alpha_*^{5/2}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^{5/2}\eta_2^3 + \beta_3\epsilon_1^6\alpha_*^{7/2}\mu\epsilon_1^6\dfrac{9}{16}\dfrac{\bar{m}_1}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^{7/2}\eta_3^3 + \beta_3\epsilon_1^6\alpha_*^3\mu\dfrac{105}{32}\dfrac{\bar{m}_2}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^3\eta_3^2\eta_4\\ \qquad + \beta_3\epsilon_1^6\alpha_*^3\mu\dfrac{315}{32}\dfrac{\bar{m}_2}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^3\eta_3\eta_4^2 + \beta_3\epsilon_1^6\alpha_*^3\mu\dfrac{105}{16}\dfrac{\bar{m}_2}{\bar{m}_0}\left(\dfrac{\hat{a}_1}{\hat{a}_2}\right)^3\eta_4^3 = 0 \end{cases}$$

It is not clear to the author if this system exhibits non-trivial solutions, so the analysis of this case is deferred to a subsequent paper.

We can now complete the

Proof of Theorem 7. It remains only to check condition (D.1), with P, ρ as in (4.18). In view of (D.3), (D.4), (4.19) and the choice of $\bar{\gamma}$ in (4.17), we have $\rho \geqslant \rho_\star\min\{\epsilon_1, \sqrt[4]{\mu}\}$ and hence

$$\begin{aligned} M_\star &\geqslant \frac{\tilde{c}}{\rho}\min\{(\mu\alpha_*^2)^q, \rho^q\} \geqslant \frac{\tilde{c}}{\rho}\min\{(\mu\alpha_*^2)^q, \epsilon_1^q, \ \bar{\rho}^q\}\\ &\geqslant \frac{\tilde{c}}{\rho}\min\left\{(\mu\alpha_*^2)^q, \epsilon_1^q, \ \left(\frac{\bar{\gamma}_\star}{2\bar{M}}\right)^q\mu^{q/4}\right\} \geqslant \frac{c_\star}{\rho}\min\left\{(\mu\alpha_*^2)^q, \epsilon_1^q\right\} \end{aligned}$$

for some $q > 1 > c_\star$ depending only on n_1, n_2, $\mathfrak{a}_1$, $\mathfrak{a}_2$. Noticing that (4.16) and Cauchy inequality imply

$$M := \sup |\partial P| = \mu \sup |\partial \overline{f}_G| \leqslant \tilde{C} \frac{\mu}{\rho} \max\{e^{-\bar{K} s_0/6},\ \epsilon_1^{2m+1}\}$$

one sees that condition (D.1) is met, provided one previously fixes, in Lemma 5, $2m+1 \geqslant q$, $\bar{K}$ as in (4.3), with a suitable $\bar{K}_\star$ and takes $\epsilon_1 < (\frac{c_\star}{\tilde{C}})^{1/(2m+1)}(\mu\alpha_*^2)^{q/(2m+1)}$. The thesis then follows, with α, ϵ replaced by α_*, ϵ_1 and $\phi := \breve{\phi} \circ \hat{\phi} \circ \overline{\phi} \circ \hat{\phi}^{-1}$. □

APPENDIX A. THE FUNDAMENTAL THEOREM AND ANOTHER RESULT IN ARNOLD'S 1963 PAPER

Here we recall two theorems in [5]. The former is named the "Fundamental Theorem" in [5] and is as follows.

Recall the definition of $\mathcal{P}_{\epsilon_0}$ in (3.16).

Theorem 8 (V. I. Arnold, [5, p. 143]). *Consider a Hamiltonian of the form*

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

which is real-analytic on $\mathcal{P}_{\epsilon_0}$ where $V \subset \mathbb{R}^{n_1}$ is open and connected, $B^{2n_2}_{\epsilon_0} \subset \mathbb{R}^{2n_2}$ is a ball of radius ϵ_0 around the origin and $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$. Assume that

(i) *$I \in V \to \partial_I H_0$ is a diffeomorphism;*

(ii) *$P_{\rm av}$ is in Birkhoff normal form*[56] *of order 6;*

(iii) *the matrix β of the "second order Birkhoff invariants": is not singular: $|\det \beta| \neq 0$ on V.*

Then, there exists $\epsilon_0 > 0$ such that, for

$$0 < \epsilon < \epsilon_0, \qquad 0 < \mu < \epsilon^8, \tag{A.1}$$

one can find a set $\mathcal{K}_{\mu,\epsilon} \subset \mathcal{P}_\epsilon \subset \mathcal{P}_{\epsilon_0}$, with

$$\operatorname{meas} \mathcal{K}_{\mu,\epsilon} \geqslant (1 - \epsilon^{16(n_1+n_2)}) \operatorname{meas} \mathcal{P}_\epsilon$$

formed by the union of H-invariant (n_1+n_2)-dimensional tori on which the H-motion is analytically conjugated to linear Diophantine[57] *quasi-periodic motions.*

The latter is less general, but used in [5] to prove Theorem 2.

Theorem 9 (V. I. Arnold, [5, Ch. I, Section 8, p. 107]). *Under the same assumptions as in Theorem 8, but replacing* (ii), (iii) *and* (A.1) *with*

(ii)′ *$P_{\rm av}$ has the form*

$$P_{\rm av}(I, p, q) = N(I, J) + \tilde{N}(I, p, q) \text{ where } J = \left(\frac{p_1^2+q_1^2}{2}, \cdots, \frac{p_{n_2}^2+q_{n_2}^2}{2}\right) \text{ and } \tilde{N} = {\rm o}(\mu);$$

(iii)′ *the Hessian $\partial^2_{(I,J)} N$ is non-singular: $\det \partial^2_{(I,J)} N \neq 0$ on $V \times B^{2n_2}_{\epsilon_0}$*

and condition (A.1) *with condition*

$$|\mu| < \mu_*$$

one can find a set $\mathcal{K}_\mu \subset \mathcal{P}_{\epsilon_0}$, with

$$\operatorname{meas} \mathcal{K}_\mu \geqslant (1 - \mu^a) \operatorname{meas} \mathcal{P}_{\epsilon_0}$$

(where a decreases with n_1+n_2) having the same properties as the set $\mathcal{K}_\epsilon$ of Theorem 8.

---

[56] We refer to [23] for information on Birkhoff Theory.

[57] *I.e.*, the flow is conjugated to the Kronecker flow $\theta \in \mathbb{T}^{n_1+n_2} \to \theta + \omega t \in \mathbb{T}^{n_1+n_2}$, with $\omega \in \mathbb{R}^{n_1+n_2}$ satisfying $|\omega \cdot k| \geqslant \gamma |k|_1^{-\tau}$ for all $k \neq 0$, for suitable γ, $\tau > 0$.

APPENDIX B. PROOF OF (1.22), (1.23) AND (4.5)

The formulae in (1.22) and (1.23) are a consequence of Proposition 1 and the formulae developed in [16, 38] (see, *e.g.*, [16, Appendix A]). Indeed, from such papers there results that, if

$$\begin{aligned}
&a_i := \frac{1}{M_i}\left(\frac{\Lambda_i}{m_i}\right)^2, \quad e_i^2 = \frac{\eta_i^2+\xi_i^2}{\Lambda_i} - \left(\frac{\eta_i^2+\xi_i^2}{2\Lambda_i}\right)^2 =: 2\mathrm{i}u_i u_i^\star \bar{e}_i^2\\
&\zeta_i:\ \zeta_i - e_i\sin\zeta_i = \lambda_i + \arg(\eta_i,\xi_i)\\
&\bar{\mathfrak{c}} := \frac{2\Lambda_1+2\Lambda_2-2\mathrm{i}u_1u_1^\star-2\mathrm{i}u_2u_2^\star-\mathrm{i}vv^\star}{4(\Lambda_1-\mathrm{i}u_1u_1^\star)(\Lambda_2-\mathrm{i}u_2u_2^\star)}, \quad \bar{\mathfrak{s}} := \sqrt{2\bar{\mathfrak{c}}(1-\mathrm{i}vv^\star\bar{\mathfrak{c}})}
\end{aligned}$$

then, the expressions of $\mathrm{C}^{(2)}\cdot x^{(1)}$, $|\mathrm{C}^{(2)}|$, $\mathrm{r}_1 = |x^{(1)}|$ and $\mathrm{r}_2 = |x^{(2)}|$ in terms of the RPS variables are

$$\begin{aligned}
&\mathrm{C}^{(2)}\cdot x^{(1)} = \Big((\hat{u}_1v^\star - \hat{u}_1^\star v)x^{(1)} + \mathrm{i}(\hat{u}_1v^\star+\hat{u}_1^\star v)x^{(2)}\Big)\bar{\mathfrak{s}}|\mathrm{C}^{(2)}|\\
&|\mathrm{C}^{(2)}| = \Lambda_2 - \mathrm{i}u_2u_2^\star, \quad |x^{(i)}| = a_i(1-e_i\cos\zeta_i)\\
&x_1^{(1)} := \frac{1}{M_1}\left(\frac{\Lambda_1}{m_1}\right)^2(\cos\zeta_1 - e_1), \quad x_2^{(1)} := \frac{1}{M_1}\left(\frac{\Lambda_1}{m_1}\right)^2\sqrt{1-e_1^2}\sin\zeta_1
\end{aligned}$$

with

$$\hat{u}_i := \frac{u_i}{\sqrt{\mathrm{i}u_iu_i^\star}} = \frac{\eta_i - \mathrm{i}\xi_i}{\sqrt{2}\sqrt{\eta_i^2+\xi_i^2}}, \quad \hat{u}_i^\star := \frac{u_i^\star}{\sqrt{\mathrm{i}u_iu_i^\star}} = \frac{\eta_i+\mathrm{i}\xi_i}{\sqrt{2}\mathrm{i}\sqrt{\eta_i^2+\xi_i^2}}.$$

Then we have

$$\begin{aligned}
(\mathrm{C}^{(2)}\cdot x^{(1)})^2 = \Big(&\mathrm{i}vv^\star\left((x_1^{(1)})^2+(x_2^{(1)})^2\right) + \left((\hat{u}_1^\star)^2v^2+(v^\star)^2\hat{u}_1^2\right)\left((x_1^{(1)})^2-(x_2^{(1)})^2\right)\\
&+2\mathrm{i}\left((\hat{u}_1^\star)^2v^2-(v^\star)^2\hat{u}_1^2\right)x_1^{(1)}x_2^{(1)}\Big)\bar{\mathfrak{s}}^2|\mathrm{C}^{(2)}|^2
\end{aligned}$$

and hence, taking the λ_1-average (recall the relation $d\lambda_2 = (1-e_2\cos\zeta_2)d\zeta_2$)

$$\frac{1}{2\pi}\int_{\mathbb{T}}(\mathrm{C}^{(2)}\cdot x^{(1)})^2 d\lambda_1 = \Big(\mathrm{i}vv^\star a_1^2(1+\frac{3}{2}e_1^2) + \frac{5}{2}\left((u_1^\star)^2v^2+(v^\star)^2u_1^2\right)\frac{a_1^2e_1^2}{2\mathrm{i}u_1u_1^\star}\Big)\bar{\mathfrak{s}}^2|\mathrm{C}^{(2)}|^2. \tag{B.1}$$

Here, we have used

$$\begin{aligned}
&\frac{1}{2\pi}\int_{\mathbb{T}}\left((x_1^{(1)})^2+(x_2^{(1)})^2\right)d\lambda_1 = \frac{1}{2\pi}\int_{\mathbb{T}}\mathrm{r}_1^2 = \frac{1}{2\pi}\int_{\mathbb{T}}a_1^2(1-e_1\cos\zeta_1)^3d\zeta_1\\
&\quad = a_1^2(1+\frac{3}{2}e_1^2)\\
&\frac{1}{2\pi}\int_{\mathbb{T}}\left((x_1^{(1)})^2-(x_2^{(1)})^2\right)d\lambda_1 = \frac{1}{2\pi}\int_{\mathbb{T}}d\zeta_1\big(a_1^2(\cos 2\zeta_1+e_1^2\\
&\quad +e_1^2\sin^2\zeta_1 - 2e_1\cos\zeta_1)(1-e_1\cos\zeta_1)\big) = \frac{5}{2}a_1^2e_1^2\\
&\frac{1}{2\pi}\int_{\mathbb{T}}x_1^{(1)}x_2^{(1)}d\lambda_1 = \frac{1}{2\pi}\int_{\mathbb{T}}a_1^2\sqrt{1-e_1^2}(\cos\zeta_1-e_1)(1-e_1\cos\zeta_1)\sin\zeta_1 d\zeta_1 = 0.
\end{aligned}$$

Note that (B.1) implies (1.23). In turn, (1.22) for $n = 2$ and (1.23) give (4.5), with $\mathrm{f} := \dfrac{\frac{1}{2\pi}\int_{\mathbb{T}}\frac{d\zeta}{(1-e_2\cos\zeta)}}{(1-\frac{\eta_2^2+\xi_2^2}{2\Lambda_2})^2} = \dfrac{1}{(1-\frac{\eta_2^2+\xi_2^2}{2\Lambda_2})^3}$ (recall the relation $d\lambda_2 = (1-e_2\cos\zeta_2)d\zeta_2$ and use $\frac{1}{2\pi}\int_{\mathbb{T}}\frac{d\zeta}{(1-e_2\cos\zeta)} = \frac{1}{\sqrt{1-e_2^2}} = \frac{1}{1-\frac{\eta_2^2+\xi_2^2}{2\Lambda_2}}$).

APPENDIX C. PROOF OF THEOREM 6

Theorem 6 is an easy consequence[58)] of the following more technical statement.

Theorem 10. *Under the same notations and assumptions as in Theorem 6, one can find γ_*, C_* such that, for any ϵ_0, one can find positive numbers $\epsilon_1 < \epsilon_0$, μ_* and α_* such that, for any α, μ, γ_1, $\bar\gamma_2$, $\bar\gamma$ verifying*

$$|\alpha| < \alpha_*, \quad |\mu| < \mu_*, \quad \mu\bar\gamma_2 \leqslant \gamma_1$$

and

$$\begin{cases} \gamma_* \sqrt[4]{\mu}(\log\alpha^{-1})^{\tau+1} \leqslant \bar\gamma \leqslant \gamma_* \\ \gamma_* \max\left\{\alpha^2, \ \frac{\sqrt{\mu}(\log\alpha^{-1})^{\tau+1}}{\bar\gamma}\right\} < \gamma_1 < \gamma_* \\ \gamma_* \max\left\{\alpha^2(\log({\gamma_1}^2/\alpha^3))^{\tau_*+1},\right. \\ \left. \ \sqrt{\mu}(\log\alpha^{-1})^{\tau+1}\bar\gamma^{-1}\left(\log\left(\frac{{\gamma_1}^2}{\mu(\log\alpha^{-1})^{2\tau+1}\bar\gamma^{-2}}\right)\right)^{\tau+1}\right\} < \bar\gamma_2 < \gamma_*\epsilon_0^2, \end{cases} \tag{C.1}$$

where $\tau > n := n_1 + n_2$, then, one can find a set $\mathcal{K} \subset \mathcal{P}$ formed by the union of H-invariant n-dimensional tori, on which the H-motion is analytically conjugated to linear Diophantine quasi-periodic motions. The set $\mathcal{K}$ is of positive measure and satisfies

$$\operatorname{meas}\mathcal{K} > \left[1 - C\left(\bar\gamma + \gamma_1 + \frac{\bar\gamma_2}{\epsilon_0^2} + \alpha^{n_2}\right)\right]\operatorname{meas}\mathcal{P}_{\epsilon_1}.$$

Furthermore, the flow on each H-invariant torus in $\mathcal{K}$ is analytically conjugated to a translation $\psi \in \mathbb{T}^n \to \psi + \omega t \in \mathbb{T}^n$ with Diophantine frequencies.

This result is a slight modification of [12, Theorem 1.4] (which, in turn, had been obtained in [38]). Then here we briefly sketch its proof, describing only the necessary changes with respect to [12, Proof of Theorem 1.4] and referring the reader to that paper for more details.

To proceed, we need to recall

– the definition of "two velocities" Diophantine vector[59)] in [12, Eq. (19)];

– the functional setting and notations described at the beginning of [12, Section 2];

– the "averaging (iterative) Theorem" [12, Lemma A.1];

– the "two-scale KAM Theorem" [12, Proposition 3].

Sketch of proof of Theorem 10. Let ρ_0, s_0, ϵ_0 (possibly with a smaller value of ϵ_0) be positive numbers such that H in (3.17) has analytic extension on the complex set

$$\mathcal{P}_{\rho_0,s_0,\epsilon_0} = V_{\rho_0} \times \mathbb{T}^{n_1}_{s_0} \times B^{2n_2}_{\epsilon_0}.$$

Take three numbers $\bar\gamma$, γ_1, $\gamma_2 = \mu\bar\gamma_2$ verifying (C.1) and $\mu\bar\gamma_2 < \gamma_1$, where γ_* is some large number, depending only on n_1, n_2, to be chosen below.

As in [12, Proof of Theorem 1.4, Step 1], start with removing, in H, the dependence on φ up to high orders. But, at difference with [12, Proof of Theorem 1.4, Step 1], apply [12, Lemma A.1]

---

58) To obtain Theorem 6 from Theorem 10, it is sufficient to choose

$$\bar\gamma = \gamma_* \sqrt[4]{\mu}\log(\alpha^{-1})^{\tau+1}, \quad \gamma_1 = \gamma_2 = \gamma_*^2 \max\{\alpha^2, \ \sqrt[4]{\mu}\} < \gamma_*\epsilon_0^2, \quad C_* := \frac{C}{\epsilon_0^2}.$$

59) This is a suitable generalization of the standard definition of Diophantine numbers, introduced in [5].

(instead of [12, Proposition 1]), with $\ell_1 = n_1$, $\ell_2 = 0$, $m = n_2$ $h = H_0$, $g \equiv 0$, $f = \mu P$, $B = B' = \{0\}$, $r_p = r_q = \epsilon_0$, $s = s_0$, $\rho_p = \rho_q = \epsilon_0/3$, $\sigma = s_0/3$, $\Lambda = \{0\}$,

$$e^{-\bar{K}s_0/3} := \kappa \qquad \text{i.e.}, \qquad \bar{K} = \frac{3}{s_0}\log\kappa^{-1}, \tag{C.2}$$

$A = \bar{D}$, $r = \bar{\rho}$, $\rho = \bar{\rho}/3$, where $\bar{D}$, $\bar{\rho}$ are defined as in [12, (37)] By [12, (38)], and the choice of $\bar{\gamma}$, the following standard measure estimate holds

$$\operatorname{meas}\left(V \setminus \bar{D}\right) \leqslant C\gamma_*\sqrt{\mu}(\log\kappa^{-1})^{\tau+1}\operatorname{meas} V$$

where C depends on the C^1-norm of H_0. Proceeding as [12, (39)] and the immediately following formula, one sees that the "non-resonance" condition [12, (64)] on $\bar{D}_{\bar{\rho}}$ and the "smallness" condition [12, (65)] are then verified, provided μ is chosen small enough, because of the choice of $\bar{\gamma}$ and γ_*. By the thesis of [12, Lemma A.1], we find a real-analytic symplectomorphism

$$\bar{\phi}: \ (\bar{I}, \bar{\varphi}, \bar{p}, \bar{q}) \in W_{(\bar{\rho},\epsilon_0)/3, s_0/3} \to (I, \varphi, p, q) \in W_{v_0, s_0}$$

where $W_{v_0,s_0} := \bar{D}_{\rho_0} \times \mathbb{T}^{n_1}_{s_0} \times B_{\epsilon_0}$ $(v_0 = (\rho_0, \epsilon_0))$, and, by the choice of $\bar{K}$ in (C.2), H is transformed into[60)]

$$\begin{aligned}\bar{H} := H \circ \bar{\phi} &= h + \mu P_{\text{av}} + \mu\bar{P}\\ &= h + \mu N + \mu\tilde{N} + \mu\bar{P}\end{aligned} \tag{C.3}$$

where $P_{\text{av}} = N + \tilde{N}$ corresponds to g_+ of [12, Lemma A.1], $\bar{P}$ corresponds to f_+ and hence, by the choice of $\bar{K}$ in (C.2), the assumption on $\tilde{N}$ and the thesis [12, (68)] of [12, Lemma A.1], one has that the new perturbation $\mu\tilde{N} + \mu\bar{P}$ verifies

$$\begin{aligned}\|\mu\tilde{N} + \mu\bar{P}\|_{v_0/3, s_0/3} &\leqslant C\mu\max\left\{\frac{\bar{K}^{2\tau+1}}{\bar{\gamma}^2}\mu,\ e^{-\bar{K}s_0/3},\ \kappa\right\}\\ &\leqslant C\mu\max\left\{\frac{\bar{K}^{2\tau+1}}{\bar{\gamma}^2}\mu,\ \kappa\right\}.\end{aligned} \tag{C.4}$$

In view of [12, (69)], the transformation $\bar{\phi}$ verifies

$$|I - \bar{I}|,\ |p - \bar{p}|,\ |q - \bar{q}| \leqslant C\frac{\mu(\log\kappa^{-1})^{\tau}}{\bar{\gamma}}, \quad |\varphi - \bar{\varphi}| \leqslant C\frac{\mu(\log\kappa^{-1})^{2\tau+1}}{\bar{\gamma}^2}.$$

Continue as in [12, Proof of Theorem 1.4, Step 5], but replacing the set in [12, (47)]. with the set

$$\mathcal{A} := \left\{J \in \mathbb{R}^{n_2} : \rho_1 < J_i < \epsilon_0^2/9, \quad 1 \leqslant i \leqslant n_2\right\} \tag{C.5}$$

where $\rho_1 < \epsilon_0^2/9$ will be fixed in the next step, on so as to maximize the measure of preserved tori. Next define $\mathcal{D}$ as in [12, (48)] (but with $\mathcal{A}$ as in (C.5)) and

$$\rho := \min\{\rho_1,\ \bar{\rho}/3\}, \quad s := s_0/3. \tag{C.6}$$

Introduce the change of variables

$$(J, \psi) = \big((J_1, J_2), (\psi_1, \psi_2)\big) \in \mathcal{D}_\rho \times \mathbb{T}^{n_1+n_2}_s \to (\bar{I}, \bar{\varphi}, \bar{p}, \bar{q})$$

defined as in [12, (49)], but replacing "checks" with "bars", This lets the Hamiltonian (C.3) into

$$H(J, \psi) = H_0(J_1) + \mu N(J) + \mu(\bar{P} + \tilde{N}), \quad (J, \psi) \in \mathcal{D}_\rho \times \mathbb{T}^{n_1+n_2}_s.$$

Next, analogously to [12, Proof of Theorem 1.4, Step 6], construct the Kolmogorov set and estimate its measure via [12, Proposition 3].

[60)] $\Pi_0 T_{\bar{K}} P = P_{\text{av}} = \dfrac{1}{(2\pi)^n}\displaystyle\int_{\mathbb{T}^n} P d\varphi.$

To this end, fix γ_1 and $\gamma_2 = \mu\bar{\gamma}_2$, with γ_1, $\bar{\gamma}_2$ satisfying $\mu\bar{\gamma}_2 \leqslant \gamma_1$ and (C.1). Let ρ_1 in(C.5)–(C.6) be chosen so that

$$\rho_1 = \check{c}_1 \max\left\{\sqrt{\kappa},\ \frac{\sqrt{\mu}(\log\kappa^{-1})^{\tau+1/2}}{\bar{\gamma}}\right\}$$

with $\check{c}_1$ some large number depending only on n_1, n_2 to be fixed below. Note that the needed condition $\rho_1 < \epsilon_0^2/9$ (compare the previous step; Eq. (C.5)) is satisfied for $\kappa < (\epsilon_0/(3\sqrt{\check{c}_1}))^4$ and[61] $\mu < \gamma_\star^4(\epsilon_0/(3\sqrt{\check{c}_1}))^8$. The assumption that the frequency map $\omega := \partial(H_0(J_1) + \mu N(J))$ is a diffeomorphism of $\mathcal{D}_\rho$ is trivially satisfied. Moreover, the numbers M, $\hat{M}$, $\cdots$, $\bar{M}_2$ involved in [12, Proposition 3] may be chosen as in [12, Proof of Theorem 1.4, Step 6], apart for E, which is chosen as[62]

$$E = C\max\{\mu\kappa,\ \bar{K}^{2\tau+1}\mu^2\bar{\gamma}^{-2}\}.$$

Then, we can take L as in [12, Proof of Theorem 1.4, Step 6], while

$$K = C\log{(E/(\mu{\gamma_1}^2))^{-1}}$$

and

$$\hat{\rho} = c\,\min\left\{\frac{\gamma_1}{(\log{(E/(\mu{\gamma_1}^2))^{-1}})^{\tau+1}}, \frac{\bar{\gamma}_2}{(\log{(E/(\mu{\gamma_1}^2))^{-1}})^{\tau+1}}, \frac{\bar{\gamma}}{(\log\kappa^{-1})^{\bar{\tau}+1}}, \rho_1,\ \rho_0\right\}.$$

To check the "KAM-smallness condition" [12, (32)], we divide the two cases $E = C\mu\kappa$ or $E = C\bar{K}^{2\tau+1}\mu^2\bar{\gamma}^{-2}$. If $E = \mu\kappa$,

$$\hat{c}\hat{E} \leqslant C\max\left\{\kappa\Big(\log\Big(\frac{{\gamma_1}^2}{\kappa}\Big)\Big)^{2(\tau+1)}\max\left\{\frac{1}{\gamma_1^2}, \frac{1}{\bar{\gamma}_2^2}\right\}, \frac{\kappa(\log\kappa^{-1})^{2(\tau+1)}}{\bar{\gamma}^2},\ \frac{\kappa}{\rho_1^2},\ \frac{\kappa}{\rho_0^2}\right\},$$

with a constant C not involving $\check{c}_1$. Then, from (C.1) and $\rho_1 \geqslant \check{c}_1\sqrt{\kappa}$ there follows

$$\hat{c}\hat{E} < C\max\left\{\frac{1}{\gamma_*}, \frac{1}{\check{c}_1^2}, \frac{\kappa}{\rho_0^2}\right\} < 1 \tag{C.7}$$

provided γ_*, $\check{c}_1^2 > C$ and $\kappa < C^{-1}\rho_0^2$. On the other hand, in the case $E = C\mu^2\bar{K}^{2\tau+1}\bar{\gamma}^{-2}$

$$\begin{aligned}\hat{c}\hat{E} \leqslant C\max\Big\{&\mu(\log\kappa^{-1})^{2\tau+1}\bar{\gamma}^{-2}\ \Big(\log\Big(\frac{{\gamma_1}^2}{\mu(\log\kappa^{-1})^{2\tau+1}\bar{\gamma}^{-2}}\Big)\Big)^{2(\tau+1)}\max\{\frac{1}{\gamma_1^2}, \frac{1}{\bar{\gamma}_2^2}\},\\ &\frac{\mu(\log\kappa^{-1})^{4(\tau+1)}}{\bar{\gamma}^4},\ \frac{\mu(\log\kappa^{-1})^{2\tau+1}\bar{\gamma}^{-2}}{\rho_1^2},\ \frac{\mu(\log\kappa^{-1})^{2\tau+1}\bar{\gamma}^{-2}}{\rho_0^2}\Big\}.\end{aligned}$$

Using now that $\rho_1 \geqslant \check{c}_1\frac{\sqrt{\mu}(\log\kappa^{-1})^{\tau+1/2}}{\bar{\gamma}}$ and again the definition of $\bar{\gamma}$ in (C.1), we again find an inequality like in (C.7), but with $\frac{\kappa}{\rho_0^2}$ replaced by $\frac{\sqrt{\mu}}{\rho_0^2\gamma_*^2}$

Finally, since the KAM condition $\hat{c}\hat{E} < 1$ is met, [12, Proposition 3] holds in this case. Then, we can find a set of invariant tori

$$\mathcal{K}_* \subset \bar{D}_r \times \mathbb{T}^{n_1} \times \left\{2\rho_1 < p_i^2 + q_i^2 < 2(\epsilon_0/3)^2,\ \forall\ i\right\}_r \subset (\mathcal{P}_{\sqrt{2}\epsilon_0/3})_r$$

(with $r < C\bar{\gamma}_2$) satisfying the measure estimate

$$\begin{aligned}\operatorname{meas}\left(\mathcal{P}_{\sqrt{2\check{c}_2}\epsilon_0} \setminus \mathcal{K}_*\right) &\leqslant \operatorname{meas}\left(\mathcal{P}_{\sqrt{2\check{c}_2}\epsilon_0}\right)_r \setminus \mathcal{K}_*)\\ &\leqslant C(\bar{\gamma} + \gamma_1 + \frac{\bar{\gamma}_2}{\epsilon_0^2} + \kappa^{n_2/4})\operatorname{meas}\mathcal{P}_{\sqrt{2}\epsilon_0/3}).\end{aligned} \tag{C.8}$$

---

[61] Use the definition of $\bar{\gamma}$ in (C.1).

[62] Compare, in particular, (C.4) for the choice of E and recall Equation (C.6) and the definition of $\bar{\rho}$ and of $\bar{K}$ in (C.2).

We omit to detail how (C.8) follows from [12, (34)]. For example, the reader may easily modify the end of [12, Proof of Theorem 1.4, Step 6].

The theorem is so proved with $\mathcal{K} := \mathcal{K}_* \cap \mathcal{P}_{\epsilon_0/3}$, $\epsilon_1 = \sqrt{2}\epsilon_0/3$, $\kappa_* := \min\{C^{-1/4}\sqrt{\rho_0},\, \epsilon_0/(3\sqrt{\check c_1})\}$, $\mu_* := \min\{C^{-2}\rho_0^4\gamma_*^4,\ \gamma_\star^4(\epsilon_0/(3\sqrt{\check c_1}))^8\}$. □

APPENDIX D. THE THEOREM BY N. N. NEHOROŠEV

Below is a more technical statement of Theorem 3, as it follows from [33] and, especially, [34].

The statement in [33]–[34] is based on the notion of "steepness" for a given smooth function $H_0(I) = H_0(I_1, \cdots, I_{n_1})$ of n_1 arguments. We shall adopt the definition given in [33]. This definition involves a number of parameters, denoted, in [33], as $(g,\, m,\, C_1,\, \cdots,\, C_{n_1-1},\, \delta_1,\, \cdots,\, \delta_{n_1-1},\, \mathfrak{a}_1,\, \cdots,\, \mathfrak{a}_{n_1-1})$. Accordingly, we shall call a given function $(g,\, m,\, C_1,\, \cdots,\, C_{n_1-1},\, \delta_1,\, \cdots,\, \delta_{n_1-1},\, \mathfrak{a}_1,\, \cdots,\, \mathfrak{a}_{n_1-1})$-steep, if it is steep with such parameters. See [33, p. 28 and p. 36] for details.

Theorem 11 ([33], p. 30; [34]). *Let $H = H_0(I) + P(I, \varphi, p, q)$ be real-analytic on $\mathcal{P}_\rho := V_\rho \times \mathbb{T}^{n_1}_\rho \times B^{2n_2}_\rho$ and assume that $I \in V \to H_0(I)$ is $(g,\, m,\, C_1,\, \cdots,\, C_{n_1-1},\, \delta_1,\, \cdots,\, \delta_{n_1-1},\, \mathfrak{a}_1,\, \cdots,\, \mathfrak{a}_{n_1-1})$-steep, with $\rho < 1 < m$. Then, one can find a, $b \in (0,1)$ and*[63] *$0 < M_\star < \rho^{1/b}$ such that, if*

$$M := \sup_{\mathcal{P}_\rho} |\partial P| \in (0, M_\star) \tag{D.1}$$

any trajectory $t \to \gamma(t) = (I(t), \varphi(t), p(t), q(t))$ solution of H such that

$$(p(t), q(t)) \in B^{2n_2}, \quad \forall\ 0 \leqslant t \leqslant T := \frac{1}{M} e^{\frac{1}{M^a}} \tag{D.2}$$

verifies

$$|I(t) - I(0)| \leqslant r := \frac{1}{2} M^b \qquad \forall\ 0 \leqslant t \leqslant T.$$

The number $M_\star$ can be taken to be[64]

$$M_\star = \min\left\{ \left(\frac{\rho}{2}\right)^{1/b}, \quad M_0 \right\} \tag{D.3}$$

where M_0 verifies

$$M_0 \geqslant \frac{c_0}{\rho} \min\left\{ \left(\frac{C_{n_1-1}}{g}\right)^p,\ \left(\frac{\rho}{m}\right)^p,\ \left(\frac{m}{C_r}\right)^p,\ \left(\frac{1}{\max_r \delta_r}\right)^p,\ 1 \right\} \tag{D.4}$$

for some $c_0 < 1 < p$ depending only on n_1, n_2 and $\mathfrak{a}_1, \cdots, \mathfrak{a}_{n_1-1}$.

---

[63] We changed a bit notations of [33]. Let us call $\bar{\mathcal{P}}$, $\bar\rho$ the quantities that in the statement of [33, The main theorem, p. 30] are called F, ρ (clearly, s, n, H_1, G, D of [33] correspond to our n_1, n_2, P, V, B^{2n_2}). In the statement of [33, The main theorem, p. 30], condition (D.2) is required, with $\mathcal{P}$ replaced by $\bar{\mathcal{P}}_{-2r}$, where $\bar{\mathcal{P}}_{-2r}$ is a real set defined as the biggest subset $\mathcal{A} \subset \bar{\mathcal{P}}$ for which $\mathcal{A}_{2r} \subset \bar{\mathcal{P}}$. Plainly $(\bar{\mathcal{P}}_{-2r})_{2r+\bar\rho} = \bar{\mathcal{P}}_{\bar\rho}$. Letting $\mathcal{P} := \bar{\mathcal{P}}_{-2r}$ and $\rho := 2r + \bar\rho$ we have our statement. Our condition $M_\star < \rho^{1/b}$ corresponds to [33] 's assumption $\bar\rho > 0$.

[64] See [34, p. 53]. By the previous note, we have to replace ρ in [34, p. 53] with $\bar\rho := \rho - 2r$. Note that condition $M_\star < (\frac{\rho}{2})^{1/b}$ implies $\rho \geqslant \rho - 2r = \rho - M^b \geqslant \rho - M_\star^b \geqslant \frac{\rho}{2}$. With this observation, we are allowed to identify ρ of [34, p. 53] with our ρ. Letting then M_0, M_1 and M_2 as in [34, p. 53], one sees, using the formulae in [34, pp. 48–57], that $M_1 = \frac{c_1}{\rho^8 m^4}(\frac{C_{n_1-1}}{g})^p$, while, since $\rho < 1 < m$, $M_2 = c_2 \min\left\{\frac{1}{m\rho^2}\left(\frac{\rho}{m}\right)^p, \frac{1}{m\rho^2}\left(\frac{C_{n_1-1}}{g}\right)^p, \frac{1}{\rho^2 m}\left(\frac{m}{C_r}\right)^p, \frac{1}{m\rho^2}\left(\frac{1}{\max_r \delta_r}\right)^p, \frac{1}{m\rho^2}\right\}$. Therefore, $M_0 := \min\{M_1,\ M_2\}$ verifies the inequality in (D.4).

Steepness Conditions

In [33], a function $H_0 = H_0(I)$ of n_1 variables $(I_1, \cdots, I_{n_1})$ is called "quasi-convex" in I if the system

$$\begin{cases} \sum_{j=1}^{n_1} \partial_{I_j} H_0(I)\eta_j = 0 \\ \sum_{j,k=1}^{n_1} \partial^2_{I_j I_k} H_0(I)\eta_j\eta_k = 0 \end{cases}$$

has the only trivial solution. Concave or convex functions, having definite in sign Hessian $\partial^2_{I_j I_k} H_0(I)$, are in particular quasi-convex. Moreover, H_0 is said to satisfy the three-jet conditions if, again, the system

$$\begin{cases} \sum_{j=1}^{n_1} \partial_{I_j} H_0(I)\eta_j = 0 \\ \sum_{j,k=1}^{n_1} \partial^2_{I_j I_k} H_0(I)\eta_j\eta_k = 0 \\ \sum_{j,k,h=1}^{n_1} \partial^3_{I_j I_k, I_h} H_0(I)\eta_j\eta_k\eta_h = 0 \end{cases} \tag{D.5}$$

has the only trivial solution.

In [32] it is proved that quasi-convex functions and functions satisfying the three-jet condition are steep.

ACKNOWLEDGMENTS

I wish to express my regards and deep admiration to Professor Alain Chenciner in occasion of his seventieth birthday and thank warmly Alain Albouy and Alexey Borisov for inviting me to present this paper in that important circumstance.

Moreover, I would like to thank Massimiliano Berti, Jacques Féjoz, Boris Khesin, Sergei Kuksin, Edi Zehnder and Vittorio Coti Zelati for honoring me, since my PhD, of their comments on the planetary problem and their encouragement.

Doubly thanks to Jacques Féjoz, referée of my PhD thesis, for mentioning in [19] my contribution to the proof of Arnold's Theorem, especially, my rediscovery of Deprit's reduction. I remember with much pleasure the long, relaxing and enlightening discussions on the planetary problem since we met for the first time in early Spring of 2008. I feel deeply indebted with him, since then.

Thanks to Alessandra Celletti, Giancarlo Benettin, Luca Biasco, Francesco Fassò, Massimiliano Guzzo and Lei Zhao, who helped me with precious bibliographic advices.

To the anonymous Referees, for their thoughtful advices, without which this paper would not have this form. In particular, for advertising me on the possibility of more global results than I stated in Theorem B, in the case of the planar three-body problem.

To my early adviser Luigi Chierchia, without whom nothing of what I did since my PhD would exist.

To my husband, without whom nothing would have meaning.

Research Supported by "Prin 2009 project Critical Point Theory and Perturbative Methods for Nonlinear Differential Equations" and European Research Council under FP7 project.

REFERENCES

1. Abdullah, K. and Albouy, A., On a Strange Resonance Noticed by M. Herman, *Regul. Chaotic Dyn.,* 2001, vol. 6, no. 4, pp. 421–432.
2. Arnol'd, V. I., Proof of a Theorem of A. N. Kolmogorov on the Invariance of Quasi-Periodic Motions under Small Perturbations of the Hamiltonian, *Uspekhi Mat. Nauk,* 1963, vol. 18, no. 5(113), pp. 13–40 [*Russian Math. Surveys,* 1963, vol. 18, no. 5(113), pp. 9–36].
3. Arnol'd, V. I., *Mathematical Methods of Classical Mechanics,* 2nd ed., Grad. Texts in Math., vol. 60, New York: Springer, 1989.
4. Arnol'd, V. I., Kozlov, V. V., and Neishtadt, A. I., *Mathematical Aspects of Classical and Celestial Mechanics,* 3rd ed., Encyclopaedia Math. Sci., vol. 3, Berlin: Springer, 2006.
5. Arnol'd, V. I., Small Denominators and Problems of Stability of Motion in Classical and Celestial Mechanics, *Uspekhi Mat. Nauk,* 1963, vol. 18, no. 6(114), pp. 91–192 [*Russian Math. Surveys,* 1963, vol. 18, no. 6(114), pp. 85–191].
6. Benettin, G., Guzzo, M., and Marini, V., Adiabatic Chaos in the Spin-Orbit Problem, *Celestial Mech. Dynam. Astronom.,* 2008, vol. 101, nos. 1–2, pp. 203–224.
7. Biasco, L., Chierchia, L., and Valdinoci, E., Elliptic Two-Dimensional Invariant Tori for the Planetary Three-Body Problem, *Arch. Ration. Mech. Anal.,* 2003, vol. 170, no. 2, pp. 91–135. See also: Biasco, L., Chierchia, L., and Valdinoci, E., Corrigendum to: "Elliptic Two-Dimensional Invariant Tori for the Planetary Three-Body Problem", *Arch. Ration. Mech. Anal.,* 2006, vol. 180, no. 3, pp. 507–509.
8. Boigey, F., Élimination des nœuds dans le problème newtonien des quatre corps, *Celestial Mech.,* 1982, vol. 27, no. 4, pp. 399–414.
9. Bounemoura, A. and Niederman, L., Generic Nekhoroshev Theory without Small Divisors, *Ann. Inst. Fourier (Grenoble),* 2012, vol. 62, no. 1, pp. 277–324.
10. Chierchia, L., The Planetary N-Body Problem, in *Celestial Mechanics,* A. Celletti (Ed.), Encyclopedia of Life Support Systems (EOLSS), vol. 6.119.55 (2012).
11. Chierchia, L. and Gallavotti, G., Drift and Diffusion in Phase Space, *Ann. Inst. H. Poincaré Phys. Théor.,* 1994, vol. 60, no. 1, 144 pp.
12. Chierchia, L. and Pinzari, G., Properly-Degenerate KAM Theory (Following V. I. Arnold), *Discrete Contin. Dyn. Syst. Ser. S,* 2010, vol. 3, no. 4, pp. 545–578.
13. Chierchia, L. and Pinzari, G., Deprit's Reduction of the Nodes Revised, *Celestial. Mech.,* 2011, vol. 109, no. 3, pp. 285–301.
14. Chierchia, L. and Pusateri, F., Analytic Lagrangian Tori for the Planetary Many-Body Problem, *Ergodic Theory Dynam. Systems,* 2009, vol. 29, no. 3, pp. 849–873.
15. Chierchia, L. and Pinzari, G., Planetary Birkhoff Normal Forms, *J. Mod. Dyn.,* 2011, vol. 5, no. 4, pp. 623–664.
16. Chierchia, L. and Pinzari, G., The Planetary N-Body Problem: Symplectic Foliation, Reductions and Invariant Tori, *Invent. Math.,* 2011, vol. 186, no. 1, pp. 1–77.
17. Deprit, A., Elimination of the Nodes in Problems of n Bodies, *Celestial Mech.,* 1983, vol. 30, no. 2, pp. 181–195.
18. Féjoz, J., Démonstration du "théorème d'Arnold' sur la stabilite" du système planétaire (d'après Herman), *Ergodic Theory Dynam. Systems,* 2004, vol. 24, no. 5, pp. 1521–1582.
19. Féjoz, J., On "Arnold's theorem" on the Stability of the Solar System, *Discrete Contin. Dyn. Syst.,* 2013, vol. 33, no. 8, pp. 3555–3565.
20. Ferrer, S. and Osácar, C., Harrington's Hamiltonian in the Stellar Problem of Three Bodies: Reductions, Relative Equilibria and Bifurcations, *Celestial Mech. Dynam. Astronom.,* 1994, vol. 58, no. 3, pp. 245–275.
21. Giorgilli, A., Locatelli, U., and Sansottera, M., Kolmogorov and Nekhoroshev Theory for the Problem of Three Bodies, *Celestial Mech. Dynam. Astronom.,* 2009, vol. 104, nos. 1–2, pp. 159–173.
22. Herman, M. R., Torsion du problème planétaire, http://www.college-de-france.fr/default/EN/all/equ_dif/archives_michel_herman.htm (edited by J. Féjoz in 2009).
23. Hofer, H. and Zehnder, E., *Symplectic Invariants and Hamiltonian Dynamics,* Birkhäuser Advanced Texts: Basler Lehrbücher, Basel: Birkhäuser, 1994.
24. *Proceedings of the International Congress of Mathematicians (Djursholm, Sweden, 15–22 August 1962),* V. Stenstrom (Ed.), Djursholm: Inst. Mittag-Leffler, 1963.
25. Jacobi, C. G. J., Sur l'élimination des noeuds dans le problème des trois corps, *Astron. Nachr.,* 1842, vol. 20, no. 462, pp. 81–98.
26. Kolmogorov, A. N., On the Conservation of Conditionally Periodic Motions under Small Perturbation of the Hamiltonian, *Dokl. Akad. Nauk SSSR,* 1954, vol. 98, pp. 527–530 (Russian).
27. Laskar, J. and Robutel, P., Stability of the Planetary Three-Body Problem: 1. Expansion of the Planetary Hamiltonian, *Celestial Mech. Dynam. Astronom.,* 1995, vol. 62, no. 3, pp. 193–217.
28. Laskar, J., Is the Solar System Stable?, arXiv: 1209.5996 (2012).

29. Lidov, M. L. and Ziglin, S. L., Non-Restricted Double-Averaged Three Body Problem in Hill's Case, *Celestial Mech.,* 1976, vol. 13, no. 4, pp. 471–489.
30. Malige, F., Robutel, P., and Laskar, J., Partial Reduction in the n-Body Planetary Problem Using the Angular Momentum Integral, *Celestial Mech. Dynam. Astronom.,* 2002, vol. 84, no. 3, pp. 283–316.
31. Moser, J., On Invariant Curves of Area-Preserving Mappings of an Annulus, *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II,* 1962, vol. 1962, pp. 1–20.
32. Nekhoroshev, N. N., Stable Lower Estimates for Smooth Mappings and for Gradients of Smooth Function, *Mat. Sb. (N. S.),* 1973, vol. 90(132), no. 3, pp. 432–478 [*Math. USSR Sbornik,* 1973, vol. 19, no. 3, pp. 425–467].
33. Nehorošev, N. N., An Exponential Estimate of the Time of Stability of Nearly Integrable Hamiltonian Systems, *Uspekhi Mat. Nauk,* 1977, vol. 32, no. 6 (198), pp. 5–66, 287 (Russian).
34. Nehorošev, N. N., An Exponential Estimate of the Time of Stability of Nearly Integrable Hamiltonian Systems: 2, *Trudy Sem. Petrovsk.,* 1979, no. 5, pp. 5–50 (Russian). See also: *Topics in Modern Mathematics, Petrovskii Seminar,* no. 5, O. A. Oleinik (Ed.), New York: Consultant Bureau, 1985.
35. Niederman, L., Stability over Exponentially Long Times in the Planetary Problem, *Nonlinearity,* 1996, vol. 9, no. 6, pp. 1703–1751.
36. Niederman, L., Hamiltonian Stability and Subanalytic Geometry, *Ann. Inst. Fourier (Grenoble),* 2006, vol. 56, no. 3, pp. 795–813.
37. Palacián, J. F., Sayas, F., and Yanguas, P., Regular and Singular Reductions in the Spatial Three-Body Problem, *Qual. Theory Dyn. Syst.,* 2013, vol. 12, no. 1, pp. 143–182.
38. Pinzari, G., On the Kolmogorov Set for Many-Body Problems, Università Roma Tre, 2009.
39. Pöschel, J., Nekhoroshev Estimates for Quasi-Convex Hamiltonian Systems, *Math. Z.,* 1993, vol. 213, no. 2, pp. 187–216.
40. Pusateri, F., Analytic KAM Tori for the Planetary $(n+1)$-Body Problem, Università degli Studi Roma Tre, 2006.
41. Radau, R., Sur une transformation des équations différentielles de la dynamique, *Ann. Sci. École Norm. Sup.,* 1868, vol. 5, pp. 311–375.
42. Robutel, Ph., Stability of the Planetary Three-Body Problem: 2. KAM Theory and Existence of Quasiperiodic Motions, *Celestial Mech. Dynam. Astronom.,* 1995, vol. 62, no. 3, pp. 219–261.
43. Rüssmann, H., Invariant Tori in Non-Degenerate Nearly Integrable Hamiltonian Systems, *Regul. Chaotic Dyn.,* 2001, vol. 6, no. 2, pp. 119–204.
44. Sansottera, M., Locatelli, U., and Giorgilli, A., On the Stability of the Secular Evolution of the Planar Sun − Jupiter − Saturn − Uranus System, *Math. Comput. Simulation,* 2013, vol. 88, pp. 1–14.
45. Schirinzi, G. and Guzzo, M., On the Formulation of New Explicit Conditions for Steepness from a Former Result of N. N. Nekhoroshev, *J. Math. Phys.,* 2013, vol. 54, no. 7, 072702, 22 pp.
46. Williamson, J., On the Algebraic Problem Concerning the Normal Forms of Linear Dynamical Systems, *Amer. J. Math.,* 1936, vol. 58, no. 1, pp. 141–163.
47. Zhao, L., Solutions quasi-périodiques et solutions de quasi-collision du problème spatial des trois corps, Université Paris VII Denis Diderot, 2013.