Springer Proceedings in Mathematics & Statistics

Giulio Baù · Sara Di Ruzza ·
Rocío Isabel Páez · Tiziano Penati ·
Marco Sansottera *Editors*

New Frontiers of Celestial Mechanics: Theory and Applications

I-CELMECH Training School, Milan, Italy, February 3–7, 2020

Springer

Springer Proceedings in Mathematics & Statistics

Volume 399

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including data science, operations research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

Giulio Baù · Sara Di Ruzza · Rocío Isabel Páez ·
Tiziano Penati · Marco Sansottera
Editors

New Frontiers of Celestial Mechanics: Theory and Applications

I-CELMECH Training School, Milan, Italy, February 3–7, 2020

Springer

Editors
Giulio Baù
Department of Mathematics
University of Pisa
Pisa, Italy

Rocío Isabel Páez
School of Computer Science
and Information Technology
University College Cork UCC
Cork, Ireland

Marco Sansottera
Department of Mathematics
University of Milan
Milan, Italy

Sara Di Ruzza
Department of Mathematics
University of Palermo
Palermo, Italy

Tiziano Penati
Department of Mathematics
University of Milan
Milan, Italy

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-031-13114-1 ISBN 978-3-031-13115-8 (eBook)
https://doi.org/10.1007/978-3-031-13115-8

Mathematics Subject Classification: 70F15, 37N05, 37J40

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains the detailed text of the major lectures delivered during the I-CELMECH Training School 2020 held in Milan, whose aim was to present a contemporary review of recent results in the field of Celestial Mechanics, with special emphasis on theoretical aspects. In fact, the modern developments of Celestial Mechanics have their roots in the mathematics of Hamiltonian perturbation theory and of the dynamical systems. It is remarkable how in the last decades, in parallel to the celebrated theorems of these disciplines, constructive methods have been developed in order to tackle problems with a long tradition, such as the stability of the Solar System, the rotations of celestial bodies and orbit determination, as well as the novel scientific needs raised by the discovery of exoplanetary systems, the management of the space debris problem and the modern space mission design.

Celestial Mechanics training schools have a great tradition that dates back to 1981 with the school of the NATO Advanced Study Institute held at Cortina d'Ampezzo and that is still going on with the recent summer school SDSM 2017 "Satellite Dynamics and Space Missions: Theory and Applications of Celestial Mechanics", held in San Martino al Cimino.

Quoting the preface of the Proceedings of training school in Cortina d'Ampezzo in 1981, where it all began, *the Editors of the Proceedings were extremely fortunate to have obtained the cooperation of outstanding lecturers who were clear, thorough, understandable, patient to answer questions, but above all, had knowledge of the most recent developments in our field of interest.* This demonstrates a continuity of tradition in our community of which we are extremely proud. This preface will briefly review the content of the lectures.

The contribution by Ugo Locatelli focuses on the explicit construction of invariant tori exploiting suitable Hamiltonian normal forms, with particular emphasis on applications to Celestial Mechanics. First, the algorithm constructing the Kolmogorov normal form is described in detail. Then the extension to lower-dimensional elliptic tori is provided. Both algorithms are then combined so as to accurately approximate the long-term dynamics of the HD 4732 extrasolar system.

The contribution by Gabriella Pinzari presents a review of some results of their research group, regarding the relation between some particular motions of the Three-Body problem (3BP) and the motions of the so-called Euler (or two-centre) problem, which is integrable. For the analysis of such relation, the authors make use of two novel results: on one hand, the two-centre problem (2CP) bears a remarkable property, here called *renormalizable integrability*, which states that the simple averaged potential of the 2CP and the Euler integral are one function of the other; on the other hand, the motions of the Euler integral are at least qualitatively explicit, and the averaged Newtonian potential is a prominent part of the 3BP Hamiltonian. In previous works, the authors give partial answers to the question whether the motions of the Euler integral can be traced in the 3BP, here revisited.

The contribution by Alessandra Celletti deals with dissipative systems, a key topic in Celestial Mechanics. In particular the problem of the existence of invariant tori for *conformally symplectic* systems, which have the property to transform the symplectic form into a multiple of itself, is studied. Two different models are presented: a discrete system known as the standard map and a continuous system known as the spin-orbit problem. In both cases, both the conservative and dissipative versions are considered, in order to highlight the differences between the symplectic and conformally symplectic dynamics. Finally, a short survey of the present state of KAM (Kolmogorov-Arnold-Moser) estimates for the existence of invariant tori in the conservative and dissipative standard maps and spin-orbit problems is provided.

The contribution by Gwenaël Boué provides basic tools to understand the rotational dynamics of extended bodies which could be either rigid or deformable by tides. The problem is described in a Lagrangian formalism as it was developed by H. Poincaré in 1901. The case of rigid body is also presented in the corresponding Hamiltonian formalism. The mathematical description of the deformation of the extended body follows the approach used by C. Ragazzo and L. Ruiz in their two papers of 2015, 2017 due to the compactness and clarity of their formalism. In this Chapter, many applications to the rotation and the libration of celestial bodies are illustrated.

The contribution by Christos Efthymiopoulos concerns the phenomenon of Arnold diffusion. The authors begin with the famous example given by Arnold to describe the slow diffusion taking place in the action-space in Hamiltonian nonlinear dynamical systems with three or more degrees of freedom. The text introduces basic concepts related to our current understanding of the mechanisms leading to Arnold diffusion and at the same time performed a qualitative investigation of the phenomenon of Arnold diffusion with many examples. The problem of the speed of diffusion is investigated using methods of perturbation theory, with particular emphasis on Nekhoroshev's theorem. The choice of the authors of mixing the analytical and numerical parts trying to explain in a very rigorous way numerical results makes the paper self-consistent.

The contribution by Giovanni F. Gronchi deals with the problem of initial orbit determination of a solar system body, i.e. the determination of a preliminary orbit from observations collected for example by a telescope. The two methods that are presented, named `Link2` and `Link3`, try to link together two and three, respectively,

short arcs of optical observations of the same object which can possibly be quite far apart in time. The conservation laws of Kepler's problem are used to derive a polynomial equation of degree 9 (`Link2`) and 8 (`Link3`) for the distance of the body from the observer. Some numerical examples show the performance of the two linkage methods when they are applied to real observations of asteroids.

The contribution by Catalin Gales provides an overview of some recent developments in the study of dynamics of space debris with focus on specific resonant interactions, in particular those related to the tesseral resonances. After an historical introduction to the topic, the authors provide a long-term picture of the dynamics that can help in the modeling and mitigation of the space debris problem, both in term of Cartesian coordinates and in the Hamiltonian framework. Some key terms in the perturbing functions are classified, while the effect of the dissipative force of the atmospheric drag is also formulated. In the following sections, the authors describe the effects of the tesseral resonances in the LEO, MEO and GEO regions. Finally, the effect of lunisolar resonances in high-altitude orbits and future perspectives are discussed.

The contribution by Antonio Giorgilli provides an answer to a simple question, *how did Kepler discover his celebrated laws?*. The answer however is not that simple and the present paper guides the reader by a short walk along the main works of Kepler, notably the *Astronomia Nova*, trying to follow his search of the perfection of the World till the discovery of his celebrated laws. At the end of the road, the consciousness that the finish line had not yet been reached.

The I-CELMECH Training School was attended by 86 participants from all over the world and it was made possible, thanks to the support of project MIUR-PRIN 20178CJA2B "New frontiers of Celestial Mechanics: theory and applications", the Department of Mathematics "Federigo Enriques" of the Università degli Studi di Milano and the MusAB "Museo Astronomico di Brera".

Pisa, Italy — Giulio Baù
Palermo, Italy — Sara Di Ruzza
Cork, Ireland — Rocío Isabel Páez
Milan, Italy — Tiziano Penati
Milan, Italy — Marco Sansottera

Contents

About the Editors

Giulio Baù is Associate Professor at the Department of Mathematics of the University of Pisa. His research activity deals with the development of new orbit propagation and determination methods for small celestial bodies (in particular asteroids and space debris), the study of their dynamics, and regularization techniques in the N-body problem.

Sara Di Ruzza is a Researcher of Mathematical Physics at the Department of Mathematics at the University of Palermo. Her main field of interest is Celestial Mechanics. She worked on the spin-orbit problem, on some particular cases of the three-body problem both from an analytical and numerical point of view. She applied theory to real scenarios such as asteroids motion and space missions. The latest works are focused on the occurrence of chaos in the planar three-body problem.

Rocío Isabel Páez is a Senior Postdoctoral Researcher at the University College Cork, in Ireland. She received her Ph.D. in Mathematics from the University of Rome "Tor Vergata" in 2016. Since then, she has held academic positions at the University of Rome "Tor Vergata", the Academy of Athens, and the University of Padova. Her research area is in Applied Mathematics, with a focus on the application of advanced methods of perturbation theory and numerical simulations in Dynamical Astronomy.

Tiziano Penati is Associate Professor of Mathematical Physics at the Department of Mathematics, University of Milan. His main field of interest is Hamiltonian Perturbation Theory, with special attention to Normal Form techniques. His work focuses on the investigation of time-periodic localized structures and metastability phenomena in Hamiltonian Lattices, like the Fermi–Pasta–Ulam–Tsingou, the discrete Klein–Gordon, or the discrete Nonlinear Schroedinger models.

Marco Sansottera is a Researcher of Mathematical Physics at the Department of Mathematics, University of Milan. His research activity is mainly focused on Dynamical Systems and Celestial Mechanics. In particular, he studied the stability properties of planetary systems, investigating the dynamics in the neighborhood of some invariant objects, like maximal dimension KAM tori, equilibrium points, and lower dimensional elliptic tori.

Invariant KAM Tori: From Theory to Applications to Exoplanetary Systems

Ugo Locatelli, Chiara Caracciolo, Marco Sansottera and Mara Volpi

Abstract We consider the classical problem of the construction of invariant tori exploiting suitable Hamiltonian normal forms. This kind of approach can be translated by means of the Lie series method into explicit computational algorithms, which are particularly suitable for applications in the field of Celestial Mechanics. First, the algorithm constructing the Kolmogorov normal form is described in detail. Then, the extension to lower-dimensional elliptic tori is provided. We adopt the same formalism and notations in both cases, with the aim of making the latter easier to understand. Finally, they are both used in a combined way in order to approximate carefully the secular dynamics of the extrasolar system hosting two planets orbiting around the HD 4732 star.

Keywords Elliptic lower-dimensional tori · KAM theory · Normal forms · Hamiltonian perturbation theory · Exoplanets · n-body planetary problem · Celestial mechanics

U. Locatelli (✉) · M. Volpi
Dipartimento di Matematica dell'Università degli Studi di Roma "Tor Vergata", via della ricerca scientifica 1, 00133 Roma, Italy
e-mail: locatell@mat.uniroma2.it

C. Caracciolo · M. Sansottera
Dipartimento di Matematica dell'Università degli Studi di Milano, via Saldini 50, 20133 Milano, Italy

G. Baù et al. (eds.), *New Frontiers of Celestial Mechanics: Theory and Applications*, Springer Proceedings in Mathematics & Statistics 399,
https://doi.org/10.1007/978-3-031-13115-8_1

1 Introduction

The birth of KAM[1] theory was marked by a famous article written in 1954 by A.N. Kolmogorov, i.e., [22]. At that epoch the great potential of KAM theorem in order to solve interesting problems in the field of Celestial Mechanics was immediately understood. In fact, it was applied just a few years later in order to prove the stability of the triangular Lagrangian points in the planar restricted problem of three bodies for almost all admissible mass ratios (see [25]). Since then, several applications have ensured the existence of invariant tori in the context of other Hamiltonian models that are of interest in Celestial Mechanics (see, e.g., [10]). Nevertheless, the applications of KAM theory to physically realistic models have never been straightforward. This is mainly due to a few severe constraints that appear in the hypotheses of KAM theorem (e.g., concerning the smallness on the parameter ruling the size of the perturbation).

In the last few decades, the successful applications of KAM theory to Celestial Mechanics introduced more and more refinements in the preliminary work to adapt the Hamiltonian model in such a way to bypass the aforementioned difficulties (see, e.g., [27, 28]). In some other works, the novelty concerns the design of a new approach strategy. In particular, this has been made by combining the results provided by two different theorems; for instance, in [18, 20] the estimates *à la* Nekhoroshev have been applied in the neighborhood of an invariant KAM torus, by following the proof scheme described in [31]. This kind of strategy can be implemented in a natural way by adopting an approach based on suitable *normal forms*. Indeed, different normal form algorithms can be applied one after the other. This work has the ambitious goal of fully explaining a very recent type of applications in the field of Celestial Mechanics, where the computational procedure leading to the Kolmogorov normal form is performed in the neighborhood of a periodic orbit. In turn, such an invariant manifold is preliminarly located by a corresponding normal form for an elliptic torus. The addition of this intermediate step is crucial in order to successfully apply our

[1] It is worth to repeat, here and once again, the story explaining the choice of the acronym KAM. In 1954, during the International Congress of Mathematicians in Amsterdam, Kolmogorov presented his version of the (KAM) theorem. In the same year, he also wrote the very short article [22], where he provided just a scheme of the proof. According to a few direct witnesses, a few years later Kolmogorov explained all the details of his proof in a cycle of lectures delivered at the Moscow University. This was based on a sequence of canonical transformations coherently defined on a so called scale of Banach spaces; a modern reformulation of the proof that should be very similar to the original one is included in [11]. In 1963, V.I. Arnold (who had been a student of Kolmogorov) published a complete proof of the theorem, based on a different approach able to ensure the existence of a Cantor set including many invariant tori and having positive Lebesgue measure (see the statement of Corollary 1 and [1]). In the meantime, the german mathematician J. Moser developed a completely independent version of the proof in the case of symplectic mappings (see [33]). Let us also recall that at the beginning the correctness of the Kolmogorov's approach was doubtful for Moser. Indeed, also because of a famous sentence included in the report he wrote for *Mathematical Reviews* on the Kolmogorov's article (see MR0097508, 20 n. 4066), for many years Arnold's approach was thought to be the only viable one, in order to prove KAM theorem for quasi-integrable Hamiltonian systems.

computational algorithm in its entirety to extrasolar planetary systems with rather eccentric orbits (i.e., whose eccentricity values are significantly larger than those observed for the gaseous planets of our Solar System).

The first theoretical results about the existence of elliptic tori go back to [13, 29, 36]. In the last two decades, similar statements have been proved also in the context of Hamiltonian planetary systems (see [3, 4, 19]). In the present notes, we aim to develop an approach that is far from being purely theoretical. Indeed, we will explain how to extract from the proof schemes the information that is fundamental in order to properly design a computational procedure, which allows to determine invariant manifolds that are in good agreement with the orbital motions of extrasolar planets.

In the following, Sect. 2 contains a quick introduction of a few elementary notions concerning the Hamiltonian perturbation theory and a careful description of the normal form method constructing KAM tori. In Sect. 3, we show how that approach can be adapted for the construction of lower-dimensional invariant manifolds of elliptic type. In the final Sect. 4 our new application to an exoplanetary system is explained in detail; this is designed by combining the two kind of normal forms previously discussed, whose constructions are performed one after each other.

2 Basics of KAM Theory

2.1 Near to the Identity Canonical Transformations by Lie Series

Let us consider two generic dynamical functions $f = f(\boldsymbol{p}, \boldsymbol{q})$ and $\chi = \chi(\boldsymbol{p}, \boldsymbol{q})$, that are defined on all the phase space endowed by n pairs of conjugate canonical variables $(\boldsymbol{p}, \boldsymbol{q}) = (p_1, \ldots, p_n, q_1, \ldots, q_n)$. It is well known that the time evolution of f under the flow induced by χ is ruled by the Poisson bracket between these two functions, i.e., $\dot{f} = \frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{p}(t), \boldsymbol{q}(t)) = \{f, \chi\}$, where

$$\{f, \chi\} = \sum_{j=1}^{n} \frac{\partial f}{\partial q_j} \frac{\partial \chi}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial \chi}{\partial q_j} \tag{1}$$

and the flow $(\boldsymbol{p}(t), \boldsymbol{q}(t)) = \Phi_{\chi}^{t}\big(\boldsymbol{p}(0), \boldsymbol{q}(0)\big)$ is defined by the solution of the corresponding Hamilton equations

$$\dot{p}_j = -\frac{\partial \chi}{\partial q_j}, \qquad \dot{q}_j = \frac{\partial \chi}{\partial p_j}, \qquad \forall\, j = 1, \ldots, n \tag{2}$$

(being $(\boldsymbol{p}(0), \boldsymbol{q}(0))$ regarded as initial conditions).

Let us now focus on the Taylor expansion with respect to time of the generic dynamical function f, i.e., $f + t\dot{f} + \frac{t^2}{2}\frac{\mathrm{d}}{\mathrm{d}t}\dot{f} + \ldots = f + t\{f, \chi\} + \frac{t^2}{2}\{\{f, \chi\}, \chi\} + \ldots$, that can be reformulated in terms of Lie series. First, let us introduce the so called Lie derivative operator: $\mathcal{L}_\chi \cdot = \{\cdot, \chi\}$; in the present context, it is usual to refer to χ as the generating function of the corresponding Lie derivative. Thus, the previous Taylor expansion in time can be expressed as $\exp\left(t\mathcal{L}_\chi\right) f = \sum_{j=0}^{\infty} \frac{t^j}{j!}\mathcal{L}_\chi^j f$. It is common to define the Lie series operator just in the case with $t = 1$, i.e., it acts on the generic dynamical function f in such a way that

$$\exp\left(\mathcal{L}_\chi\right) f = \sum_{j=0}^{\infty} \frac{1}{j!}\mathcal{L}_\chi^j f \ ;$$

let us emphasize that this formula must be interpreted at a purely formal level, in the sense that we do not wonder about the convergence of the series. However, it can be ensured if the sup-norm of the generating function χ is small enough,[2] as it is natural to expect.

Since any single canonical coordinate can be seen as a particular dynamical function, we can express the Hamiltonian flow at time 1 in terms of Lie series in the following way:

$$\Phi_\chi^1(\boldsymbol{p}, \boldsymbol{q}) = \exp\left(\mathcal{L}_\chi\right)(\boldsymbol{p}, \boldsymbol{q}), \tag{3}$$

where, for every pair of canonical variables $(p_i\ ,\ q_i)$ (being $i = 1, \ldots, n$), we put

$$\Phi_\chi^1 p_i = \exp\left(\mathcal{L}_\chi\right) p_i \ , \qquad \Phi_\chi^1 q_i = \exp\left(\mathcal{L}_\chi\right) q_i \ .$$

It is well known that the Hamiltonian flow is canonical, then we readily obtain that the map defined by the Lie series operator in the right hand side of (3) is canonical as well. Moreover, such a change of coordinates is obviously close to the identity in the limit of the generating functions shrinking to zero.

The canonical formalism makes very convenient the writing of the equations of motion in the new variables. Let us assume that the evolution in the original set of coordinates $(\boldsymbol{p}, \boldsymbol{q})$ is ruled by a single function $H = H(\boldsymbol{p}, \boldsymbol{q})$ entering the Hamilton equations (2) in place of χ; moreover, let $(\boldsymbol{p}, \boldsymbol{q}) = \mathcal{C}(\boldsymbol{P}, \boldsymbol{Q})$ be a canonical transformation. Therefore, the new equations of motions can be written as follows:

$$\dot{P}_j = -\frac{\partial \mathcal{K}}{\partial Q_j} \ , \qquad \dot{Q}_j = \frac{\partial \mathcal{K}}{\partial P_j} \ , \qquad \forall\ j = 1, \ldots, n, \tag{4}$$

being $\mathcal{K}(\boldsymbol{P}, \boldsymbol{Q}) = H\big(\mathcal{C}(\boldsymbol{P}, \boldsymbol{Q})\big)$ the new Hamiltonian function. In such a context, the Lie series formalism makes automatic (and, then, somehow easier) the procedure

[2] The convergence of the Lie series is carefully discussed in [14, 15]; in particular, the explanatory notes in [15] contains also a rather self-consistent introduction to the Lie series formalism in the Hamiltonian framework.

of substitution, because of the so called "exchange theorem" (see [14]). In fact, if χ is a small enough generating function, the new Hamiltonian can be expressed as

$$\mathcal{K}(\boldsymbol{P},\,\boldsymbol{Q}) = \exp\left(\mathcal{L}_\chi\right) H\Big|_{(\boldsymbol{p},\boldsymbol{q})=(\boldsymbol{P},\,\boldsymbol{Q})}\,,$$

this means that we can apply the Lie series to the old Hamiltonian function so as to rename the variables, only at the end. For more detailed explanations we defer to the whole Sect. 4.1 of [15]. Of course, the same computational procedure holds also for the corresponding canonical transformation, that is given by

$$(\boldsymbol{p},\boldsymbol{q}) = \mathcal{C}(\boldsymbol{P},\,\boldsymbol{Q}) = \exp\left(\mathcal{L}_\chi\right)(\boldsymbol{p},\boldsymbol{q})\Big|_{(\boldsymbol{p},\boldsymbol{q})=(\boldsymbol{P},\,\boldsymbol{Q})}\,.$$

2.2 Statement(s) of KAM Theorem

First, let us recall the statement of KAM theorem as in its very first version introduced by Kolmogorov (see [22]).

Theorem 1 (KAM, according to the version due to Kolmogorov) *Consider a Hamiltonian function* $H : \mathcal{A} \times \mathbb{T}^n \mapsto \mathbb{R}$ *(being* $\mathcal{A} \subseteq \mathbb{R}^n$ *an open set) of the form* $H(\boldsymbol{p},\boldsymbol{q}) = \boldsymbol{\omega}\cdot\boldsymbol{p} + h(\boldsymbol{p}) + \varepsilon f(\boldsymbol{p},\boldsymbol{q})$ *where* h *is at least quadratic with respect to the actions* $\boldsymbol{p}$*, i.e.,* $h(\boldsymbol{p}) = \mathcal{O}(\|\boldsymbol{p}\|^2)$ *for* $\boldsymbol{p} \to \boldsymbol{0}$*. Moreover, let us assume the following hypotheses:*

(a) $\boldsymbol{\omega}$ *is Diophantine; this means that there are two positive constants*[3] γ *and* τ *such that* $|\boldsymbol{k}\cdot\boldsymbol{\omega}| \geq \frac{\gamma}{|\boldsymbol{k}|^\tau}\ \forall\, \boldsymbol{k} \in \mathbb{Z}^n \setminus \{\boldsymbol{0}\}$*;*

(b) H *is analytic on its action–angle*[4] *domain of definition* $\mathcal{A} \times \mathbb{T}^n$*;*

(c) $h(\boldsymbol{p})$ *is non-degenerate, i.e.,* $\det\left(\frac{\partial^2 h}{\partial p_i \partial p_j}(\boldsymbol{p})\right)_{i,j} \neq 0\ \forall\, \boldsymbol{p} \in \mathcal{A}$*;*

(d) ε *is a small enough parameter.*

Therefore, there is a canonical transformation $(\boldsymbol{p},\boldsymbol{q}) = \psi_\varepsilon(\boldsymbol{P},\,\boldsymbol{Q})$*, leading* H *in the so called Kolmogorov normal form* $\mathcal{K}(\boldsymbol{P},\,\boldsymbol{Q}) = \omega\cdot\boldsymbol{P} + \mathcal{O}(\|\boldsymbol{P}\|^2)$*, being* $\mathcal{K} = H \circ \psi_\varepsilon$*.*

In our exposition of these topics, we do not consider all the very interesting mathematical work that has been done in the last fifty years in order to weaken the

[3] Indeed, in order to satisfy the Diophantine inequality, it is essential that $\tau \geq n-1$.

[4] Although there exist formulations of the KAM theorem that are not dealing with action–angle canonical coordinates (see, e.g., [12]), we stress that this is a rather natural framework to assume. In fact, by definition a n–dimensional torus $\mathbb{T}^n$ is in a bijective correspondence with n angles, denoted as $(q_1\,,\,\ldots\,,\,q_n)$ in agreement with the text. Thus, they can be adopted as coordinates. Let us recall that in Hamiltonian mechanics the product between each conjugate pair of canonical variables has the physical dimension of an action, that is the same as an angular momentum. Therefore, $\forall\ j = 1,\,\ldots\,n$, the conjugate momentum p_j is an action, because $(\boldsymbol{p},\boldsymbol{q})$ are assumed to be canonical coordinates.

assumptions on the KAM theorem. We prefer to focus on what makes the theorem suitable to apply to interesting physical problems. This is somehow hidden in the thesis of the statement and, mainly, in the proof scheme. Let us highlight such a content.

One can easily verify that, if the Hamiltonian is in the Kolmogorov normal form $\mathcal{K}(\boldsymbol{P},\,\boldsymbol{Q}) = \omega\cdot\boldsymbol{P} + \mathcal{O}(\|\boldsymbol{P}\|^2)$, then $t \mapsto (\boldsymbol{P}(t) = \boldsymbol{0}\,,\ \boldsymbol{Q}(t) = \boldsymbol{Q}_0 + \omega t)$ is the solution for the equations of motion (4) starting from the generic initial conditions $(\boldsymbol{P}(0)\,,\ \boldsymbol{Q}(0)) = (\boldsymbol{0}\,,\ \boldsymbol{Q}_0)$. Since the canonical transformations enjoy the property of preserving solutions, this allows us to design the following integration scheme for the equations of motion (3), when the generic Hamiltonian χ is replaced by H, that describes the problem we are considering:

$$\begin{array}{ccc} \big(\boldsymbol{p}(0),\boldsymbol{q}(0)\big) & \overset{\psi_\varepsilon^{-1}}{\longrightarrow} & (\boldsymbol{P}(0)\,,\ \boldsymbol{Q}(0)) \\ & & \Big\downarrow \Phi_{\mathcal{K}}^t \\ \big(\boldsymbol{p}(t),\boldsymbol{q}(t)\big) & \overset{\psi_\varepsilon}{\longleftarrow} & (\boldsymbol{P}(t)\,,\ \boldsymbol{Q}(t)) \end{array} \quad . \tag{5}$$

In the scientific literature, this way to compute the motion law $t \mapsto (\boldsymbol{p}(t)\,,\ \boldsymbol{q}(t)) = \Phi_H^t(\boldsymbol{p}(0)\,,\ \boldsymbol{q}(0))$ is often said to be semi-analytic. Such a name is due to the fact that the schematic procedure above is usually performed after having determined the Fourier expansions of the canonical transformation ψ_ε, by using a software package designed for doing computer algebra manipulations.

In spite of the fact that the very first version of the KAM theorem ensures the existence of a single invariant torus, the statement can be extended so as to cover a very generic situation. Indeed, in his very short but incredibly seminal article [22], Kolmogorov recalled a well known result of number theory: almost all n–dimensional vectors are Diophantine. This remark jointly with the uniform non-degeneracy of the so called action-frequency map in the integrable approximation, i.e., $\boldsymbol{p} \mapsto \boldsymbol{\omega}(\boldsymbol{p}) = \left(\frac{\partial h}{\partial p_i}(\boldsymbol{p})\right)_{i=1,\dots,n}$, allowed him to state the following result in [22].

Corollary 1 (KAM, according to the version proved by Arnold) *Consider a quasi-integrable Hamiltonian depending on action–angle variables, i.e., $H : \mathcal{A}\times\mathbb{T}^n \mapsto \mathbb{R}$ (being $\mathcal{A}\subseteq\mathbb{R}^n$ an open set) of the form $H(\boldsymbol{p},\boldsymbol{q}) = h(\boldsymbol{p}) + \varepsilon f(\boldsymbol{p},\boldsymbol{q})$. If we assume the same hypotheses (b)–(d) of Theorem 1, then there is a set $\mathcal{S}_\varepsilon$ that is made by invariant tori and is such that its Lebesgue measure $\mu\big(\mathcal{S}_\varepsilon\big)$ is positive. Moreover,*

$$\lim_{\varepsilon\to 0}\mu\Big(\big(\mathcal{A}\times\mathbb{T}^n\big)\setminus\mathcal{S}_\varepsilon\Big) = 0\,.$$

Let us emphasize that this statement highlights one of the main merits of the KAM theorem: it shows that there is a sort of continuity (in terms of the Lebesgue measure) between integrable systems and quasi-integrable ones. From one hand, this sort of intuitive concept was (and still is) considered to be extremely natural; on the other

hand, at that epoch such an expectation was in contrast with the famous theorem by Poincaré (that can be felt as somehow paradoxical, see [34]) on the non-existence of integrals of motion apart from the energy for a generic quasi-integrable Hamiltonian system.

Although the statement of the Corollary above can be easily deduced from the original version of the KAM theorem that is due to Kolmogorov, the proof scheme introduced by Arnold in [1] is extremely deep, because it provides a more global picture of the dynamics. This approach has been further extended, for instance, in [35], where it is proved that quasi-integrable Hamiltonian satisfying the usual hypotheses (b)–(d) of Theorem 1 can be conjugated to integrable ones via a canonical transformation that is not analytic, but it is $\mathcal{C}^{(\infty)}$.

2.3 Algorithmic Construction of the Kolmogorov Normal Form

These notes are focusing more on the applications based on the KAM theory rather than on the theory itself. Therefore, it is important to describe carefully the so called formal algorithm constructing the Kolmogorov normal form. The results about the convergence of such a computational procedure are very well established (see, e.g., [17]) and in the following we will just briefly recall them.

For the sake of definiteness, we need to introduce some notations. For a fixed positive integer K we introduce the distinct classes of functions $\mathcal{P}_{\ell,sK}$, for all non-negative indexes $\ell,\ s \geq 0$. Any generic function $g \in \mathcal{P}_{\ell,sK}$ can be written as

$$g(\boldsymbol{p},\boldsymbol{q}) = \sum_{\substack{\boldsymbol{j}\in\mathbb{N}^n \\ |\boldsymbol{j}|=\ell}} \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^n \\ |\boldsymbol{k}|\leq sK}} c_{\boldsymbol{j},\boldsymbol{k}}\, \boldsymbol{p}^{\boldsymbol{j}} \exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{q}\right), \tag{6}$$

where $(\boldsymbol{p},\boldsymbol{q})$ are action–angle canonical variables and the coefficients $c_{\boldsymbol{j},\boldsymbol{k}} \in \mathbb{C}$ satisfy the following relation: $c_{\boldsymbol{j},-\boldsymbol{k}} = \bar{c}_{\boldsymbol{j},\boldsymbol{k}}$ so that $g : \mathbb{R}^n \times \mathbb{T}^n \mapsto \mathbb{R}$. Moreover, in the previous formula, we have introduced the symbol $|\cdot|$ to denote the ℓ_1-norm (i.e., $|\boldsymbol{k}| = |k_1| + \ldots + |k_n|$) and we have adopted the multi-index notation, i.e., $\boldsymbol{p}^{\boldsymbol{j}} = p_1^{j_1} \cdot \ldots \cdot p_n^{j_n}$. In the following, we will adopt the usual notation for the average of a function g with respect to the generic angles $\boldsymbol{\vartheta} \in \mathbb{T}^n$, i.e., $\langle g \rangle_{\boldsymbol{\vartheta}} = \int_{\mathbb{T}^n} \mathrm{d}\vartheta_1 \ldots \mathrm{d}\vartheta_n\, g/(2\pi)^n$.

We will start the formal algorithm from a Hamiltonian of the following type:

$$\begin{aligned} H^{(0)}(\boldsymbol{p},\boldsymbol{q};\boldsymbol{\omega}^{(0)}) \;=\; & E^{(0)} + \boldsymbol{\omega}^{(0)}\cdot\boldsymbol{p} + \sum_{s\geq 0}\sum_{\ell\geq 2} f_\ell^{(0,\,s)}(\boldsymbol{p},\boldsymbol{q};\boldsymbol{\omega}^{(0)}) \\ & + \sum_{s\geq 1}\sum_{\ell=0}^{1} f_\ell^{(0,\,s)}(\boldsymbol{p},\boldsymbol{q};\boldsymbol{\omega}^{(0)}), \end{aligned} \tag{7}$$

where $f_\ell^{(0,s)} \in \mathcal{P}_{\ell,sK}$, being the first upper index related to the normalization step, and $E^{(0)} \in \mathcal{P}_{0,0}$ is a constant meaning the energy level of the torus $\left\{(\boldsymbol{p},\boldsymbol{q}) : \boldsymbol{p}=\boldsymbol{0},\ \boldsymbol{q} \in \mathbb{T}^n\right\}$ that is invariant in the integrable approximation. The occurrence of $\boldsymbol{\omega}^{(0)}$ at the end of the list of the arguments emphasizes that those functions depend also on that angular velocity vector in a parametric way. We also stress that the terms appearing in the second row of formula (7) have to be considered as the small perturbation we aim to remove in order to bring the Hamiltonian in Kolmogorov normal form. According to the definition given by Poincaré (see [34]), the *general problem of the dynamics* is described by a real analytic Hamiltonian of type $H(\boldsymbol{I},\boldsymbol{\varphi};\varepsilon) = h(\boldsymbol{I}) + \varepsilon f(\boldsymbol{I},\boldsymbol{\varphi})$, being $(\boldsymbol{I},\boldsymbol{\varphi})$ action–angle coordinates and ε a small parameter. It is well known that such an Hamiltonian can be put in the form (7) provided that the Hessian of the integrable part h is non-degenerate on its open domain, say $\mathcal{A} \subseteq \mathbb{R}^n$. Indeed, it is just matter of performing a canonical change of coordinates that translates the origin of the actions in correspondence to $\boldsymbol{I}^\star \in \mathcal{A}$, because

$$\left.\frac{\partial h(\boldsymbol{I})}{\partial I_j}\right|_{\boldsymbol{I}=\boldsymbol{I}^\star} = \left.\frac{\partial h(\boldsymbol{I}(\boldsymbol{p}))}{\partial p_j}\right|_{\boldsymbol{p}=\boldsymbol{0}} = \omega_j^{(0)} \quad \forall\ j = 1,\dots,n\,,$$

where $\boldsymbol{I} = \boldsymbol{p} + \boldsymbol{I}^\star$. Obviously, the so called action–frequency map in the integrable approximation, i.e., $\boldsymbol{I}^\star \mapsto \boldsymbol{\omega}^{(0)}$, can be inverted because the Hessian of h is non-degenerate. Therefore, the angular velocity vector $\boldsymbol{\omega}^{(0)}$ can be used instead of $\boldsymbol{I}^\star$ in order to parameterize the whole Hamiltonian. Moreover the Fourier decay of the coefficients with respect to the angles $\boldsymbol{q} = \boldsymbol{\vartheta}$ allows to perform the expansion (7) in such a way that $f_\ell^{(0,s)} = \mathcal{O}(\varepsilon^s)$. In other words, the positive integer parameter K can be chosen in such a way that the superscript s refers at the same time to both the order of magnitude and the trigonometric degree (being $f_\ell^{(0,s)} \in \mathcal{P}_{\ell,sK}$); more details about that can be found in [17].

We are now ready for the description of the (generic) r-th step of the normalization procedure, which defines the Hamiltonian $H^{(r)}$ starting from $H^{(r-1)}$, whose expansion is written as follows:

$$\begin{aligned} H^{(r-1)}(\boldsymbol{p},\boldsymbol{q}) \;=\; & E^{(r-1)} + \boldsymbol{\omega}^{(r-1)}\cdot\boldsymbol{p} + \sum_{s\geq 0}\sum_{\ell\geq 2} f_\ell^{(r-1,s)}(\boldsymbol{p},\boldsymbol{q}) \\ & + \sum_{s\geq r}\sum_{\ell=0}^{1} f_\ell^{(r-1,s)}(\boldsymbol{p},\boldsymbol{q})\,. \end{aligned} \tag{8}$$

Hereafter, we omit the dependence of the function from the parameters, unless it has some special meaning. Let us assume that some fundamental properties that hold true for $H^{(0)}$ are satisfied also for the expansion above of $H^{(r-1)}$, i.e., $f_\ell^{(r-1,s)} \in \mathcal{P}_{\ell,sK}$ and $f_\ell^{(r-1,s)} = \mathcal{O}(\varepsilon^s)$. Since the r-th normalization step aims to remove the main perturbing terms, that are $f_0^{(r-1,r)}$ and $f_1^{(r-1,r)}$, we introduce a first generating function $\chi_1^{(r)}$, that is determined by solving the following (first) homological equation:

$$\left\{\boldsymbol{\omega}^{(r-1)}\cdot\boldsymbol{p},\,\chi_1^{(r)}\right\}+f_0^{(r-1,\,r)}(\boldsymbol{q})=\langle f_0^{(r-1,r)}(\boldsymbol{q})\rangle_{\boldsymbol{q}}\,. \tag{9}$$

Since $f_0^{(r-1,\,r)}\in\mathcal{P}_{0,rK}$, its expansion is written as

$$f_0^{(r-1,\,r)}(\boldsymbol{q})=\sum_{|\boldsymbol{k}|\le rK}c_{\boldsymbol{k}}\exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{q}\right),$$

where the complex coefficients are such that $c_{-\boldsymbol{k}}=\bar{c}_{\boldsymbol{k}}$. Therefore, one can easily check that the first homological equation (9) is solved by putting $\langle f_0^{(r-1,r)}(\boldsymbol{q})\rangle_{\boldsymbol{q}}=c_{\boldsymbol{0}}$ and

$$\chi_1^{(r)}(\boldsymbol{q})=\sum_{0<|\boldsymbol{k}|\le rK}\frac{c_{\boldsymbol{k}}\exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{q}\right)}{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)}}\,. \tag{10}$$

In order to preserve the validity of the solution above, of course, we have to require that none of the divisors can eventually vanish; thus we assume the following non-resonance condition:

$$\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)}\neq 0\quad\forall\,0<\boldsymbol{k}\le rK\,. \tag{11}$$

The first half of the r-th normalization step is completed by introducing $\hat{H}^{(r)}=\exp\left(\mathcal{L}_{\chi_1^{(r)}}\right)H^{(r-1)}$. Such an intermediate Hamiltonian can be written in a form similar to formula (8), i.e.,

$$\hat{H}^{(r)}(\boldsymbol{p},\boldsymbol{q})=E^{(r)}+\boldsymbol{\omega}^{(r)}\cdot\boldsymbol{p}+\sum_{s\ge 0}\sum_{\ell\ge 2}\hat{f}_\ell^{(r,\,s)}(\boldsymbol{p},\boldsymbol{q})+\sum_{s\ge r}\sum_{\ell=0}^{1}\hat{f}_\ell^{(r,\,s)}(\boldsymbol{p},\boldsymbol{q})\,, \tag{12}$$

where the recursive definitions of the new summands $\hat{f}_\ell^{(r,\,s)}$ (in terms of $f_\ell^{(r-1,\,s)}$) can be given by exploiting the linearity of the Lie series and by separating the functions according to the different classes $\mathcal{P}_{\ell,sK}$ they belong to. We think it is convenient to formulate these definitions in a rather unconventional way, by using a notation similar to that commonly used in the programming languages; in our opinion, such a choice should make easier the translation of the formal algorithm in any code to be executed in a computational environment. For this purpose, we first define[5] $\hat{f}_\ell^{(r,\,s)}(\boldsymbol{p},\boldsymbol{q})=f_\ell^{(r-1,\,s)}(\boldsymbol{p},\boldsymbol{q})\;\forall\,\ell\ge 0,\;s\ge 0$. Then, *by abuse of notation*, we update $\lfloor s/r\rfloor$ times the definition of the terms $\hat{f}_\ell^{(r,\,s)}$ appearing in the expansion of the new Hamiltonian according to the following rule:

$$\hat{f}_{\ell-j}^{(r,\,s+jr)}\hookleftarrow\frac{1}{j!}\mathcal{L}^j_{\chi_1^{(r)}}f_\ell^{(r-1,\,s)}\quad\forall\,\ell\ge 1,\;1\le j\le\ell,\;s\ge 0\,, \tag{13}$$

[5] We remark that $f_\ell^{(r-1,\,s)}$ do not enter in the expansion (8) if $\ell=0,\;1$ and $s<r$. The same applies to the terms $\hat{f}_\ell^{(r,\,s)}$ that do not make part of the expression of $\hat{H}^{(r)}$, which is written in (12). However, the recursive definitions described in the present subsection are such that $f_\ell^{(r-1,\,s)}=\hat{f}_\ell^{(r,\,s)}=0$ $\forall\,0\le s<r$ when $\ell=0,\;1$.

where with the notation $a \hookleftarrow b$ we mean that the quantity a is redefined so as to be equal to $a + b$. Moreover, there is a last additional contribution that is due to the application of the Lie series to the Hamiltonian $H^{(r-1)}$, and in order to take it into account we write

$$\hat{f}_0^{(r,r)} \hookleftarrow \mathcal{L}_{\chi_1^{(r)}} \boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p} \,. \tag{14}$$

However, because of the homological equation (9), we can finally put $\hat{f}_0^{(r,r)} = 0$ and update the constant energy value so that

$$E^{(r)} = E^{(r-1)} + \langle f_0^{(r-1,r)} \rangle_{\boldsymbol{q}} \,. \tag{15}$$

At this point, it is important to remark that the angular average of the remaining perturbing term that is $\mathcal{O}(\varepsilon^r)$, i.e., $\langle \hat{f}_1^{(r,r)} \rangle_{\boldsymbol{q}}$ is exactly of the same type as $\boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p}$ (this means that both of them are linear with respect to the actions and do not depend on the angles). Therefore, it is useful to update also the angular velocity vector[6] by joining together these two terms. This can be done, by redefining

$$\boldsymbol{\omega}^{(r)} \cdot \boldsymbol{p} = \boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p} + \langle \hat{f}_1^{(r,r)} \rangle_{\boldsymbol{q}} \tag{16}$$

and

$$\hat{f}_1^{(r,r)} = \hat{f}_1^{(r,r)} - \langle \hat{f}_1^{(r,r)} \rangle_{\boldsymbol{q}} \,. \tag{17}$$

Let us recall that all the terms $\hat{f}_\ell^{(r,s)}$ that appear in formula (12) are organized so that they belong to different classes of functions. In order to prove that these structures are suitably preserved by the normalization algorithm, the following statement is essential.

Lemma 1 *Let us consider two generic functions $g \in \mathcal{P}_{\ell,sK}$ and $h \in \mathcal{P}_{m,rK}$, where K is a fixed positive integer number. Then, the following inclusion property holds true*[7]*:*

$$\{g, h\} = \mathcal{L}_h \, g \in \mathcal{P}_{\ell+m-1,(r+s)K} \quad \forall\, \ell,\, m,\, r,\, s \in \mathbb{N} \,.$$

The proof is omitted, because it can be obtained as a straightforward consequence of the definition of the Poisson brackets. By applying repeatedly the lemma above and a trivial induction argument to formulæ (13)–(17), one can easily prove that $E^{(r)} \in \mathcal{P}_{0,0}$ and $\hat{f}_\ell^{(r,s)} \in \mathcal{P}_{\ell,sK}$ for all the terms of type $\hat{f}_\ell^{(r,s)}$ that appear in formula (12). Moreover, it can be ensured that $\left| E^{(r)} - E^{(r-1)} \right| = \mathcal{O}(\varepsilon^r)$ and $\hat{f}_\ell^{(r,s)} = \mathcal{O}(\varepsilon^s)$, if the same relation is assumed to be true at the end of the previous normalization step, i.e., $f_\ell^{(r-1,s)} = \mathcal{O}(\varepsilon^s)$.

---

[6] We emphasize that this is one of the main differences with respect to the original proof scheme designed by Kolmogorov, where the angular velocity vector is kept fixed at every normalization step (see [5] for a fully consistent translation of such an approach, that is implemented by using the Lie series technique).

[7] The statement can be considered as valid also in the trivial case with $\ell = m = 0$, by enlarging the definition of the classes of functions so that $\mathcal{P}_{-1,sK} = \{0\}\ \forall\, s \in \mathbb{N}$.

In order to complete the r-th normalization step, we have to remove the remaining perturbing term that is $\mathcal{O}(\varepsilon^r)$ and appears in the expansion (12) of Hamiltonian $\hat{H}^{(r)}$, i.e., $\hat{f}_1^{(r,r)}$. For such a purpose, we determine a second generating function $\chi_2^{(r)}$, by solving the following (second) homological equation:

$$\left\{\boldsymbol{\omega}^{(r)} \cdot \boldsymbol{p},\ \chi_2^{(r)}\right\} + \hat{f}_1^{(r,r)}(\boldsymbol{p},\boldsymbol{q}) = 0\ . \tag{18}$$

We can deal with the equation above in a very similar way with respect to what has been done for the first homological equation (9). In fact, the solution of (18) can be written as follows:

$$\chi_2^{(r)}(\boldsymbol{p},\boldsymbol{q}) = \sum_{|\boldsymbol{j}|=1} \sum_{0<|\boldsymbol{k}|\le rK} \frac{c_{\boldsymbol{j},\boldsymbol{k}}\, \boldsymbol{p}^{\boldsymbol{j}} \exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{q}\right)}{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r)}}\ , \tag{19}$$

where the expansion of the perturbing term $\hat{f}_1^{(r,r)} \in \mathcal{P}_{1,rK}$ is of type

$$\hat{f}_1^{(r,r)}(\boldsymbol{p},\boldsymbol{q}) = \sum_{|\boldsymbol{j}|=1} \sum_{0<|\boldsymbol{k}|\le rK} c_{\boldsymbol{j},\boldsymbol{k}}\, \boldsymbol{p}^{\boldsymbol{j}} \exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{q}\right)\ .$$

Let us recall that the angular average of $\hat{f}_1^{(r,r)}$ is equal to zero, because of the redefinition (17). Of course, the solution written in (19) is valid provided that the following non-resonance condition is satisfied:

$$\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r)} \neq 0 \quad \forall\ 0 < \boldsymbol{k} \le rK\ . \tag{20}$$

Finally, $H^{(r)} = \exp\left(\mathcal{L}_{\chi_2^{(r)}}\right)\hat{H}^{(r)}$ is the new Hamiltonian that is defined by the canonical transformation of coordinates that is introduced by the r-th normalization step. Also the expansion of such a Hamiltonian can be written in a form similar to (8), i.e.,

$$H^{(r)}(\boldsymbol{p},\boldsymbol{q}) = E^{(r)} + \boldsymbol{\omega}^{(r)}\cdot\boldsymbol{p} + \sum_{s\ge 0}\sum_{\ell\ge 2} f_\ell^{(r,s)}(\boldsymbol{p},\boldsymbol{q}) + \sum_{s\ge r+1}\sum_{\ell=0}^{1} f_\ell^{(r,s)}(\boldsymbol{p},\boldsymbol{q})\ . \tag{21}$$

In this case too, the recursive definitions of the new summands $f_\ell^{(r,s)}$ can be given by exploiting the linearity of the Lie series and by separating the functions according to the different classes they belong to. Let us start by introducing $f_\ell^{(r,s)}(\boldsymbol{p},\boldsymbol{q}) = \hat{f}_\ell^{(r,s)}(\boldsymbol{p},\boldsymbol{q})\ \forall\ \ell \ge 0,\ s \ge 0$. By a new abuse of notation, we update many times the definition of the terms appearing in the expansion of Hamiltonian $H^{(r)}$ according to the following rule:

$$f_\ell^{(r,s+jr)} \hookleftarrow \frac{1}{j!}\mathcal{L}^j_{\chi_2^{(r)}} \hat{f}_\ell^{(r,s)} \quad \forall\ \ell \ge 2,\ j \ge 1,\ s \ge 0 \text{ or } \forall\ \ell = 0, 1,\ j \ge 1,\ s > r\ . \tag{22}$$

In order to take into account also the summands that are generated by the application of the Lie series $\exp\left(\mathcal{L}_{\chi_2^{(r)}}\right)$ to both the terms $\boldsymbol{\omega}^{(r)} \cdot \boldsymbol{p}$ and $f_1^{(r,r)}(\boldsymbol{p}, \boldsymbol{q})$, we add the prescription

$$f_1^{(r,\,(j+1)r)} \hookleftarrow \frac{j}{(j+1)!} \mathcal{L}^j_{\chi_2^{(r)}} \hat{f}_\ell^{(r,r)} \quad \forall\, j \geq 1\,, \tag{23}$$

where we make use of formula (18). Also the last redefinition, i.e.,

$$f_1^{(r,r)} = 0\,, \tag{24}$$

is a straightforward consequence of the second homological equation. By applying again Lemma 1 and a trivial induction argument to formulæ (22)–(23), one can easily prove that $f_\ell^{(r,\,s)} \in \mathcal{P}_{\ell,\,sK}$ for all the summands $f_\ell^{(r,\,s)} = \mathcal{O}(\varepsilon^s)$ that appear in formula (21).

This final remark ends the description of the r-th normalization step of the algorithm that can be iterated so as to determine the next Hamiltonian $H^{(r+1)}$, starting from $H^{(r)}$, and so on.

Let us add a few further comments about the algorithm constructing the Kolmogorov normal form in order to understand its applicability. In practice, one is often interested in determining an approximation up to a fixed order, say $R_{\mathrm{I}} \in \mathbb{N}$, of the motions travelling an invariant KAM torus. For this purpose, starting from $H^{(0)}$, one has to preliminarly compute the Taylor-Fourier truncated expansions of the following type, for all the Hamiltonian $H^{(r)}$ that are introduced by the normalization algorithm with $r = 1, \ldots, R_{\mathrm{I}}$:

$$H^{(r)}(\boldsymbol{p}, \boldsymbol{q}) \simeq E^{(r)} + \boldsymbol{\omega}^{(r)} \cdot \boldsymbol{p} + \sum_{s=0}^{R_{\mathrm{I}}} \sum_{\ell=0}^{\ell_{\max}} f_\ell^{(r,\,s)}(\boldsymbol{p}, \boldsymbol{q})\,, \tag{25}$$

where all the terms that are $o\left(\varepsilon^{R_{\mathrm{I}}}\right)$ or of polynomial degree larger than $\ell_{\max}$ with respect to the actions[8] have been neglected. Let us recall that the algorithm works in such a way to define $f_\ell^{(r,\,s)} = 0\, \forall\, \ell = 0, 1\,,\ 0 \leq s \leq r$. When the first R_{I} normalization steps are performed, all the generating functions $\chi_1^{(r)}$ and $\chi_2^{(r)}\ \forall\, r = 1, \ldots, R_{\mathrm{I}}$ are fully determined. Their composition allows to compute the expansion of ψ_ε that enters in the definition of the semi-analytic scheme of integration (5) and is truncated,

[8] In the practical applications, it is very common to truncate this kind of Taylor series expansions up to a finite degree. In this framework, it is important to remark that the upper limit on the degree in actions is preserved by the Lie series having $\chi_1^{(r)} \in \mathcal{P}_{0,rK}$ and $\chi_2^{(r)} \in \mathcal{P}_{1,rK}$ as generating functions. This can be easily checked by applying repeatedly Lemma 1, that can be used also to prove that just functions of type $f_\ell^{(r,\,s)}$ with $\ell \leq R_{\mathrm{I}} + 1$ are involved in the definitions of $\chi_1^{(r)}$ and $\chi_2^{(r)}\ \forall\, r = 1, \ldots, R_{\mathrm{I}}$. In other terms, this means that the request of determining an approximation up to a fixed order of magnitude $\mathcal{O}\left(\varepsilon^{R_{\mathrm{I}}}\right)$ (for what concerns the canonical transformation that conjugates some orbits to an invariant torus) yields in a fully consistent way also a truncation limit on the polynomial degree in the actions.

once again, so as to neglect all the summands that are $o\left(\varepsilon^{R_1}\right)$. Therefore, the wanted approximation of the motions travelling an invariant KAM torus up to a fixed order of magnitude $\mathcal{O}\left(\varepsilon^{R_1}\right)$ can be provided by the scheme (5) where also the normal form Hamiltonian $\mathcal{K}$ is replaced by $H^{(R_1)}$, which requires $\ell_{\max}\left(R_1+1\right)^2$ functions of type $f_\ell^{(r,\,s)} \in \mathcal{P}_{\ell,sK}$ to be determined. Since their expansions in Taylor-Fourier series are finite (recall definition (6)), all their coefficients are *representable on a computer* (that is equipped with a large enough memory). Therefore, it is *finite* also the number of elementary operations that are defined by the Poisson brackets prescribed by normalization algorithm. The same conclusion applies also for the aforementioned expansion of the canonical transformation ψ_ε . As a whole, we can conclude that the wanted approximation of the motions travelling an invariant KAM torus is *explicitly computable*, because the total amount of operations that are defined by the normalization algorithm is *finite*.

2.4 On the Convergence of the Algorithm Constructing the Kolmogorov Normal Form

In the present context, it is useful to introduce another version of the KAM theorem.

Proposition 1 *Consider the family of Hamiltonians $H^{(0)}(\boldsymbol{p},\boldsymbol{q};\boldsymbol{\omega}^{(0)})$ of the type described in (7). Those functions are defined so that $H^{(0)}:\ \mathcal{A}\times\mathbb{T}^n\times\mathcal{U}\mapsto\mathbb{R}$, where both $\mathcal{A}$ and $\mathcal{U}$ are open subsets of $\mathbb{R}^n$, being $\mathbf{0}\in\mathcal{A}$ and $\mathcal{U}$ bounded. Therefore, $(\boldsymbol{p},\boldsymbol{q})$ are action-angle canonical coordinates and the family of Hamiltonians is parameterized with respect to $\boldsymbol{\omega}^{(0)}\in\mathcal{U}$. Let us also assume that for some fixed and positive values of $K\in\mathbb{N}$, $\varepsilon\in\mathbb{R}$ and $E\in\mathbb{R}$, the following inequalities are satisfied by the functions $f_\ell^{(0,s)}\in\mathcal{P}_{\ell,sK}$:*

$$\sup_{(\boldsymbol{p},\boldsymbol{q};\boldsymbol{\omega}^{(0)})\in\mathcal{A}\times\mathbb{T}^n\times\mathcal{U}}\left|f_\ell^{(0,s)}(\boldsymbol{p},\boldsymbol{q};\boldsymbol{\omega}^{(0)})\right|\le E\,\varepsilon^s \tag{26}$$

$\forall\ s\ge 1,\ \ell\ge 0$ *and* $\forall\ \ell\ge 2$ when $s=0$.

Then, there is a positive $\varepsilon^\star$ such that for $0\le\varepsilon<\varepsilon^\star$ the following statement holds true: there exists a non-resonant set $\mathcal{U}^{(\infty)}\subset\mathcal{U}$ such that the Lebesgue measure μ of the complementary set $\mathcal{U}\setminus\mathcal{U}^{(\infty)}$ goes to zero for $\varepsilon\to 0$ and for each $\boldsymbol{\omega}^{(0)}\in\mathcal{U}^{(\infty)}$ there is an analytic canonical transformation $(\boldsymbol{p},\boldsymbol{q})=\psi^{(\infty)}_{\varepsilon;\boldsymbol{\omega}^{(0)}}(\boldsymbol{P},\boldsymbol{Q})$ leading the Hamiltonian to the normal form

$$H^{(\infty)}(\boldsymbol{P},\boldsymbol{Q};\boldsymbol{\omega}^{(0)})=E^{(\infty)}+\boldsymbol{\omega}^{(\infty)}\cdot\boldsymbol{P}+\sum_{s\ge 0}\sum_{\ell\ge 2}f_\ell^{(\infty,\,s)}(\boldsymbol{P},\boldsymbol{Q};\boldsymbol{\omega}^{(0)})\ , \tag{27}$$

where $f_\ell^{(\infty,\,s)}\in\mathcal{P}_{\ell,sK}\ \forall\ s\ge 0,\ \ell\ge 2$ and $E^{(\infty)}$ is a finite real value fixing the constant energy level that corresponds to the invariant torus $\left\{(\boldsymbol{P}=\mathbf{0},\ \boldsymbol{Q}\in\mathbb{T}^n)\right\}$.

Moreover, the canonical change of coordinates is close to the identity in the sense that $\left\| \psi^{(\infty)}_{\varepsilon;\boldsymbol{\omega}^{(0)}}(\boldsymbol{P},\boldsymbol{Q}) - (\boldsymbol{P},\boldsymbol{Q}) \right\| = \mathcal{O}(\varepsilon)$ *and the same applies also to both the energy level and the detuning of the angular velocity vector (that are* $\left| E^{(\infty)} - E^{(0)} \right| = \mathcal{O}(\varepsilon)$ *and* $\left\| \boldsymbol{\omega}^{(\infty)} - \boldsymbol{\omega}^{(0)} \right\| = \mathcal{O}(\varepsilon)$*, respectively).*

The statement above is substantially equivalent to that claimed in theorem C of [36] (which is considered as a classical version of the KAM theorem, in the very own words of the Author, J. Pöschel). The proof of Proposition 1 can be obtained by adapting the one described in [7] in such a way to prove the convergence of the normalization algorithm described in the previous Sect. 2.3. Indeed, both articles [7, 36] deal only with the more complicate proof of existence for invariant tori that are of dimension smaller than the number n of degrees of freedom and have elliptic character in the transverse directions. The construction of the normal form corresponding to such a type of invariant manifolds will be widely discussed in the next Sect. 3. As a main difference between the approaches developed in those two works, let us recall that the proof adopted in [36] is based on a fast convergence scheme of quadratic type (a so called Newton-like method, where perturbing terms of order of magnitude $\mathcal{O}\left(\varepsilon^{2^{r-1}}\right)$ are removed during the r-th normalization step). Such a technique has been adopted since the very first works in KAM theory, but the convergence of the normalization algorithm described in Sect. 2.3 is of linear type (because perturbing terms of order of magnitude $\mathcal{O}\left(\varepsilon^{r}\right)$ are removed during the r-th normalization step). The latter is in a better position for the applications[9] and a complete proof of the KAM theorem adopting a convergence method of linear type is available since the last decade of the past century (see [17]). Rather curiously, the best way to translate the algorithm constructing the Kolmogorov normal form in a computer-assisted proof requires to join the convergence scheme of linear type (in order to explicitly perform on a computer the largest possible number R_{I} of preliminary steps) with that of quadratic type (that provides a statement of KAM theorem that is very suitable to rigorously complete the proof). This is one of the main conclusions discussed in a recent work (see [40]).

The statement of Proposition 1 highlights that we are forced to provide a result which holds true with respect to the Lebesgue measure, because we have *chosen* to adopt a version of the normalization algorithm where the angular velocity vector is allowed to vary at each step (recall formula (16) that defines the detuning shift $\boldsymbol{\omega}^{(r)} - \boldsymbol{\omega}^{(r-1)}$). This means that such a statement has to be understood in a probabilistic sense, because we are not able to describe in detail the structure of the non-resonant set $\mathcal{U}^{(\infty)}$. In particular, for a fixed initial value of the angular velocity vector $\boldsymbol{\omega}^{(0)}$ we cannot establish whether the specific Hamiltonian $H^{(0)}(\boldsymbol{p},\boldsymbol{q};\boldsymbol{\omega}^{(0)})$ can be brought in Kolmogorov normal form or not. We can just claim that the normalization algorithm can converge with a rate of success (i.e., $\mu\left(\mathcal{U} \setminus \mathcal{U}^{(\infty)}\right) / \mu(\mathcal{U})$) that gets larger and larger when the small parameter ε which rules the size of the

[9] This is the main reason why the present work is focusing on approaches based on a convergence scheme of linear type. A very far from being exhaustive list of references to applications of KAM theorem has been discussed in the Introduction.

perturbation is decreasing. On the other hand, we can characterize very well the set of the final values of the angular velocities, i.e., $\left\{\boldsymbol{\omega}^{(\infty)}\left(\boldsymbol{\omega}^{(0)}\right) \,:\, \boldsymbol{\omega}^{(0)} \in \mathcal{U}^{(\infty)}\right\}$, because they are Diophantine. In the recent work [37], the problem of the convergence of this type of normalization algorithms is revisited so as to provide a KAM-like statement. It is proved by fixing since the beginning the final value $\boldsymbol{\omega}^{(\infty)}$ and its non-resonance properties (that allow to explicitly solve the homological equations at every step of the algorithm). Moreover, the total detuning $\boldsymbol{\omega}^{(\infty)} - \boldsymbol{\omega}^{(0)}$ is given in terms of series whose coefficients are defined in a recursive way. Therefore, the convergence of the normalization algorithm is ensured (provided that the perturbation is small enough), the total detuning is estimated explicitly, while the exact location of $\boldsymbol{\omega}^{(0)}$ remains partially unknown, because it can be determined just by iterating *ad infinitum* the computational procedure.

3 Construction of Invariant Elliptic Tori by a Normal Form Algorithm

Elliptic tori are compact invariant manifolds of dimension smaller than the maximal one, that is equal to the number n of degrees of freedom. In order to better imagine them, let us consider a phase space $\mathcal{F}$ that is endowed by the canonical coordinates $(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{X}, \boldsymbol{Y})$, where $(\boldsymbol{P}, \boldsymbol{Q}) \in \mathbb{R}^{n_1} \times \mathbb{T}^{n_1}$ are action-angle variables and also $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathbb{R}^{n_2} \times \mathbb{R}^{n_2}$ denote pairs of conjugate (momenta and) coordinates, while $n = n_1 + n_2$ with both n_1 and n_2 positive integers. Let us consider a Hamiltonian of the following type:

$$\mathcal{H}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{\omega} \cdot \boldsymbol{P} + \sum_{j=1}^{n_2} \frac{\Omega_j}{2}(X_j^2 + Y_j^2) + \mathcal{R}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{X}, \boldsymbol{Y}) \,,$$

where $\boldsymbol{\Omega} \in \mathbb{R}^{n_2}$ and the remainder $\mathcal{R}$ is an analytic function with respect to its arguments and is such that $\mathcal{R}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{X}, \boldsymbol{Y}) = o\left(\|\boldsymbol{P}\| + \|(\boldsymbol{X}, \boldsymbol{Y})\|^2\right)$, when $(\boldsymbol{P}, \boldsymbol{X}, \boldsymbol{Y}) \to (\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})$. It is easy to check that

$$(\boldsymbol{P}(t), \boldsymbol{Q}(t), \boldsymbol{X}(t), \boldsymbol{Y}(t)) = \left(\boldsymbol{0}, \boldsymbol{Q}(0) + \boldsymbol{\omega} t, \boldsymbol{0}, \boldsymbol{0}\right) \tag{28}$$

is a solution of Hamilton equations, since the function $\mathcal{H}$, except for its main part, contains terms of type $\mathcal{O}(\|\boldsymbol{P}\|^2)$, $\mathcal{O}(\|\boldsymbol{P}\|\|(\boldsymbol{X}, \boldsymbol{Y})\|)$ and $\mathcal{O}(\|(\boldsymbol{X}, \boldsymbol{Y})\|^3)$ only. Because of this remark, it is evident that the n_1–dimensional manifold $\left\{(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{X}, \boldsymbol{Y}) \,:\, \boldsymbol{P} = \boldsymbol{0},\ \boldsymbol{Q} \in \mathbb{T}^{n_1},\ \boldsymbol{X} = \boldsymbol{Y} = \boldsymbol{0}\right\}$ is invariant. The elliptical character is given by the fact that, in the remaining $n_2 = n - n_1$ degrees of freedom, the dynamics that is transverse with respect to such an invariant manifold is given by the composition of n_2 oscillatory motions whose periods tend to the values $2\pi/\Omega_1, \ldots, 2\pi/\Omega_{n_2}$, in the limit of $(\boldsymbol{P}, \boldsymbol{X}, \boldsymbol{Y}) \to (\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})$. Of course, this is due to the occurrence of the term

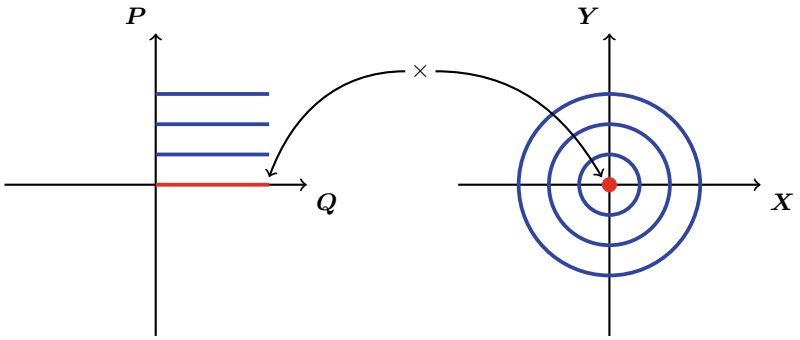


Fig. 1 Schematic representation of an elliptic torus. The orbit is given by the cartesian product of the two invariant surfaces that are marked in red, that are a torus (in the left panel) and a stable equilibrium point (on the right, resp.)

$\sum_{j=1}^{n_2} \Omega_j (X_j^2 + Y_j^2)/2$ which overwhelms the effect of the remainder $\mathcal{R}$ in the so called limit of small oscillations.

The name of elliptic torus is well justified by all the remarks discussed since the beginning of the present section. A schematic representation of such kind of invariant manifolds is sketched in Fig. 1.

3.1 Algorithmic Construction of the Normal Form for Elliptic Tori

Since we aim at introducing the algorithm constructing the normal form for invariant elliptic tori in a way that is as much as possible coherent with what we have already done in Sect. 2.3 for KAM tori, we prefer to not adopt canonical coordinates $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{n_1} \times \mathbb{T}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_2}$ that are substantially the ones considered in the discussion at the beginning of the present section. Indeed, we think it is convenient to introduce the so called action-angle coordinates for harmonic oscillators, in order to replace the polynomial ones, that are $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{n_2} \times \mathbb{R}^{n_2}$; this means that we define $(\boldsymbol{J}, \boldsymbol{\varphi}) \in \left(\mathbb{R}^{n_2}_+ \cup \{\boldsymbol{0}\}\right) \times \mathbb{T}^{n_2}$ so that $x_j = \sqrt{2J_j} \cos \varphi_j$ and $y_j = \sqrt{2J_j} \sin \varphi_j$, where this change of coordinates is canonical $\forall\ j = 1, \ldots, n_2$.

We are now ready to introduce classes of functions depending on $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi}) \in \mathbb{R}^{n_1} \times \mathbb{T}^{n_1} \times \left(\mathbb{R}^{n_2}_+ \cup \{\boldsymbol{0}\}\right) \times \mathbb{T}^{n_2}$ in a very similar way to what has been previously done. For some fixed positive integer K we introduce the distinct classes of functions $\widehat{\mathfrak{P}}_{\hat{m}, \hat{\ell}, sK}$, with integers $\hat{m},\ \hat{\ell},\ s \geq 0$; any generic function $g \in \widehat{\mathfrak{P}}_{\hat{m}, \hat{\ell}, sK}$ can be written as

$$g\,(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi}) = \sum_{\substack{\boldsymbol{m}\in\mathbb{N}^{n_1}\\ |\boldsymbol{m}|=\hat{m}}} \sum_{\substack{\boldsymbol{\ell}\in\mathbb{N}^{n_2}\\ |\boldsymbol{\ell}|=\hat{\ell}}} \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^{n_1}\\ |\boldsymbol{k}|\le sK}} \sum_{\substack{\hat{\ell}_j=-\ell_j,\,-\ell_j+2,\ldots,\,\ell_j\\ j=1,\,\ldots\,,n_2}} c_{\boldsymbol{m},\boldsymbol{\ell},\boldsymbol{k},\hat{\boldsymbol{\ell}}}\,\boldsymbol{p}^{\boldsymbol{m}}\left(\sqrt{\boldsymbol{J}}\right)^{\boldsymbol{\ell}} \exp\left[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{q}+\hat{\boldsymbol{\ell}}\cdot\boldsymbol{\varphi})\right], \tag{29}$$

where the complex coefficients are such that $c_{\boldsymbol{m},\boldsymbol{\ell},-\boldsymbol{k},-\hat{\boldsymbol{\ell}}} = \bar{c}_{\boldsymbol{m},\boldsymbol{\ell},\boldsymbol{k},\hat{\boldsymbol{\ell}}}$, then the codomain of any $g \in \widehat{\mathfrak{P}}_{\hat{m},\,\hat{\ell},\,sK}$ is included in $\mathbb{R}$. Let us emphasize that, in each term appearing in the Taylor-Fourier expansion of a function belonging to a class of type $\widehat{\mathfrak{P}}_{\hat{m},\,\hat{\ell},\,sK}$, the indexes vector $(\hat{\ell}_1\,,\,\ldots\,,\,\hat{\ell}_{n_2})$ are subject to special restrictions that are inherited by the corresponding polynomial structure with respect to the variables $(\boldsymbol{x},\boldsymbol{y}) = \left(\sqrt{2\boldsymbol{J}}\cos\boldsymbol{\varphi},\sqrt{2\boldsymbol{J}}\sin\boldsymbol{\varphi}\right)$. In fact, they are such that $\forall\ j = 1,\,\ldots\,,n_2$ the j-th component of the Fourier harmonic $\hat{\ell}_j$ must have the same parity with respect to the corresponding degree ℓ_j of $\sqrt{J_j}$ and also the inequality $\left|\hat{\ell}_j\right| \le \ell_j$ must be satisfied.[10] Furthermore, we will say that $g \in \mathfrak{P}_{\ell,sK}$ if

$$g \in \bigcup_{\substack{\hat{m}\ge 0,\,\hat{\ell}\ge 0\\ 2\hat{m}+\hat{\ell}=\ell}} \widehat{\mathfrak{P}}_{\hat{m},\hat{\ell},sK}\,. \tag{30}$$

In other words, a function belonging to the class $\mathfrak{P}_{\ell,sK}$ depends on the actions so as to be homogeneous polynomials of total degree ℓ in the square roots of $\boldsymbol{p}$ and $\boldsymbol{J}$, while its Fourier expansion contain harmonics of total trigonometric degree in $\boldsymbol{q}$ that are not larger than sK.

In order to extend the approach described in Sect. 2.3 with the aim to design an efficient algorithm constructing the normal form in the case of elliptic tori, we are also forced to reformulate the Lemma 1 in a suitable version to describe the action of the Poisson brackets on these new classes of functions, that are defined thanks to formulæ (29)–(30). This is made as it follows.

Lemma 2 *Let us consider two generic functions $g \in \mathfrak{P}_{\ell,sK}$ and $h \in \mathfrak{P}_{m,rK}$, where K is a fixed positive integer number. Then,*[11]

$$\{g,h\} = \mathcal{L}_h\, g \in \mathfrak{P}_{\ell+m-2,(r+s)K} \quad \forall\ \ell,\ m,\ r,\ s \in \mathbb{N}\,.$$

[10] When there are variables such that they appear in the Taylor-Fourier expansions of a function so that they follow this kind of restrictions, then they are often said to be of D'Alembert type. This name is given by analogy, because in Celestial Mechanics the secular part of the Hamiltonian perturbing terms due to the interactions between planets shows the same kind of expansions, since they satisfy the so called D'Alembert rules.

[11] The statement can be considered as valid also in the trivial cases with $\ell + m = 0,\,1$, by enlarging the definition of the classes of functions so that $\mathfrak{P}_{-2,sK} = \mathfrak{P}_{-1,sK} = \{0\}\ \forall\, s \in \mathbb{N}$.

Also in this case the proof is omitted, because it can be obtained by simply applying[12] the definition of the Poisson brackets.

As an environment where it is natural to properly define the algorithm constructing the normal form for elliptic tori, let us start to consider a Hamiltonian $\mathcal{H}^{(0)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})$ that can be written in the following way:

$$\begin{aligned}\mathcal{H}^{(0)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi}) = \mathcal{E}^{(0)} + \boldsymbol{\omega}^{(0)} \cdot \boldsymbol{p} + \boldsymbol{\Omega}^{(0)} \cdot \boldsymbol{J} + \sum_{s\geq 0} \sum_{\ell \geq 3} f_\ell^{(0,\,s)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi}) \\ + \sum_{s\geq 1} \sum_{\ell=0}^{2} f_\ell^{(0,\,s)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi}) \,,\end{aligned} \tag{31}$$

where $\mathcal{E}^{(0)} \in \mathfrak{P}_{0,0}$ is a constant[13] and $f_\ell^{(0,\,s)} \in \mathfrak{P}_{\ell, sK}$, being the first upper index related to the normalization step. For instance, in [8] it is shown how to bring an FPU chain of $N + 1$ particles in the form above, by following a procedure that is valid for a generic Hamiltonian problem in the neighborhood of a stable equilibrium point. In other words, the Hamiltonian describing that model can be expanded as $\mathcal{H}^{(0)}$ in (31), with $f_\ell^{(0,\,s)} = 0$ when $s \geq 3$ and $f_\ell^{(0,\,1)} \in \mathfrak{P}_{\ell,K}$, $f_\ell^{(0,\,2)} \in \mathfrak{P}_{\ell,2K}$ $\forall\ \ell \geq 0$, being[14] $K = 2$. This holds true, both for the so called α–model and the β one. Let us also emphasize that the energy value $\mathcal{E}^{(0)}$, the angular velocity vector $\boldsymbol{\Omega}^{(0)} \in \mathbb{R}^{n_2}$ and all the functions $f_\ell^{(0,\,s)}$ depend on $\boldsymbol{\omega}^{(0)} \in \mathbb{R}^{n_1}$ in a parametric way. In order to keep the notation so that it does not get too cumbersome, in the present subsection we do not include $\boldsymbol{\omega}^{(0)}$ among the arguments of the terms appearing in the expansions of the Hamiltonians. Moreover, for a generic problem in the neighborhood of a stable equilibrium point one can also easily show that $f_\ell^{(0,\,s)} = \mathcal{O}(\varepsilon^s)$, where ε is the natural

---

[12] Actually, it looks natural to be doubtful about the fact that Poisson brackets always preserve the restrictions on the Fourier harmonics that must be satisfied by variables of D'Alembert type. However, one can immediately realize that the only tricky case occurs when the Poisson brackets include also the following terms:

$$\begin{aligned}&\frac{\partial\big(\sqrt{J_j}\big)^{|\hat{\ell}_j|}\exp(\mathrm{i}\hat{\ell}_j\varphi_j)}{\partial\varphi_j}\,\frac{\partial\big(\sqrt{J_j}\big)^{|\hat{m}_j|}\exp(\mathrm{i}\hat{m}_j\varphi_j)}{\partial J_j} \\ &-\frac{\partial\big(\sqrt{J_j}\big)^{|\hat{\ell}_j|}\exp(\mathrm{i}\hat{\ell}_j\varphi_j)}{\partial J_j}\,\frac{\partial\big(\sqrt{J_j}\big)^{|\hat{m}_j|}\exp(\mathrm{i}\hat{m}_j\varphi_j)}{\partial\varphi_j} \\ &= \frac{\mathrm{i}}{2}\big(\hat{\ell}_j|\hat{m}_j| - \hat{m}_j|\hat{\ell}_j|\big)\big(\sqrt{J_j}\big)^{|\hat{\ell}_j|+|\hat{m}_j|-2}\exp\big(\mathrm{i}(\hat{\ell}_j+\hat{m}_j)\varphi_j\big)\,.\end{aligned}$$

However, if $\hat{\ell}_j$ and $\hat{m}_j$ have opposite signs then $\big|\hat{\ell}_j + \hat{m}_j\big| \leq \big|\hat{\ell}_j\big| + \big|\hat{m}_j\big| - 2$ (let us remark that the term above vanishes if $\hat{\ell}_j = 0$ or $\hat{m}_j = 0$). In the remaining case (i.e., $\hat{\ell}_j \neq 0$ and $\hat{m}_j \neq 0$ have the same sign), the coefficient $\hat{\ell}_j|\hat{m}_j| - \hat{m}_j|\hat{\ell}_j|$ is always equal to zero.

[13] $\mathcal{E}^{(0)}$ denotes the energy level of the elliptic torus that is invariant in the approximation given by the angular average, i.e., when $f_\ell^{(0,\,s)} = 0\ \forall\, s > 0$.

[14] Setting $K = 2$ is quite natural for Hamiltonian systems close to stable equilibria, see, e.g., [20].

small parameter for this kind of models, because it denotes the first approximation of the distance (expressed in terms of the actions) between the wanted elliptic torus and the stable equilibrium point.

In a strict analogy with what has been done to construct the Kolmogorov normal form, here our main purpose is to eliminate from the Hamiltonian all the terms having total degree less than three in the square root of the actions; by referring to the paradigmatic form described in (31), the unwanted terms are appearing in its last row. Actually, such a goal can be achieved by performing an infinite sequence of canonical transformations, so as to bring the Hamiltonian to the following final normal form:

$$\mathcal{H}^{(\infty)}(\boldsymbol{P},\boldsymbol{Q},\boldsymbol{\Xi},\boldsymbol{\Theta}) = \mathcal{E}^{(\infty)} + \boldsymbol{\omega}^{(\infty)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(\infty)}\cdot\boldsymbol{\Xi} + \sum_{s\geq 0}\sum_{\ell\geq 3} f_\ell^{(\infty,s)}(\boldsymbol{P},\boldsymbol{Q},\boldsymbol{\Xi},\boldsymbol{\Theta})\,, \tag{32}$$

with $f_\ell^{(\infty,s)} \in \mathfrak{P}_{\ell,sK}$ and $\mathcal{E}^{(\infty)} \in \mathfrak{P}_{0,0}$. The motion law $(\boldsymbol{P}(t),\,\boldsymbol{Q}(t),\,\boldsymbol{\Xi}(t),\,\boldsymbol{\Theta}(t)) = \big(\mathbf{0},\,\boldsymbol{Q}_0+\boldsymbol{\omega}^{(\infty)}t,\,\mathbf{0},\,\boldsymbol{\Theta}\big)$ is a solution of the Hamilton equations related to the normal form $\mathcal{H}^{(\infty)}$ and is equivalent[15] to (28). Such a motion law is generated by the initial condition $(\mathbf{0},\,\boldsymbol{Q}_0,\,\mathbf{0},\,\boldsymbol{\Theta})$, is quasi-periodic with an angular velocity vector equal to $\boldsymbol{\omega}^{(\infty)}$ and the corresponding orbit lies on the n_1-dimensional invariant torus $\boldsymbol{P}=\mathbf{0}$, $\boldsymbol{\Xi}=\mathbf{0}$. The energy level of such a manifold is $H^{(\infty)}(\mathbf{0},\,\boldsymbol{Q},\,\mathbf{0},\,\boldsymbol{\Theta}) = \mathcal{E}^{(\infty)}$. Moreover, it is elliptic in the sense that the transverse dynamics in a neighborhood of the invariant torus itself is given by oscillations whose corresponding angular velocity vector is approaching $\boldsymbol{\Omega}^{(\infty)}$ in the limit of $\|(\boldsymbol{P},\boldsymbol{\Xi})\|$ going to zero.

Also in the present case, that is concerning the elliptic tori, the formal algorithm for the construction of the normal form is composed by a sequence of canonical transformations, defined using the formalism of Lie series. We can summarize the r-th normalization step, by giving the formula defining the canonical change of coordinates that transforms the intermediate Hamiltonian $\mathcal{H}^{(r-1)}$ into the subsequent $\mathcal{H}^{(r)}$. The expansion of the former is of the following type:

$$\begin{aligned}\mathcal{H}^{(r-1)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi}) =& \mathcal{E}^{(r-1)} + \boldsymbol{\omega}^{(r-1)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(r-1)}\cdot\boldsymbol{J} + \sum_{s\geq 0}\sum_{\ell\geq 3} f_\ell^{(r-1,s)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi}) \\ &+ \sum_{s\geq r}\sum_{\ell=0}^{2} f_\ell^{(r-1,\,s)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi})\,,\end{aligned} \tag{33}$$

[15] We remark that $\dot{\boldsymbol{P}} = \big\{\boldsymbol{P}\,,\,\mathcal{H}^{(\infty)}\big\} = \mathbf{0}$ and $\dot{\boldsymbol{\Xi}} = \big\{\boldsymbol{\Xi}\,,\,\mathcal{H}^{(\infty)}\big\} = \mathbf{0}$ when $\boldsymbol{P}=\mathbf{0}$ and $\boldsymbol{\Xi}=\mathbf{0}$. Because of the well known degeneracy of the change of coordinates $(\boldsymbol{X},\boldsymbol{Y}) = \big(\sqrt{2\boldsymbol{\Xi}}\cos\boldsymbol{\Theta},\sqrt{2\boldsymbol{\Xi}}\sin\boldsymbol{\Theta}\big)$, all the set $\big\{(\boldsymbol{\Xi}=\mathbf{0},\boldsymbol{\Theta}\in\mathbb{T}^{n_2})\big\}$ correspond to a single point $\big\{(\boldsymbol{X}=\mathbf{0},\boldsymbol{Y}=\mathbf{0})\big\}$ of the reduced phase space that considers just the last n_2 degrees of freedom. By the way, we emphasize that such a degeneracy is completely harmless in the framework we have adopted. In order to conclude the check of the solution of the Hamilton equations related to the normal form $\mathcal{H}^{(\infty)}$ when $\boldsymbol{P}=\mathbf{0}$ and $\boldsymbol{\Xi}=\mathbf{0}$, it is enough to remark that $\dot{\boldsymbol{Q}} = \big\{\boldsymbol{Q}\,,\,\mathcal{H}^{(\infty)}\big\} = \boldsymbol{\omega}^{(\infty)}$.

being $f_\ell^{(r-1,s)} \in \mathfrak{P}_{\ell,sK}$ and $\mathcal{E}^{(r-1)} \in \mathfrak{P}_{0,0}$, i.e., it is a constant referring to the level of the energy in the approximation that is valid up to terms $\mathcal{O}(\varepsilon^r)$. Let us emphasize that the starting Hamiltonian $\mathcal{H}^{(0)}$ written in Eq. (31) is exactly in the form (33) with $r = 1$. The conjugacy relation which allows to write the Hamiltonian defined at the end of the r-th normalization step as a function of the previous one is given by

$$\mathcal{H}^{(r)} = \mathcal{H}^{(r-1)} \circ \exp\left(\mathcal{L}_{\chi_0^{(r)}}\right) \circ \exp\left(\mathcal{L}_{\chi_1^{(r)}}\right) \circ \exp\left(\mathcal{L}_{\chi_2^{(r)}}\right) \circ \mathfrak{D}^{(r)} , \tag{34}$$

where the Lie series[16] operator $\exp\left(\mathcal{L}_{\chi_j^{(r)}}\right) \cdot$ removes the Hamiltonian terms with total degree in the square root of the actions equal to j and with trigonometric degree in the angles $\boldsymbol{q}$ up to rK. Moreover, by a linear canonical transformation $\mathfrak{D}^{(r)}$, the terms that are quadratic in $\sqrt{\boldsymbol{J}}$ and do not depend on both the actions $\boldsymbol{p}$ and the angles $\boldsymbol{q}$ are brought to a diagonal form. At the end of this r-th normalization step, the ineliminable terms that are independent on the angles $\boldsymbol{q}$ and linear either in $\boldsymbol{p}$ or in $\boldsymbol{J}$ are added to the normal form part. This requires to update the angular velocities from $\left(\boldsymbol{\omega}^{(r-1)}, \boldsymbol{\Omega}^{(r-1)}\right)$ to $\left(\boldsymbol{\omega}^{(r)}, \boldsymbol{\Omega}^{(r)}\right)$, that is why in (32) the Hamiltonian in Kolmogorov normal form has new frequency vectors $\boldsymbol{\omega}^{(\infty)}$ and $\boldsymbol{\Omega}^{(\infty)}$.

All the details that properly define how the algorithm actually works are exhaustively described in the following.

First Stage of the *r*-th Normalization Step

In the context of the r-th normalization step, the first stage aims to remove the terms depending just on the angles $\boldsymbol{q}$ up to the trigonometrical degree rK, i.e. the terms collected in $f_0^{(r-1,r)} = \mathcal{O}(\varepsilon^r)$. We determine the generating function $\chi_0^{(r)}$ by solving the homological equation

$$\left\{\boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p},\ \chi_0^{(r)}\right\} + f_0^{(r-1,r)}(\boldsymbol{q}) = \langle f_0^{(r-1,r)}(\boldsymbol{q})\rangle_{\boldsymbol{q}} . \tag{35}$$

Let us remark that the equation above is perfectly equivalent to that in formula (9), because $f_0^{(r-1,r)} \in \mathfrak{P}_{0,rK}$ depends on $\boldsymbol{q}$ only and, therefore, $f_0^{(r-1,r)} \in \mathcal{P}_{0,rK} = \mathfrak{P}_{0,rK}$. Thus, we can write the solution of this new (first) homological equation (35) exactly in the same way as we have done for what concerns (10), i.e., we put $\langle f_0^{(r-1,r)}(\boldsymbol{q})\rangle_{\boldsymbol{q}} = c_{\boldsymbol{0}}$ and

$$\chi_0^{(r)}(\boldsymbol{q}) = \sum_{0<|\boldsymbol{k}|\le rK} \frac{c_{\boldsymbol{k}} \exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{q}\right)}{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)}} , \tag{36}$$

[16] Because of the so called "exchange theorem" (see [14]), the new Hamiltonian $H^{(r)}$ is obtained from the old one, by applying the Lie series to $H^{(r-1)}$ in reverse order with respect to what is written in (34). This is consistent with the order of the discussion in the following subsections: the first stage of the r-th normalization step deals with the canonical transformation generated by $\chi_0^{(r)}$, the second one with $\chi_1^{(r)}$ and the last one with both $\chi_2^{(r)}$ and $\mathfrak{D}^{(r)}$.

being $f_0^{(r-1,\,r)}(\boldsymbol{q}) = \sum_{|\boldsymbol{k}|\leq rK} c_{\boldsymbol{k}} \exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{q}\right)$. Of course, such a solution is certainly valid provided the non-resonance condition (11) is satisfied.

Now, we apply the canonical transformation $\exp \mathcal{L}_{\chi_0^{(r)}}$ to the Hamiltonian which is defined at the end of the $r-1$-th normalization step. By the usual abuse of notation, we choose to rename the new variables as the old ones. This allows to write the transformed Hamiltonian $H^{(\mathrm{I};\,r)} = \exp\left(\mathcal{L}_{\chi_0^{(r)}}\right) H^{(r-1)}$ as follows:

$$\begin{aligned}\mathcal{H}^{(\mathrm{I};\,r)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi}) =& \mathcal{E}^{(r)} + \boldsymbol{\omega}^{(r-1)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(r-1)}\cdot\boldsymbol{J} + \sum_{s\geq 0}\sum_{\ell\geq 3} f_\ell^{(\mathrm{I};\,r,\,s)} \\ &+ \sum_{s\geq r}\sum_{\ell=0}^{2} f_\ell^{(\mathrm{I};\,r,\,s)}\,,\end{aligned} \tag{37}$$

where for the sake of brevity we have omitted to list the arguments of the functions $f_\ell^{(\mathrm{I};\,r,\,s)}$. Let us introduce them in the same unconventional way we have adopted in Sect. 2.3 to describe the algorithm constructing the Kolmogorov normal form. First, we define[17] $f_\ell^{(\mathrm{I};\,r,\,s)} = f_\ell^{(r-1,\,s)}\ \forall\ \ell\geq 0,\ s\geq 0$. By further abuses of notation, we update many times the definition of the terms appearing in the expansion of the new Hamiltonian according to the following rule:

$$f_{\ell-2i}^{(\mathrm{I};\,r,\,s+jr)} \hookleftarrow \frac{1}{j!}\mathcal{L}^j_{\chi_0^{(r)}} f_\ell^{(r-1,\,s)} \quad \forall\ \ell\geq 0,\ 1\leq j\leq \lfloor \ell/2\rfloor,\ s\geq 0\,. \tag{38}$$

By applying repeatedly Lemma 2 and a trivial induction argument to the formula above, one can easily prove that $f_\ell^{(\mathrm{I};\,r,\,s)} \in \mathfrak{P}_{\ell,sK}\ \forall\ \ell\geq 0,\ s\geq 0$. In order to end the description of the first stage of the r-th normalization step, we have to take into account also the effects induced by the homological equation (35). For such a purpose, we finally set $f_0^{(\mathrm{I};\,r,\,r)} = 0$ and we update the approximated value referring to the energy of the wanted elliptic torus exactly in the same way we have done to write formula (15), i.e., we put $\mathcal{E}^{(r)} = \mathcal{E}^{(r-1)} + \langle f_0^{(r-1,\,r)}\rangle_{\boldsymbol{q}}$.

Second Stage of the *r*-th Normalization Step

The second stage of the r-th normalization step acts on the Hamiltonian that is initially expanded as in (37), with the goal to remove the perturbing term which is linear in $\sqrt{\boldsymbol{J}}$ and independent of $\boldsymbol{p}$, i.e., $f_1^{(\mathrm{I};\,r,\,r)}$. Thus, we have to solve the following homological equation:

$$\left\{\boldsymbol{\omega}^{(r-1)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(r-1)}\cdot\boldsymbol{J},\ \chi_1^{(r)}\right\} + f_1^{(\mathrm{I};\,r,\,r)}(\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi}) = 0\,. \tag{39}$$

[17] We remark that the terms $f_\ell^{(r-1,\,s)}$ do not enter in the expansion (33) when $\ell = 0,\ 1,\ 2$ and $s < r$. However, the recursive definitions described in the present subsection are such that all those functions are equal to zero. Keeping in mind this fact allows to write in a rather compact way both formula (38) and the analogous ones in the following.

Let us write the expansion of $f_1^{(\mathrm{I};\,r,\,r)}(\boldsymbol{q},\,\boldsymbol{J},\,\boldsymbol{\varphi})$ as follows:

$$f_1^{(\mathrm{I};\,r,\,r)}(\boldsymbol{q},\,\boldsymbol{J},\,\boldsymbol{\varphi}) = \sum_{0\le \boldsymbol{k}\le rK} \sum_{j=1}^{n_2} \sqrt{J_j}\left[c^{(+)}_{\boldsymbol{k},\,j}\, e^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{q}+\varphi_j)} + c^{(-)}_{\boldsymbol{k},\,j}\, e^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{q}-\varphi_j)}\right], \tag{40}$$

where every coefficients $c^{(+)}_{\boldsymbol{k},\,j} \in \mathbb{C}$ is equal to the complex conjugate of $c^{(-)}_{-\boldsymbol{k},\,j}$ $\forall\, 0 \le \boldsymbol{k} \le rK$, $1 \le j \le n_2$. Therefore, the generating function $\chi_1^{(r)}$ solving Eq. (39) is determined in such a way that

$$\chi_1^{(r)}(\boldsymbol{q},\,\boldsymbol{J},\,\boldsymbol{\varphi}) = \sum_{0\le \boldsymbol{k}\le rK} \sum_{j=1}^{n_2} \frac{\sqrt{J_j}}{\mathrm{i}}\left[\frac{c^{(+)}_{\boldsymbol{k},\,j}\, e^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{q}+\varphi_j)}}{\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)} + \Omega_j^{(r-1)}} + \frac{c^{(-)}_{\boldsymbol{k},\,j}\, e^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{q}-\varphi_j)}}{\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)} - \Omega_j^{(r-1)}}\right]. \tag{41}$$

This expression is well-defined, provided that the frequency vector $\boldsymbol{\omega}^{(r-1)}$ satisfies the so-called first Melnikov non-resonance condition up to order rK (see [29]), i.e.,

$$\min_{\substack{0<|\boldsymbol{k}|\le rK,\\ |\boldsymbol{\ell}|=1}} \left|\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)} + \boldsymbol{\ell}\cdot\boldsymbol{\Omega}^{(r-1)}\right| \ge \frac{\gamma}{(rK)^\tau} \quad \text{and} \quad \min_{|\boldsymbol{\ell}|=1} \left|\boldsymbol{\ell}\cdot\boldsymbol{\Omega}^{(r-1)}\right| \ge \gamma\,, \tag{42}$$

for some fixed values of both $\gamma > 0$ and $\tau > n_1 - 1$. By applying the Lie series $\exp\left(\mathcal{L}_{\chi_1^{(r)}}\right)$ to the old Hamiltonian $H^{(\mathrm{I};\,r)}$, we have a new one, which we denote as $H^{(\mathrm{II};\,r)} = \exp\left(\mathcal{L}_{\chi_1^{(r)}}\right) H^{(\mathrm{I};\,r)}$ and have the same structure as that described in (37), i.e.,

$$\begin{aligned}\mathcal{H}^{(\mathrm{II};\,r)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi}) =& \mathcal{E}^{(r)} + \boldsymbol{\omega}^{(r-1)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(r-1)}\cdot\boldsymbol{J} + \sum_{s\ge 0}\sum_{\ell\ge 3} f_\ell^{(\mathrm{II};\,r,\,s)} \\ &+ \sum_{s\ge r}\sum_{\ell=0}^{2} f_\ell^{(\mathrm{II};\,r,\,s)}\,,\end{aligned} \tag{43}$$

The functions $f_\ell^{(\mathrm{II};\,r,\,s)}$ that compose the new Hamiltonian can be determined with calculations similar to those listed during the description of the first stage of normalization. This means that we initially define $f_\ell^{(\mathrm{II};\,r,\,s)} = f_\ell^{(\mathrm{I};\,r,\,s)}$ $\forall\, \ell \ge 0$, $s \ge 0$. Then, (by abuse of notation) we redefine them many times according to the following rules:

$$\begin{aligned} f_{\ell-j}^{(\mathrm{II};\,r,\,s+jr)} &\hookleftarrow \frac{1}{j!}\mathcal{L}^j_{\chi_1^{(r)}} f_\ell^{(\mathrm{I};\,r,\,s)} \quad \forall\, \ell \ge 0\,,\ 1 \le j \le \ell\,,\ s \ge 0\,, \\ f_0^{(r,\,2r)} &\hookleftarrow \frac{1}{2}\mathcal{L}^2_{\chi_1^{(r)}}\left(\boldsymbol{\omega}^{(r)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(r)}\cdot\boldsymbol{J}\right). \end{aligned} \tag{44}$$

Because of the homological equation (39), we add also a further redefinition so that $f_1^{(\mathrm{II};\,r,\,r)} = 0$. By applying Lemma 2 to formula (44), it is easy to check that $f_\ell^{(\mathrm{II};\,r,\,s)} \in \mathfrak{P}_{\ell,sK}$ $\forall\, \ell \ge 0$, $s \ge 0$.

Third Stage of the *r*-th Normalization Step

The third and last stage of normalization is more elaborated. It aims to remove terms belonging to two different classes: first, those linear in $\boldsymbol{p}$ and independent of $(\boldsymbol{J}, \boldsymbol{\varphi})$, moreover, other terms that are quadratic in $\sqrt{\boldsymbol{J}}$ and independent of $\boldsymbol{p}$. Such a part of the perturbation is removed by the composition of two canonical transformations expressed by Lie series, being the corresponding generating functions $X_2^{(r)}(\boldsymbol{p}, \boldsymbol{q}) \in \widehat{\mathfrak{P}}_{1,0,rK}$ and $Y_2^{(r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi}) \in \widehat{\mathfrak{P}}_{0,2,rK}$, respectively. Moreover, the third stage is ended by a linear canonical transformation $\mathfrak{D}^{(r)}$ that leaves the pair $(\boldsymbol{p}, \boldsymbol{q})$ unchanged and it aims to diagonalize the terms that are quadratic in $\sqrt{\boldsymbol{J}}$ and independent of the angles $\boldsymbol{q}$. Let us detail all these changes of coordinates, so that the algorithm will be unambiguously defined at the end of our discussion.

The generating functions $X_2^{(r)}$ is in charge to remove terms that are linear in $\boldsymbol{p}$ and do depend on the angles $\boldsymbol{q}$ up to the trigonometric degree rK. Therefore, it is a solution of the following homological equation:

$$\left\{\boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p},\ X_2^{(r)}\right\} + f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{p}, \boldsymbol{q}) - \langle f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{p}, \boldsymbol{q})\rangle_{\boldsymbol{q}} = 0\,. \tag{45}$$

Let us recall that $f_2^{(\mathrm{II};\, r,\, r)} \in \mathfrak{P}_{2,rK} = \widehat{\mathfrak{P}}_{1,0,rK} \cup \widehat{\mathfrak{P}}_{0,2,rK}$; indeed, such a function does depend on all the canonical variables, i.e., $f_2^{(\mathrm{II};\, r,\, r)} = f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})$. Therefore, we denote with $f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{p}, \boldsymbol{q})$ the subpart of $f_2^{(\mathrm{II};\, r,\, r)}$ that is depending just on $(\boldsymbol{p}, \boldsymbol{q})$. Analogously, in the following $f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})$ will denote the subpart of $f_2^{(\mathrm{II};\, r,\, r)}$ that does depend on all the canonical variables but the actions $\boldsymbol{p}$ and so on also for what concerns $f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{J}, \boldsymbol{\varphi})$. For the sake of clarity, this highly non-standard notation will be maintained up to the end of the present subsection. Let us here emphasize that the term $\langle f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{p}, \boldsymbol{q})\rangle_{\boldsymbol{q}}$ will be added to the part in normal form, by updating the angular velocity vector $\boldsymbol{\omega}$, in agreement with what has been done in the context of the construction of the Kolmogorov normal form. We can deal with the homological equation (45) in the same way as for (18). Indeed, the solution writes as

$$X_2^{(r)}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{|\boldsymbol{j}|=1} \sum_{0<|\boldsymbol{k}|\le rK} \frac{c_{\boldsymbol{j},\boldsymbol{k}}\, \boldsymbol{p}^{\boldsymbol{j}} \exp\left(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{q}\right)}{\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r)}}\,, \tag{46}$$

where the expansion of the perturbing term $f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{p}, \boldsymbol{q}) \in \widehat{\mathfrak{P}}_{1,0,rK}$ is such that $f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{|\boldsymbol{j}|=1} \sum_{0<|\boldsymbol{k}|\le rK} c_{\boldsymbol{j},\boldsymbol{k}}\, \boldsymbol{p}^{\boldsymbol{j}} \exp\left(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{q}\right)$. Once again, the solution written in (46) is valid provided that the non-resonance condition (11) is satisfied.

The generating function $Y_2^{(r)}$ aims to remove the part of the term of $f_2^{(\mathrm{II};\, r,\, r)}$ that is quadratic in $\sqrt{\boldsymbol{J}}$ and does depend on the angles $\boldsymbol{q}$. Therefore, $Y_2^{(r)}$ has to solve the following homological equation:

$$\left\{\boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p} + \boldsymbol{\Omega}^{(r-1)} \cdot \boldsymbol{J},\ Y_2^{(r)}\right\} + f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi}) - \langle f_2^{(\mathrm{II};\, r,\, r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})\rangle_{\boldsymbol{q}} = 0\,. \tag{47}$$

In order to describe the solution of such an equation, it is convenient to write the explicit expansion of the perturbing term $f_2^{(\mathrm{II};\,r,\,r)}(\boldsymbol{q},\,\boldsymbol{J},\,\boldsymbol{\varphi})$. For instance, this can be done in the following way:

$$f_2^{(\mathrm{II};\,r,\,r)}(\boldsymbol{q},\,\boldsymbol{J},\,\boldsymbol{\varphi}) = \sum_{0\le|\boldsymbol{k}|\le rK} \sum_{i,\,j=1}^{n_2} c_{\boldsymbol{k},\,i,\,j}^{(\pm,\pm)} \sqrt{J_i J_j}\, \exp\left[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{q} \pm \varphi_i \pm \varphi_j)\right], \tag{48}$$

where $c_{\boldsymbol{k},\,i,\,j}^{(+,+)}$ and $c_{\boldsymbol{k},\,i,\,j}^{(+,-)}$ are the coefficients referring to the Fourier harmonics $\boldsymbol{k}\cdot\boldsymbol{q} + \varphi_i + \varphi_j$ and $\boldsymbol{k}\cdot\boldsymbol{q} + \varphi_i - \varphi_j$, respectively, and so on. Thus, the generating function $Y_2^{(r)}$ is determined by Eq. (47) in such a way that

$$Y_2^{(r)}(\boldsymbol{q},\,\boldsymbol{J},\,\boldsymbol{\varphi}) = \sum_{0<|\boldsymbol{k}|\le rK} \sum_{i,\,j=1}^{n_2} \frac{c_{\boldsymbol{k},\,i,\,j}^{(\pm,\pm)} \sqrt{J_i J_j}\, \exp\left[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{q} \pm \varphi_i \pm \varphi_j)\right]}{\mathrm{i}\left(\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)} \pm \Omega_i^{(r-1)} \pm \Omega_j^{(r-1)}\right)}, \tag{49}$$

which is well defined provided that the angular velocity vector $\boldsymbol{\omega}^{(r-1)}$ satisfies both the already mentioned Diophantine inequality (11) and the so-called second Melnikov non-resonance condition up to order rK (see [29]), i.e.,

$$\min_{\substack{0<|\boldsymbol{k}|\le rK,\\ |\boldsymbol{\ell}|=2}} \left|\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)} + \boldsymbol{\ell}\cdot\boldsymbol{\Omega}^{(r-1)}\right| \ge \frac{\gamma}{(rK)^\tau} \tag{50}$$

with fixed values of both parameters $\gamma > 0$ and $\tau > n_1 - 1$.

After having performed these two changes of coordinates, we still may have terms that do not depend on $\boldsymbol{q}$ and are either linear in $\boldsymbol{p}$ or quadratic in $\sqrt{\boldsymbol{J}}$. The former ones can be directly added to the part in normal form, whereas the latter have to be preliminarily put in diagonal form. This can be done by means of a canonical transformation $\mathfrak{D}^{(r)}$ such that

$$\left(\boldsymbol{\Omega}^{(r-1)}\cdot\boldsymbol{J} + f_2^{(\mathrm{II};\,r,\,r)}(\boldsymbol{J},\boldsymbol{\varphi})\right)\Big|_{(\boldsymbol{J},\boldsymbol{\varphi})=\mathfrak{D}^{(r)}(\bar{\boldsymbol{J}},\bar{\boldsymbol{\varphi}})} = \boldsymbol{\Omega}^{(r)}\cdot\bar{\boldsymbol{J}}\,. \tag{51}$$

Such an equation in the unknown transformation $\mathfrak{D}^{(r)}$ can be solved provided that

$$\min_{|\boldsymbol{\ell}|=2} \left|\boldsymbol{\ell}\cdot\boldsymbol{\Omega}^{(r-1)}\right| \ge \gamma \tag{52}$$

and $f_2^{(\mathrm{II};\,r,\,r)}$ is small enough, as it is explained, e.g., in Sect. 7 of [16] (where this problem is considered in the equivalent case dealing with polynomial canonical coordinates). In practical implementations, such a change of coordinates $\mathfrak{D}^{(r)}$ can be conveniently defined by composing a subsequence of Lie series, each of them being related to a quadratic generating function $\mathcal{D}_2^{(r;\,m)}(\boldsymbol{J},\boldsymbol{\varphi}) \in \widehat{\mathfrak{P}}_{0,2,0}$ with $m \in \mathbb{N}\setminus\{0\}$. All these new generating functions can be determined by adopting the following

computational (sub)procedure of iterative type. First, we introduce the new angular velocity vector $\boldsymbol{\Omega}^{(r;\,0)}$ so that

$$\boldsymbol{\Omega}^{(r;\,0)} \cdot \boldsymbol{J} = \boldsymbol{\Omega}^{(r-1)} \cdot \boldsymbol{J} + \langle f_2^{(\mathrm{II};\,r,\,r)}(\boldsymbol{J},\boldsymbol{\varphi})\rangle_{\boldsymbol{\varphi}} \tag{53}$$

and the new function

$$\mathfrak{g}_2^{(r;\,0)}(\boldsymbol{J},\boldsymbol{\varphi}) = f_2^{(\mathrm{II};\,r,\,r)}(\boldsymbol{J},\boldsymbol{\varphi}) - \langle f_2^{(\mathrm{II};\,r,\,r)}(\boldsymbol{J},\boldsymbol{\varphi})\rangle_{\boldsymbol{\varphi}} \,. \tag{54}$$

The general m-th step of this iterative (sub)procedure starts by solving the following homological equation:

$$\left\{\boldsymbol{\Omega}^{(r;\,m-1)} \cdot \boldsymbol{J},\, \mathcal{D}_2^{(r;\,m)}(\boldsymbol{J},\boldsymbol{\varphi})\right\} + \mathfrak{g}_2^{(r;\,m-1)}(\boldsymbol{J},\boldsymbol{\varphi}) = 0 \,, \tag{55}$$

where $\mathfrak{g}_2^{(r;\,m-1)} \in \widehat{\mathfrak{P}}_{0,2,0}$ is such that $\langle \mathfrak{g}_2^{(r;\,m-1)}\rangle_{\boldsymbol{\varphi}} = 0$ (and, therefore, also the new generating function $\mathcal{D}_2^{(r;\,m)}$ is sharing these same properties with $\mathfrak{g}_2^{(r;\,m-1)}$). Let us now initially introduce $\mathfrak{g}_2^{(r;\,m)} = 0$ and (by the usual abuse of notation) we redefine it many times according to the following rule:

$$\mathfrak{g}_2^{(r;\,m)} \hookleftarrow \frac{j}{(j+1)!}\mathcal{L}^{j}_{\mathcal{D}_2^{(r;\,m)}}\mathfrak{g}_2^{(r;\,m-1)} \quad \forall\, j \geq 1 \,. \tag{56}$$

Actually, at this point one can easily check that

$$\exp\left(\mathcal{L}_{\mathcal{D}_2^{(r;\,m)}}\right)\left(\boldsymbol{\Omega}^{(r;\,m-1)} \cdot \boldsymbol{J} + \mathfrak{g}_2^{(r;\,m-1)}\right) = \boldsymbol{\Omega}^{(r;\,m-1)} \cdot \boldsymbol{J} + \mathfrak{g}_2^{(r;\,m)} \,,$$

by using homological equation (55). Furthermore, we set

$$\boldsymbol{\Omega}^{(r;\,m)} \cdot \boldsymbol{J} = \boldsymbol{\Omega}^{(r;\,m-1)} \cdot \boldsymbol{J} + \langle \mathfrak{g}_2^{(r;\,m)}(\boldsymbol{J},\boldsymbol{\varphi})\rangle_{\boldsymbol{\varphi}} \tag{57}$$

and we redefine one last time $\mathfrak{g}_2^{(r;\,m)}$ so that

$$\mathfrak{g}_2^{(r;\,m)}(\boldsymbol{J},\boldsymbol{\varphi}) = \mathfrak{g}_2^{(r;\,m)}(\boldsymbol{J},\boldsymbol{\varphi}) - \langle \mathfrak{g}_2^{(r;\,m)}(\boldsymbol{J},\boldsymbol{\varphi})\rangle_{\boldsymbol{\varphi}} \,. \tag{58}$$

By applying repeatedly Lemma 2 to formulæ (53)–(58), it is easy to check that both functions $\mathcal{D}_2^{(r;\,m)}$ and $\mathfrak{g}_2^{(r;\,m)}$ belong to the class $\widehat{\mathfrak{P}}_{0,2,0}$ (also because they depend on neither $\boldsymbol{p}$ nor $\boldsymbol{q}$) and their angular average is equal to zero. In principle, these remarks would allow to iterate infinitely many times this computational (sub)procedure, that we are using to solve Eq. (51). However, in practical implementations, we have to set a criterion to stop the iterations so to ensure that the algorithm can be worked out in a finite number of operations. This can be done, for instance, in such a way to end the computations when the angular velocity vector does not modify anymore. This means that the final value $\bar{m}$ of the normalization step for this iterative (sub)procedure is such

that the equation $\boldsymbol{\Omega}^{(r;\,\bar{m})} = \boldsymbol{\Omega}^{(r;\,\bar{m}-1)}$ holds true *in the framework of the numbers that are representable on a computer*[18] (for instance, the `double precision` type). By setting $\boldsymbol{\Omega}^{(r)} = \boldsymbol{\Omega}^{(r;\,\bar{m})}$ and the canonical transformation $\mathfrak{D}^{(r)}$ equal to composition of all the Lie series generated by the *finite* sequence of functions $\left\{\mathcal{D}_2^{(r;\,m)}\right\}_{m=1}^{\bar{m}}$, we determine a solution[19] of (51) that is valid up to the numerical round-off errors.

Finally, we need to understand how all these generating functions (that have been defined during the third stage of the r-th normalization step) give their contributions to the Hamiltonian terms appearing in the following expansion:

$$\begin{aligned}\mathcal{H}^{(r)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi}) =& \mathcal{E}^{(r)} + \boldsymbol{\omega}^{(r)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(r)}\cdot\boldsymbol{J} + \sum_{s\geq 0}\sum_{\ell\geq 3} f_\ell^{(r,s)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi}) \\ &+ \sum_{s\geq r+1}\sum_{\ell=0}^{2} f_\ell^{(r,\,s)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi})\,,\end{aligned} \tag{59}$$

where $\mathcal{H}^{(r)}$ is defined in (34). In order to describe the definitions of those new summands, it is convenient to introduce the intermediate functions $g_\ell^{(r,\,s)}$, $g'^{(r,\,s)}_\ell$ in the following way. First, we define $g_\ell^{(r,\,s)} = f_\ell^{(\mathrm{II};\,r,\,s)}$ for all non-negative values of the indexes ℓ and s; then, we consider the effects induced by the application of the Lie series with generating function $X_2^{(r)}$ to the Hamiltonian. In order to do that, (by abuse of notation) we redefine many times the new intermediate functions $g_\ell^{(r,\,s)}$ according to the following rules:

$$\begin{aligned} g_\ell^{(r,\,s+jr)} &\hookleftarrow \frac{1}{j!}\mathcal{L}^j_{X_2^{(r)}} f_\ell^{(\mathrm{II};\,r,\,s)} \quad \forall\ j\geq 1,\ \ell\geq 0,\ s\geq 0\,, \\ g_2^{(r,\,jr)} &\hookleftarrow \frac{1}{j!}\mathcal{L}^j_{X_2^{(r)}}\left(\boldsymbol{\omega}^{(r)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(r)}\cdot\boldsymbol{J}\right) \quad \forall\ j\geq 1\,. \end{aligned} \tag{60}$$

As usual, the prescriptions above have been set so to gather the new terms generated by the Lie series $\exp\left(\mathcal{L}_{X_2^{(r)}}\right)$ according to both their total degree in the square root of the actions and the trigonometric degree in the angles. In analogous way, we first introduce $g'^{(r,\,s)}_\ell = g'^{(r,\,s)}_\ell\ \forall\ \ell\geq 0,\ s\geq 0$; then we apply many times the following redefinitions:

$$\begin{aligned} g'^{(r,\,s+jr)}_\ell &\hookleftarrow \frac{1}{j!}\mathcal{L}^j_{Y_2^{(r)}} g_\ell^{(r,\,s)} \quad \forall\ j\geq 1,\ \ell\geq 0,\ s\geq 0\,, \\ g'^{(r,\,jr)}_2 &\hookleftarrow \frac{1}{j!}\mathcal{L}^j_{Y_2^{(r)}}\left(\boldsymbol{\omega}^{(r)}\cdot\boldsymbol{p} + \boldsymbol{\Omega}^{(r)}\cdot\boldsymbol{J}\right) \quad \forall\ j\geq 1\,. \end{aligned} \tag{61}$$

[18] A similar criterion is adopted to determine a maximum value of the index j at which the redefinitions (56) must be stopped.

[19] As an alternative computational method, when one is dealing with the estimates needed to prove the convergence of the algorithm, in [19] the use of the Lie transforms (that are equivalent to the composition of *infinite* sequences of Lie series) has been found to be very suitable.

By applying Lemma 2 to formulæ (60)–(61), it is easy to check that $g'^{(r,s)}_{\ell} \in \mathfrak{P}_{\ell,sK}$ $\forall\ \ell \geq 0,\ s \geq 0$. Let us now remark that each class of type $\mathfrak{P}_{\ell,sK}$ is preserved[20] by the diagonalization transformation $\mathfrak{D}^{(r)}$, for all non-negative values of the indexes ℓ and s. Therefore, it is natural to put

$$f^{(r,s)}_{\ell} = g'^{(r,s)}_{\ell} \circ \mathfrak{D}^{(r)} \ . \tag{62}$$

for all indexes $\ell \geq 0$ and $s \geq 0$.

At the end of the r-th normalization step, it is convenient that the terms linearly depending just on $\boldsymbol{p}$ or $\boldsymbol{J}$ are included in the main part of the Hamiltonian, because all of them belong to the same class of functions, i.e. $\mathfrak{P}_{2,0}$. For this purpose, we introduce the new angular velocity vector $\boldsymbol{\omega}^{(r)}$, in such a way that

$$\boldsymbol{\omega}^{(r)} \cdot \boldsymbol{p} = \boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p} + f^{(\mathrm{II};\,r,\,0)}_{2}(\boldsymbol{p}) \ , \tag{63}$$

while the new values of the components of $\boldsymbol{\Omega}^{(r)}$ are defined by Eq. (51), that also allows us to put $f^{(r,r)}_{2} = 0$. This ends the justification of the fact that the Hamiltonian $\mathcal{H}^{(r)}$ can be written as in formula (59) with new terms such that $f^{(r,\,s)}_{\ell} \in \mathfrak{P}_{\ell,sK}$ and $\mathcal{E}^{(r)} \in \mathfrak{P}_{0,0}$. Therefore, $\mathcal{H}^{(r)}$ has the same structure of $\mathcal{H}^{(r-1)}$ in (33); this also mean that the normalization algorithm can be iterated to the next $(r+1\text{-th})$ step. As a final comment ending the present subsection, let us also remark that the new perturbative terms $f^{(r,s)}_{\ell}$ with $\ell = 0,\ 1,\ 2$ are expected to be smaller with respect to the previous ones; this is because of the Fourier decay of the coefficients jointly with the fact that we removed the part of perturbation up to the trigonometric degree rK.

3.2 *On the Convergence of the Algorithm Constructing the Normal Form for Elliptic Tori*

As we have discussed since the introduction, in the present work we make the choice of adopting the same approach to construct two different normal forms, that are related to KAM invariant manifolds and elliptic tori, respectively. For what concerns the analysis of the convergence, such a choice now allows us to use arguments that are very similar to those described in the previous Sect. 2. In particular, also for what concerns the motion on elliptic tori, we emphasize that it can be approximated within a precision up to a fixed order of magnitude by using our procedure that is *explicitly computable*, because the total amount of operations that are defined also by this normalization algorithm is *finite*.

[20] This statement can be justified, by referring also to the definition of the canonical transformation $\mathfrak{D}^{(r)}$ as composition of all the Lie series generated by the set of functions $\left\{\mathcal{D}^{(r;\,m)}_{2}\right\}^{\tilde{m}}_{m=1}$. In fact, it can be easily done by applying Lemma 2 to all the contributions due to the repeated application of the Lie derivative with generating functions $\mathcal{D}^{(r;\,m)}_{2} \in \mathfrak{P}_{2,0}$.

The non-resonance conditions we have assumed in (11), (42), (50) and (52) can be summarized in the following way:

$$\min_{\substack{0<|\boldsymbol{k}|\le rK,\\ 0\le|\boldsymbol{\ell}|\le 2}} \left|\boldsymbol{k}\cdot\boldsymbol{\omega}^{(r-1)}+\boldsymbol{\ell}\cdot\boldsymbol{\Omega}^{(r-1)}\right| \ge \frac{\gamma}{(rK)^{\tau}} \quad \text{and} \quad \min_{0<|\boldsymbol{\ell}|\le 2} \left|\boldsymbol{\ell}\cdot\boldsymbol{\Omega}^{(r-1)}\right| \ge \gamma\,, \tag{64}$$

with $\gamma > 0$ and $\tau > n_1 - 1$. Let us here resume the parametric dependence of all the Hamiltonian terms on the initial value of the angular velocity vector $\boldsymbol{\omega}^{(0)}$, as it has been introduced at the beginning of the previous Sect. 3.1 (see the discussion following the statement of Lemma 2). In particular, in the Diophantine inequalities reported in (64) the angular velocity vectors at the r-th normalization step are functions of $\boldsymbol{\omega}^{(0)}$, i.e., $\boldsymbol{\omega}^{(r-1)} = \boldsymbol{\omega}^{(r-1)}(\boldsymbol{\omega}^{(0)})$ and $\boldsymbol{\Omega}^{(r-1)} = \boldsymbol{\Omega}^{(r-1)}(\boldsymbol{\omega}^{(0)})$. Let us recall that we do not try to keep a full control on the way for what concerns the angular velocity vectors that are modified passing from the $r-1$-th normalization step to the next one. Therefore, let us recall also here that such an approach is in contrast with the original proof scheme that was designed to construct the Kolmogorov normal form for *maximal* invariant tori, where the angular velocities are kept fixed (see [22] or, e.g., [17]), but it is somehow unavoidable because of the occurrence of the transversal angular velocities $\boldsymbol{\Omega}^{(r-1)}(\boldsymbol{\omega}^{(0)})$ that in general cannot remain constant along the normalization procedure. This seems to prevent the complete construction of the normal form and so also for what concerns the proof of the existence of an elliptic torus. Nevertheless, following the approach designed by Pöschel in [36], it can be proved that the Lebesgue measure of the resonant regions where the Melnikov conditions are not satisfied shrinks to zero with the size of the perturbation. Therefore, the chances of success in constructing the normal form for elliptic tori are described by the following statement.

Theorem 2 *Consider the family of real Hamiltonians $\mathcal{H}^{(0)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi};\boldsymbol{\omega}^{(0)})$ of the type described in (31). Those functions are defined so that $\mathcal{H}^{(0)} : \mathcal{O}_1\times\mathbb{T}^{n_1}\times\mathcal{O}_2\times\mathbb{T}^{n_2}\times\mathcal{U}\mapsto\mathbb{R}$, with $\mathcal{O}_1$ and $\mathcal{O}_2$ open neighborhoods of the origin in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}_+\cup\{\boldsymbol{0}\}$, respectively, while $\boldsymbol{\omega}^{(0)}\in\mathcal{U}$, being $\mathcal{U}$ an open subset of $\mathbb{R}^{n_1}$. Moreover, let a special class of functions include each of the terms that are of type $f_\ell^{(0,s)}$ and appear in the expansion (31), in such a way that $f_\ell^{(0,s)}\in\mathfrak{P}_{\ell,sK}$ for a fixed positive integer K. We also assume that*

(a) all the functions $\mathcal{E}^{(0)}:\mathcal{U}\mapsto\mathbb{R}$, $\boldsymbol{\Omega}^{(0)}:\mathcal{U}\mapsto\mathbb{R}^{n_1}$ and $f_\ell^{(0,s)}:\mathcal{O}_1\times\mathbb{T}^{n_1}\times\mathcal{O}_2\times\mathbb{T}^{n_2}\times\mathcal{U}\mapsto\mathbb{R}$, appearing in (31), are analytic functions with respect to $\boldsymbol{\omega}^{(0)}\in\mathcal{U}$;

(b) $\Omega_i^{(0)}(\boldsymbol{\omega}^{(0)})\neq\Omega_j^{(0)}(\boldsymbol{\omega}^{(0)})$ and $\Omega_{i_2}^{(0)}(\boldsymbol{\omega}^{(0)})\neq 0$ for $\boldsymbol{\omega}^{(0)}\in\mathcal{U}$ and $1\le i<j\le n_2$, $1\le i_2\le n_2$;

(c) for some fixed and positive values of ε and E, one has

$$\sup_{(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi};\boldsymbol{\omega}^{(0)})\in\mathcal{O}_1\times\mathbb{T}^{n_1}\times\mathcal{O}_2\times\mathbb{T}^{n_2}\times\mathcal{U}} \left|f_\ell^{(0,s)}(\boldsymbol{p},\boldsymbol{q},\boldsymbol{J},\boldsymbol{\varphi};\boldsymbol{\omega}^{(0)})\right| \le \varepsilon^s E \tag{65}$$

$\forall\, s\ge 1,\ \ell\ge 0$ and $\forall\,\ell\ge 3$ when $s=0$.

Then, there is a positive $\varepsilon^\star$ *such that for* $0 \le \varepsilon < \varepsilon^\star$ *the following statement holds true: there exists a non-resonant set* $\mathcal{U}^{(\infty)} \subset \mathcal{U}$ *of positive Lebesgue measure and with the measure of* $\mathcal{U} \setminus \mathcal{U}^{(\infty)}$ *tending to zero for* $\varepsilon \to 0$ *for bounded* $\mathcal{U}$*, such that for each* $\boldsymbol{\omega}^{(0)} \in \mathcal{U}^{(\infty)}$ *there exists an analytic canonical transformation* $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi}) = \psi^{(\infty)}_{\varepsilon;\boldsymbol{\omega}^{(0)}}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{\Xi}, \boldsymbol{\Theta})$ *leading the Hamiltonian to the normal form written in* (32)*, where* $\mathcal{E}^{(\infty)}(\boldsymbol{\omega}^{(0)})$ *is a finite real value fixing the constant energy level that corresponds to the invariant elliptic torus* $\left\{(\boldsymbol{P} = \boldsymbol{0},\ \boldsymbol{Q} \in \mathbb{T}^{n_1}, \boldsymbol{\Xi} = \boldsymbol{0}, \boldsymbol{\Theta} = \boldsymbol{0})\right\}$*. Moreover, the canonical change of coordinates is close to the identity in the sense that* $\left\| \psi^{(\infty)}_{\varepsilon;\boldsymbol{\omega}^{(0)}}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{\Xi}, \boldsymbol{\Theta}) - (\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{\Xi}, \boldsymbol{\Theta}) \right\| = \mathcal{O}(\varepsilon)$ *and the same applies also to both the energy level and the detunings of the angular velocity vectors (that are* $\left| \mathcal{E}^{(\infty)}(\boldsymbol{\omega}^{(0)}) - \mathcal{E}^{(0)}(\boldsymbol{\omega}^{(0)}) \right| = \mathcal{O}(\varepsilon)$*,* $\left\| \boldsymbol{\omega}^{(\infty)}(\boldsymbol{\omega}^{(0)}) - \boldsymbol{\omega}^{(0)} \right\| = \mathcal{O}(\varepsilon)$ *and* $\left\| \boldsymbol{\Omega}^{(\infty)}(\boldsymbol{\omega}^{(0)}) - \boldsymbol{\Omega}^{(0)}(\boldsymbol{\omega}^{(0)}) \right\| = \mathcal{O}(\varepsilon)$*, respectively).*

The complete proof of theorem above is reported in [7], where it is ensured the convergence of a normalization algorithm that is substantially the same with respect to the one described in the previous Sect. 3.1 apart some very minor modifications.[21] Therefore, the approach of that paper is based on a convergence scheme of linear type. Nevertheless, the more geometrical part of that work (which deals with the estimates of the volume covered by the resonant region) is borrowed from [36], where a statement nearly equivalent to Theorem 2 is proved by adopting a fast convergence scheme of quadratic type.

In the present case studying the elliptic tori, the choice to let the angular velocity vectors change at every normalization step is somehow more natural with respect to the original proof scheme designed by Kolmogorov. This is due to the fact that here the procedure allowing to keep fixed the angular velocities is not complete, because it involves less free parameters than the number of degrees of freedom. This is a major difference with respect to the algorithm constructing the normal form for KAM tori, where those two integer numbers are equal. For what concerns the case of the elliptic tori too, some work[22] is in progress in order to revisit the problem of the convergence of this type of normalization algorithms so as to provide a statement where the final result is not expressed in a probabilistic sense (i.e., by referring to the Lebesgue measure). This can be done by fixing since the beginning the final value of the angular velocity vectors $\left(\boldsymbol{\omega}^{(\infty)}, \boldsymbol{\Omega}^{(\infty)}\right)$ and their non-resonance properties; we emphasize that this allow to explicitly solve all the homological equations that are introduced at every step of the algorithm. Also here, the total detunings $\boldsymbol{\omega}^{(\infty)} - \boldsymbol{\omega}^{(0)}$ and $\boldsymbol{\Omega}^{(\infty)} - \boldsymbol{\Omega}^{(0)}$ are given in terms of series whose coefficients are defined in a recursive way. Such an approach is also inspired by the need to revisit what was successfully done in order to show the existence of elliptic tori in PDEs problems (see [6]).

[21] For instance, in order to describe the transverse dynamics with respect to the elliptic tori, the complex canonical coordinates $(z, \mathrm{i}\bar{z})$ instead of the action-angle ones are used, where $z_j = J_j e^{\mathrm{i}\varphi_j}$ $\forall\ j = 1, \ldots, n_2$.

[22] Danesi, V., Locatelli, U.: work in progress (2022).

4 Construction of Invariant KAM Tori in Exoplanetary Systems with Rather Eccentric Orbits

In order to properly introduce a Cauchy problem which includes the ordinary differential equations (ODE) for a planetary system, the initial conditions at a given time are needed and so also for the positions and the velocities in an astrocentric frame. It is well known that they can be replaced by the orbital elements

$$\left\{ \left(a_j \, , \, e_j \, , \, \iota_j \, , \, M_j \, , \, \omega_j \, , \, \Omega_j \right) \, : \, \forall \, j = 1, \, \dots \, , N \right\} ,$$

being N the number of the planets that are considered in the system. Orbital elements refer to the so called osculating Keplerian ellipse, which describes a fictitious motion having the same instantaneous values of both position and velocity with respect to the planet. For what concerns the Keplerian ellipse of the j-th planet, the symbols $a_j \, , \, e_j \, , \, \iota_j \, , \, M_j \, , \, \omega_j \, , \, \Omega_j$ denote the semi-major axis, the eccentricity, the inclination,[23] the mean anomaly, the argument of the pericenter[24] and the longitude of the ascending node, respectively. Of course, also the values of the masses $m_j \; \forall \; j = 0, \, 1, \, \dots \, , \, N$ (being m_0 the stellar mass) are needed in order to properly introduce the Cauchy problem for a planetary system, because they enter in the definitions of the momenta, the kinetic energy and the potential one. Unfortunately, none of the detection methods that are nowadays available to discover extrasolar planets is able to measure all the orbital elements and the masses that completely define the ODE problem (see, e.g., [2]). For the sake of simplicity, instead of considering a generic planetary problem with $N + 1$ bodies, let us focus on a specific case, i.e., the extrasolar system hosting two planets orbiting around the star named HD 4732[25] (the value of its mass is reported in the caption of the following table). The values of the known orbital elements of those exoplanets as they are given by the radial velocity detection method are reported in Table 1. Let us recall that such a detection technique is unable to provide a complete information about the mass of every j-th planet; instead, it gives its minimum value $m_j \sin(\iota_j)$.

Let us now explain how we have decided to complete the initial conditions, by also giving the motivations of our choice. Since we are interested in studying the planetary dynamics of the HD 4732 system in the framework of a secular model, we expect that its dependence on the initial values of the mean anomalies is weak. We emphasize that such an assumption does not hold true in general (see, e.g., [26]), but it is rather natural in the case of the HD 4732 planetary system because the revolution

[23] ι_j is the inclination of the Keplerian ellipse with respect to the plane orthogonal to the line of sight (i.e., the direction pointing to the object one is observing), that is usually said to be "tangent to the celestial sphere".

[24] Unfortunately, the same symbol (namely, ω) is used to denote both the angular velocity in KAM theory and the pericenter argument in astronomy. Hereafter, when the symbol ω appears *without* superscripts, it will refer just to the latter quantity.

[25] Since the detection of a fainter stellar companion in 2019 (see [32]) HD 4732 has been renamed as HD 4732A. For brevity, in the present paper we refer to such a star with the old name.

Table 1 Known orbital elements and minimal masses of the detected exoplanets orbiting around the HD 4732 star, whose mass is 1.74 times bigger than the solar one. The following data are taken from the central values of the ranges given in Table 5 of [39]. The corresponding units of measure are reported in every column between pairs of square brackets; in particular, we recall that the eccentricity of an ellipse is a pure number ranging in (0, 1) and M_{Jup} means "Jupiter mass". Since the initial time is irrelevant for an autonomous system, we have set it equal to zero in the parentheses following the orbital elements

Planet name	Planet index j	$a_j(0)$ [AU]	$e_j(0)$	$\omega_j(0)$ [°]	$m_j \sin\left(\iota_j(0)\right)$ [M_{Jup}]
HD 4732b	1	1.19	0.13	85	2.37
HD 4732c	2	4.60	0.23	118	2.37

periods are far from mean-motion resonances and they are much shorter with respect to those corresponding to the remaining angles that appear in the orbital elements list. Therefore, we simply set[26]

$$M_1(0) = M_2(0) = 0^\circ \, . \tag{66}$$

For what concerns the extrasolar system HD 4732, we plan to start a study of the dependence of its orbital dynamics on the mutual inclination i_{mut}. The present section deals with the beginning of such a research project, that will be extended in a forthcoming work. For this purpose, it is convenient to consider orbital planes initially located in such a way they are symmetric with respect to the line of sight that is also orthogonal to their intersection. As an example of this particular configuration, we can consider the case with $\iota_1(0) = 89^\circ$, $\iota_2(0) = 91^\circ$ and

$$\Omega_1(0) = \Omega_2(0) = 0^\circ \, . \tag{67}$$

In view of the general relation

$$\cos i_{\mathrm{mut}} = \cos\iota_1 \cos\iota_2 + \sin\iota_1 \sin\iota_2 \cos(\Omega_1 - \Omega_2) \, ,$$

we readily obtain that $i_{\mathrm{mut}} = 2^\circ$. More in general, we introduce the following set of initial conditions

[26] Since the times of passage at the pericenter are given by the radial velocity detection methods and they are different, we stress that our choice of defining the initial values of the mean anomalies so that $M_1(0) = M_2(0) = 0^\circ$ is not coherent with the observations about the two planets orbiting around HD 4732. However, we consider that this small inconsistency of our settings should be harmless, just because of the expectation that its secular dynamics should be very weakly affected by the initial values of the mean anomalies.

$$\mathcal{I}_{i_{\mathrm{mut}}(0)} = \Big\{ \big(a_1(0)\,,\, a_2(0)\,,\, e_1(0)\,,\, e_2(0)\,, \\ \iota_1(0) = 90^\circ - \frac{i_{\mathrm{mut}}(0)}{2}\,,\ \iota_2(0) = 90^\circ + \frac{i_{\mathrm{mut}}(0)}{2}\,, \\ M_1(0)\,,\ M_2(0)\,, \omega_1(0)\,,\ \omega_2(0)\,, \Omega_1(0)\,, \Omega_2(0)\big)\Big\}\,, \tag{68}$$

where the inclinations are parameterized with respect to $i_{\mathrm{mut}}(0)$, while the values of all the remaining orbital elements are defined according to Table 1, jointly with formulæ (66) and (67). Of course, the values of the planetary masses m_1 and m_2 can be recovered multiplying the minimal masses (that appear in the last column of Table 1) by the increasing factor $1\big/ \sin\big(\iota_j(0)\big)$. This remark helps us to understand that all the parameters and the initial conditions have been properly defined and they can eventually depend just on the value of $i_{\mathrm{mut}}(0)$. This way to parameterize the model has been introduced to better understand the properties of our (new) algorithm constructing invariant tori as a function of the mutual inclinations. A previous approach to the same problem was described in [41] and it was shown to be successful just for systems with rather small eccentricities of the exoplanets, being their initial values less than 0.1 . This is not the case of the exoplanets in the system HD 4732, because both their initial values of the eccentricities (reported in Table 1) are larger than 0.1 . We emphasize that this choice has been made with the purpose to show that our following new formulation of the constructing algorithms applies to a more extended range of models with respect to the previous approach.

Let us also recall that, in a three-body planetary problem, the longitudes of the nodes are always opposite, if they are measured with respect to the so called Laplace plane, that is invariant because it is orthogonal to the total angular momentum, by definition (see, e.g., Sect. 6.2 of [24]). Moreover, the Hamiltonian does not depend on the sum of $\Omega_1 + \Omega_2$, because of the invariance with respect to the rotations. In the following subsection, we will explain why it is preferable to consider expansions of the Hamiltonian in a frame where the Laplace plane is the horizontal one. In Celestial Mechanics the word "inclination" often refers to the angle (say, $\mathrm{i}_j \in [0^\circ, 180^\circ]$) between the angular momentum of the j-th planet and the total one. With this notation, the following relation holds true: $i_{\mathrm{mut}} = \mathrm{i}_1 + \mathrm{i}_2$.

4.1 Secular Model at Order Two in the Masses

In the present subsection, we are going to introduce a model describing the secular dynamics of a planetary system, in a way that provides results more reliable with respect to a simple average over the revolution angles (see, e.g., [38]). We emphasize that we derive the secular model at order two in the masses, by applying an approach inspired to the construction of the Kolmogorov normal form. This is a major difference with respect to other approaches providing the same level of accuracy for a

secular model (see, e.g., [23] and references therein). Here, in order to introduce our secular model, we will adopt the approach described in [41], that is summarized as follows.

A three-body Hamiltonian problem has nine degrees of freedom, but three of them can be easily separated so as to describe the uniform motion of the center of mass in an inertial frame. The untrivial part of the dynamics is represented in astrocentric canonical coordinates and its degrees of freedom can be further reduced by two using the conservation of the total angular momentum $\boldsymbol{C}$. As it is shown in Sect. 6 of [24], this allows us to write the Hamiltonian in Poincaré canonical variables, that are

$$\begin{aligned} \Lambda_j &= \frac{m_0 m_j}{m_0 + m_j}\sqrt{G(m_0+m_j)a_j}\,, \quad & \xi_j &= \sqrt{2\Lambda_j}\sqrt{1-\sqrt{1-e_j^2}}\cos(\omega_j)\,, \\ \lambda_j &= M_j + \omega_j\,, & \eta_j &= -\sqrt{2\Lambda_j}\sqrt{1-\sqrt{1-e_j^2}}\sin(\omega_j)\,. \end{aligned} \tag{69}$$

The reduction of the total angular momentum makes implicit the dependence on the inclinations i_j and on the longitudes of the nodes Ω_j. In the Laplace reference frame the mutual inclination is the sum of the two inclinations and so is given by a rather simple relation involving the Poincaré variables, i.e.,

$$i_{\mathrm{mut}} = \mathrm{i}_1 + \mathrm{i}_2 = \arccos\left(\frac{C^2 - \Lambda_1^2(1-e_1^2) - \Lambda_2^2(1-e_2^2)}{2\Lambda_1\Lambda_2\sqrt{1-e_1^2}\sqrt{1-e_2^2}}\right)\,, \tag{70}$$

being $C = \sum_{k=1}^{2}\Lambda_k\sqrt{1-e_k^2}\cos\mathrm{i}_k$, that is the (constant) module of the total angular momentum. Moreover, we introduce a translation $L_j = \Lambda_j - \Lambda_j^*$, where Λ_j^* is defined in order to obtain that in the Keplerian approximation of the motion the values of the semi-major axes are in agreement with the observations. Indeed, the expansions of a Hamiltonian representing a planetary model are usually made around the average values of the semi-major axes or their initial values. For the sake of simplicity, we will adopt this latter option. Such expansions are actually made with respect to these Poincaré variables[27] and the parameter D_2, that measures the difference between the total angular momentum of the system and the one of a similar system with circular and coplanar orbits; i.e., it is defined as $D_2 = \left[(\Lambda_1^* + \Lambda_2^*)^2 - C^2\right]/(\Lambda_1^*\Lambda_2^*)$; therefore, it is of the same order as $e_1^2 + \mathrm{i}_1^2 + e_2^2 + \mathrm{i}_2^2$. Thus, we can write the Hamiltonian of the three–body problem as

$$H_{\mathrm{3BP}} = \sum_{j_1=1}^{\infty} h_{j_1,0}^{(\mathrm{Kep})}(\boldsymbol{L}) + \mu\sum_{s=0}^{\infty}\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} D_2^s\, h_{s;j_1,j_2}^{(\mathcal{P})}(\boldsymbol{L},\boldsymbol{\lambda},\boldsymbol{\xi},\boldsymbol{\eta}) \tag{71}$$

[27] The computation of the coefficients appearing in the expansion (71) is not straightforward. For a detailed discussion of the method we have used for doing such a calculation we refer to [24].

where $\mu = \max\{m_1/m_0, m_2/m_0\}$. Moreover,

- $\mathcal{K}(\boldsymbol{L}) = \sum_{j_1=1}^{\infty} h_{j_1,0}^{(\mathrm{Kep})}(\boldsymbol{L})$ is the Keplerian part and $h_{j_1,0}^{(\mathrm{Kep})}$ is a homogeneous polynomial of degree j_1 in $\boldsymbol{L}$; in particular, $h_{1,0}^{(\mathrm{Kep})} = \boldsymbol{n}^* \cdot \boldsymbol{L}$, where the components of the angular velocity vector $\boldsymbol{n}^*$ are defined by the third Kepler law;
- $h_{s;j_1,j_2}^{(\mathcal{P})}$ is a homogeneous polynomial of degree j_1 in $\boldsymbol{L}$, degree j_2 in $(\boldsymbol{\xi}, \boldsymbol{\eta})$ and with coefficients that are trigonometric polynomials in $\boldsymbol{\lambda}$ and are related to the term D_2^s.

Clearly, in the applications we deal with finite expansions; the truncation parameters will be discussed in the following.

The expression of the Hamiltonian of the three-body problem in (71) highlights the distinction between the so called *fast variables* $(\boldsymbol{L}, \boldsymbol{\lambda})$ and the *secular variables* $(\boldsymbol{\xi}, \boldsymbol{\eta})$. Indeed, if we consider the corresponding Hamilton equations, we have that $\dot{\boldsymbol{\lambda}} = \mathcal{O}(1)$. This means that the motion of the planet along the orbit, that is in first approximation a Keplerian ellipse, has a different timescale with respect to the secular variables, whose variation is due to the interaction between the planets and, therefore, is of $\mathcal{O}(\mu)$. Since we are interested in the study of the long-time stability of the system, a common procedure consists on considering just the evolution of the secular variables, by averaging the Hamiltonian with respect to the fast angles $\boldsymbol{\lambda}$. With a simple average of $H_{3\mathrm{BP}}$ we would obtain a secular approximation with terms of order μ, namely at order 1 in the masses. Here, we consider terms up to order 2 in the masses, averaging with a close to the identity canonical change of coordinates inspired by the algorithm for the construction of the Kolmogorov normal form. Indeed, we focus on the torus corresponding to $\boldsymbol{L} = 0$. The first transformation of coordinates that we define aims at removing the perturbative terms that depend on the angles $\boldsymbol{\lambda}$ but do not depend on the actions $\boldsymbol{L}$, being $\dot{L}_j = \partial H/\partial \lambda_j$ for $j = 1, 2$. This is done by using the term linear in the actions, i.e., $\boldsymbol{n}^* \cdot \boldsymbol{L}$, to define a generating function $\chi_1^{(\mathcal{O}2)}(\boldsymbol{\lambda})$ as the solution of the following homological equation:

$$\left\{\chi_1^{(\mathcal{O}2)},\ \boldsymbol{n}^* \cdot \boldsymbol{L}\right\} + \mu \sum_{\substack{s=0,\ j_2=0 \\ 2s+j_2 \le N_S}} \left\lceil D_2^s\, h_{s;0,j_2}^{(\mathcal{P})} \right\rceil_{\boldsymbol{\lambda}:K_F} = \mu \sum_{\substack{s=0,\ j_2=0 \\ 2s+j_2 \le N_S}} D_2^s \left\langle h_{s;0,j_2}^{(\mathcal{P})} \right\rangle_{\boldsymbol{\lambda}}, \qquad (72)$$

being $\langle\cdot\rangle_{\boldsymbol{\lambda}}$ the average with respect to the angles $\boldsymbol{\lambda}$, while with the notation $\lceil\cdot\rceil_{K_F}$ we mean that the expansions are truncated at the trigonometrical degree K_F in the angles $\boldsymbol{\lambda}$. Let us add a few comments about the truncations parameters K_F and N_S. The value of K_F is defined so as to take into account the main mean-motion quasi-resonances of the system considered. For example, if the system is close to the resonance $k_1^* : k_2^*$, then K_F is defined as $K_F \ge |k_1^*| + |k_2^*|$. In the same spirit, the value N_S of the truncation of the expansions in eccentricity and inclination is set in order to consider the quasi-resonance. Let us assume that the quasi-resonant angular terms are of type $(k_1^*\lambda_1 - k_2^*\lambda_2)$, then in principle it would be convenient to consider expansions up to an order in eccentricity and inclination such that $N_S \ge 2(|k_1^*| - |k_2^*|)$, because of the D'Alembert rules (see [24]). Therefore, in the specific case of the extrasolar system HD 4732, it is rather natural to set $K_F = 9$, because the periods of the two planets

are about 0.986 yr and 7.48 yr, respectively. However, since the ratio of the angular velocities n_1^*/n_2^* is not so close to the resonance 7 : 1 or to 8 : 1 and the terms of high degree in eccentricities are not so relevant, we have found convenient to limit our expansions to $N_S = 8$, in order to reduce the computational cost of the whole procedure.

Now we have to apply the transformation of coordinates defined by the application of the Lie series operator $\exp(\mathcal{L}_{\chi_1^{(\mathcal{O}2)}}) \cdot = \sum_{j=0}^{\infty}(1/j!)\mathcal{L}^j_{\chi_1^{(\mathcal{O}2)}} \cdot$ to the Hamiltonian. Recalling that in our secular model we will not consider terms depending on $\boldsymbol{L}$ or of order greater than μ^2, the only terms we need to compute are included in the following expansion:

$$\begin{aligned}\widetilde{H} &= H_{3\mathrm{BP}} + \frac{1}{2}\left\{\chi_1^{(\mathcal{O}2)},\ \mathcal{L}_{\chi_1^{(\mathcal{O}2)}} h_{2,0}^{(\mathrm{Kep})}\right\}_{\boldsymbol{L},\boldsymbol{\lambda}} \\ &+\mu \sum_{\substack{s\geq 0,\ j_2\geq 0 \\ 2s+j_2\leq N_S}} D_2^s \left\{\chi_1^{(\mathcal{O}2)}, h_{s;1,j_2}^{(\mathcal{P})}\right\}_{\boldsymbol{L},\boldsymbol{\lambda}} + \frac{\mu}{2} \sum_{\substack{s\geq 0,\ j_2\geq 0 \\ 2s+j_2\leq N_S}} D_2^s \left\{\chi_1^{(\mathcal{O}2)}, h_{s;0,j_2}^{(\mathcal{P})}\right\}_{\boldsymbol{\xi},\boldsymbol{\eta}} ,\end{aligned} \tag{73}$$

where $\{\cdot,\cdot\}_{\boldsymbol{L},\boldsymbol{\lambda}}$ and $\{\cdot,\cdot\}_{\boldsymbol{\xi},\boldsymbol{\eta}}$ are the terms of the Poisson bracket involving only the derivatives with respect to the pairs of conjugate variables $(\boldsymbol{L},\boldsymbol{\lambda})$ and $(\boldsymbol{\xi},\boldsymbol{\eta})$, respectively. Then, according to [27], we have that

$$\langle H^{(\mathcal{O}2)}\rangle_{\boldsymbol{\lambda}}\Big|_{\boldsymbol{L}=\boldsymbol{0}} = \langle \widetilde{H}\,\rangle_{\boldsymbol{\lambda}}\Big|_{\boldsymbol{L}=\boldsymbol{0}} + \mathcal{O}(\mu^3)\,,$$

being $H^{(\mathcal{O}2)} = \exp(\mathcal{L}_{\chi_1^{(\mathcal{O}2)}}) H_{3\mathrm{BP}}$. Let us remark that for the definition of this model it is not necessary to compute the effects induced by the second generating function $\chi_2^{(\mathcal{O}_2)}(\boldsymbol{L},\boldsymbol{\lambda})$ for removing terms linear in $\boldsymbol{L}$, because the additional terms due to the application of such a Lie series operator are neglected in the secular approximation.

We can finally introduce our secular model up to order 2 in the masses by setting

$$H^{(\mathrm{sec})}(D_2,\boldsymbol{\xi},\boldsymbol{\eta}) = \left\lceil\ \langle \widetilde{H}\,\rangle_{\boldsymbol{\lambda}}\Big|_{\boldsymbol{L}=\boldsymbol{0}}\ \right\rceil_{N_S}\,, \tag{74}$$

i.e., we take the averaged expansion (over the fast angles $\boldsymbol{\lambda}$) of the part of $\widetilde{H}$ that is both independent from the actions $\boldsymbol{L}$ and truncated up to a total order of magnitude N_S in eccentricity and inclination. Since D_2 is $\mathcal{O}\big(e_1^2+\mathrm{i}_1^2+e_2^2+\mathrm{i}_2^2\big)$, this means that we keep the Hamiltonian terms $h_{s;0,j_2}^{(\mathcal{P})}$ with $2s+j_2\leq N_S$. From now on, the parameter D_2 is replaced by its explicit value that is calculated as a function of the initial conditions; thus, we can write the Hamiltonian as follows:

$$H^{(\mathrm{sec})}(\boldsymbol{\xi},\boldsymbol{\eta}) = \sum_{s=1}^{N_S/2} h_{2s}^{(\mathrm{sec})}(\boldsymbol{\xi},\boldsymbol{\eta})\,, \tag{75}$$

where h_{2s} is an homogeneous polynomial of degree $2s$. This means that the expansion contains just terms of even degree, as a further consequence of the well known

D'Alembert rules. To fix the ideas, in the case of the extrasolar system HD 4732 let us emphasize that our secular model at order two in the masses is defined by a Hamiltonian $H^{(\mathrm{sec})}$ that is a simple (even) polynomial of maximal degree 8 in the four canonical variables $(\boldsymbol{\xi}, \boldsymbol{\eta})$.

We have explicitly performed all the computations of Poisson brackets (required by Lie series formalism to express canonical transformations) and all the expansions described in the present subsection and in in the next one, by using *Χρόνος*. It is a software package especially designed for doing computer algebra manipulations into the framework of Hamiltonian perturbation theory (see [21] for an introduction to its main concepts).

4.2 Semi-analytic Computations of Invariant Tori

In the framework of Hamiltonian theory for dynamical systems, often intuition can be fruitfully helped by numerical investigations. In particular, in the case of the extrasolar system HD 4732, they allow to easily motivate the new approach that is based on normal forms and we are going to describe. In the present section, we will discuss some results provided by direct numerical integrations of the secular model $H^{(\mathrm{sec})}$ that is defined in (75); all of them have been produced by simply applying the RK4 method.

A few dynamical features of the Hamiltonian model defined by $H^{(\mathrm{sec})}$ are summarized in the plots reported in Fig. 2. They refer, as an example, to the initial conditions corresponding to the set of values $\mathcal{I}_{4^\circ}$, defined in (68). The difference of the arguments of the pericenters $\omega_2 - \omega_1$ is plotted in the bottom-right panel of such

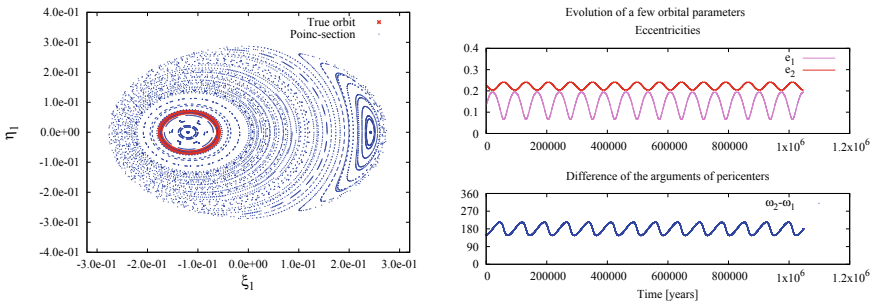


Fig. 2 On the left, Poincaré sections that are corresponding to the hyperplane $\eta_2 = 0$ (with the additional condition $\xi_2 > 0$) and are generated by the flow of the Hamiltonian secular model $H^{(\mathrm{sec})}$, which is given in (75) at order two in the masses for the exoplanetary system HD 4732; the orbit in red refers to the motion starting from the initial conditions corresponding to the set $\mathcal{I}_{4^\circ}$, that is described in (68). On the right, evolution of secular orbital elements: the eccentricities of both the exoplanets (that are e_1 and e_2) and the difference of the arguments of the pericenters (i.e., $\omega_2 - \omega_1$) are plotted on top and bottom, respectively

a figure; then, we can easily appreciate that this angle is librating around 180°. By taking into account of the fact that the nodes are opposite in the Laplace frame, this means that the pericenters of HD 4732b and HD 4732c are in the so called "apsidal locking" regime in the vicinity of the alignment of the pericenters. This phenomenon is expected to play a major role in making stable the orbits for systems where the Keplerian part of the Hamiltonian is strongly affected by the interactions between planets (see, e.g., [30] or [9]). The Poincaré sections of the motions starting from the initial conditions corresponding to $\mathcal{I}_{4^\circ}$ are plotted in red in the panel on the left of Fig. 2 and it is easy to remark that they are orbiting around a fixed point. Moreover, it looks rather close to those sections marked in red, when their distance from such a fixed point is compared with that from the orbits that are enclosing another fixed point. Let us recall that all the Poincaré sections reported in Fig. 2 refer to the same level of energy, say E, corresponding to the set of initial conditions $\mathcal{I}_{4^\circ}$. Since $H^{(\mathrm{sec})}$ is a two degrees of freedom Hamiltonian, the manifold labeled by such a value of the energy will be three-dimensional; in other words, by plotting the Poincaré sections, we automatically reduce by one the dimensions of the orbits. This is the reason why a fixed point actually corresponds to a periodic orbit. Since the fixed point with negative value of the abscissa is surrounded by closed curves, then we can argue that such a periodic orbit is linearly stable for what concerns the transverse dynamics. This means that it is a one-dimensional elliptic torus, in the terminology we have adopted in the present work. Therefore, we can conclude that the orbit generated by the set $\mathcal{I}_{4^\circ}$ of initial conditions is winding around a linearly stable periodic orbit, by remaining in its vicinity. This explains why we are going to adopt a strategy based on two different algorithms: the first one refers to the elliptic torus (that corresponds to a fixed point in the Poincaré sections) and provides a good enough approximation to start the second computational procedure that constructs the final KAM torus (which shall include also the points marked in red in Fig. 2).

Explicit Construction of the Normal Form for Elliptic Tori in the Case of the Secular Model Representing the Planetary System HD 4732

The discussion above has highlighted that it is convenient to adopt a suitable set of coordinates including also a resonant angle, that is the difference of the arguments of the pericenters. In view of such a target, we first introduce the set of action-angle variables $(\boldsymbol{\mathcal{J}}, \boldsymbol{\psi})$ via the canonical transformation

$$\xi_j = \sqrt{2\mathcal{J}_j}\cos\psi_j\ , \qquad \eta_j = \sqrt{2\mathcal{J}_j}\sin\psi_j\ , \qquad \forall\ j = 1, 2, \tag{76}$$

being $(\boldsymbol{\xi}, \boldsymbol{\eta})$ the variables appearing as arguments of the secular Hamiltonian $H^{(\mathrm{sec})}$ defined in (75). It is important to recall that the angles (ψ_1, ψ_2) associated to these secular variables are nearly equal to the arguments of the pericenters (ω_1, ω_2), apart from a small correction due to the transformation of coordinates induced by the application of the Lie series $\exp \mathcal{L}_{\chi_1^{(\mathcal{O}2)}}$ to the Hamiltonian of the three-body planetary problem. Then, it is convenient to introduce a new set of variables $(\boldsymbol{I}, \boldsymbol{\vartheta})$ such that

$$\vartheta_1 = \psi_1 - \psi_2\ , \quad \vartheta_2 = \psi_2\ , \quad I_1 = \mathcal{J}_1\ , \quad I_2 = \mathcal{J}_2 + \mathcal{J}_1\ . \tag{77}$$

We now introduce the new canonical polynomial variables $(\boldsymbol{x}, \boldsymbol{y})$ defined as

$$x_j = \sqrt{2I_j}\cos\vartheta_j\ , \qquad y_j = \sqrt{2I_j}\sin\vartheta_j\ , \qquad \forall\ j = 1, 2\ . \tag{78}$$

Let us also remark that making Poincaré sections with respect to the hyperplane $\eta_2 = 0$, when $\xi_2 > 0$ is equivalent to impose $\psi_2 = 0$, because of the definitions in (76). Therefore, looking at formulæ (77)–(78), one can easily realize that the drawing in the left panel of Fig. 2 can be seen as a plot of the Poincaré sections in coordinates $(x_1\ ,\ y_1)$ with respect to $y_2 = 0$ and with the additional condition $x_2 > 0$. Revisiting the plot in the bottom–right box of Fig. 2 in the context of the new canonical variables is interesting, because it makes clear that ϑ_1 is librating around 180°. In fact, we have that $\vartheta_1 = \psi_1 - \psi_2 \simeq \omega_1 - \omega_2$, because the relation between these differences of angles is given by the transformation induced by the application of the Lie series $\exp\mathcal{L}_{\chi_1^{(\mathcal{O}2)}}$, that is close to the identity.

By a numerical method,[28] we can easily determine the initial condition $(\boldsymbol{x}^\star, \boldsymbol{y}^\star)$ that is in correspondence with a Poincaré section and generates a periodic solution. We can now subdivide the variables in two different couples. The first one is given by $(p, q) \in \mathbb{R} \times \mathbb{T}$, i.e., the action-angle couple describing the periodic motion. Thus, we rename the angle φ_2 as q, while the action is obtained by translating the origin of I_2 so that $p = I_2 - I^\star$, where at the first trial[29] the shift value $I^\star$ is fixed so that $I^\star = \left((x_2^\star)^2 + (y_2^\star)^2\right)/2$. For what concerns the second couple of canonical coordinates, we start from the polynomial variables (x_1, y_1) in order to describe the motion transverse to the periodic orbit. The last preliminary translation is on x_1 , in order to have expansions around the value $x_1^\star$, given by the initial condition computed numerically. Let us emphasize that, since the fixed point we are trying to approximate in Fig. 2 corresponds to $\varphi_1 = 180°$, we have that $y_1^\star = 0$ and here a translation is not needed. It is now convenient to rescale the transverse variables $(\bar{x}_1, y_1)$, being $\bar{x}_1 = x_1 - x_1^\star$, in such a way that the Hamiltonian part which is quadratic in the new variables (x, y) and

[28] Let us imagine to start from an initial condition denoted by $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ that is close enough to the periodic orbit generated by the wanted solution $(\boldsymbol{x}^\star, \boldsymbol{y}^\star)$; typically, at the beginning one can put $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ equal to the values assumed by the canonical variables $(\boldsymbol{x}, \boldsymbol{y})$ in correspondence with the set $\mathcal{I}_{i_{\rm mut}(0)}$, defined in (68). During a long enough numerical integration of the Hamilton equations related to $H^{(\rm sec)}$, one can easily determine $\hat{x}_{1,-}$ and $\hat{x}_{1,+}$ that are the minimum value assumed by the variable x_1 in correspondence with the Poincaré sections and the maximum one, resp. If the difference $\hat{x}_{1,+} - \hat{x}_{1,-}$ is below a prescribed (small) threshold of tolerance, then we assume to know the solution with a good enough level of approximation and we stop this computational procedure by setting $(\boldsymbol{x}^\star, \boldsymbol{y}^\star) = (\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$. If such a "way out condition" is not satisfied, then we define $x_1^\star = (\hat{x}_{1,+} + \hat{x}_{1,-})/2$, $y_1^\star = 0$, $y_2^\star = 0$ and we determine the positive value of $x_2^\star$ so that the energy level of this new approximation of the final solution, i.e., $(\boldsymbol{x}^\star, \boldsymbol{y}^\star)$, is still equal to the value E corresponding to the set $\mathcal{I}_{i_{\rm mut}(0)}$. Let us remark that in the (re)definition of $(\boldsymbol{x}^\star, \boldsymbol{y}^\star)$ we are exploiting both the definition of the Poincaré sections and their symmetry with respect to the axis of the abscissas. At this point, we put $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) = (\boldsymbol{x}^\star, \boldsymbol{y}^\star)$ and we restart the computational procedure by performing another numerical integration so to determine new values of $\hat{x}_{1,-}$ and $\hat{x}_{1,+}$ and so on, until the "way out condition" will be satisfied.

[29] See the discussion about the solution of the implicit equation (79) by using the Newton method, which is reported at the end of these explanations.

does not depend on (p, q) is in the form $\Omega^{(0)}(x^2 + y^2)/2$. This rescaling can be done by a canonical transformation as the quadratic part does not have any mixed term $\bar{x}_1 y_1$ and the coefficients of $\bar{x}_1^2$ and y_1^2 have the same sign, because of the proximity to an elliptic equilibrium point. Thus, since such a quadratic part is in the preliminary form $a\bar{x}_1^2 + by_1^2$, it suffices to define the new variables (x, y) as $x = \sqrt[4]{\frac{a}{b}}\,\bar{x}_1$, $y = \sqrt[4]{\frac{b}{a}}\,y_1$. Finally, we introduce the second pair of canonical coordinates $(J, \varphi) \in \mathbb{R}_+ \cup \{0\} \times \mathbb{T}$ so that $x = \sqrt{2J}\cos\varphi$ and $y = \sqrt{2J}\sin\varphi$.

In the case of the secular dynamics of the planetary system HD 4732, starting from $H^{(\mathrm{sec})}$ in (75), we have applied all the canonical transformations listed above and we have expanded the Hamiltonian $\mathcal{H}^{(0)}(p, q, J, \varphi)$ up to degree 16 in the square roots of the actions (p, J). Since $\mathcal{H}^{(0)}(p, q, J, \varphi)$ is in a suitable form to apply the algorithm fully described in Sect. 3.1 in the case with $n_1 = n_2 = 1$ (this is the reason why all the variables (p, q, J, φ) are here denoted as scalar quantities instead of vectorial ones), we have applied such a computational procedure. We have performed 19 steps of the normalization algorithm so producing $\mathcal{H}^{(19)}(p, q, J, \varphi)$. During those computations, the Fourier expansions in q of all the Hamiltonians defined by the algorithm have been truncated at a maximal trigonometric degree equal to 40; since $K = 2$, this choice allows to properly determine the generating functions for the first 20 normalization steps. For the sake of brevity, we omit to report the graphs of the norms of all the generating functions that are defined by the normalization procedure, also because those plots are similar to the corresponding ones included in [8, 9]. Indeed, they show that the convergence to the identity of the canonical transformations defined at the r-th step of the algorithm is very fast with respect to r. This fact also allows to iterate a few times all the normalization procedure constructing the normal form for an elliptic torus with a computational cost which is not too expensive. We are interested in doing that in order to refine the choice of the initial shift value $I^\star$. Since all other canonical transformations are unambiguously defined, we have some remaining arbitrariness just on the translation $p = I_2 - I^\star$. We finally determine $I^\star$ in such a way that

$$\mathcal{E}^{(19)}\big(I^\star\big) = E\,, \tag{79}$$

where E is the energy level of the Poincaré sections and $\mathcal{E}^{(19)}\big(I^\star\big)$ is the energy of the elliptic torus in the approximation provided after 19 steps of normalization. The implicit equation above can be numerically solved in the unknown $I^\star$ by iterating a few times the Newton method; this is done starting from the initial guess $\big((x_2^\star)^2 + (y_2^\star)^2\big)/2$, according with the discussion above.

For brevity, we omit also the tests showing that there is an excellent agreement between the wanted periodic orbit and the nearly invariant curve, which is provided by the last execution of the normalization algorithm, that is launched during the final iteration of the Newton method targeting the solution of (79). Actually, it corresponds to the counter-image of the set $(p = 0,\ q \in \mathbb{T},\ J = 0,\ \varphi = 0)$ and is expressed in the coordinates $(\boldsymbol{\xi}, \boldsymbol{\eta})$, after having composed all the previous canonical transformations.

Explicit Construction of the Normal Form for KAM Tori in the Case of the Secular Model Representing the Planetary System HD 4732

Since the Hamiltonian $\mathcal{H}^{(19)}(p, q, J, \varphi)$ is very close to the normal form related to the wanted elliptic torus, we use it as the starting point to construct a semi-analytic solution that should provide a good approximation of the orbits generated by the initial conditions corresponding to the set $\mathcal{I}_{4^\circ}$. For such a purpose, first we translate once again the coordinates. This is made in such a way that the new invariant torus we are going to construct will be located in the proximity of these initial conditions; therefore, we define two new pairs of action-angle coordinates $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{R}^2 \times \mathbb{T}^2$. It is convenient to set $p_2 = J - J^\star$, being $J^\star$ the value of the momentum J computed in correspondence with the initial conditions related to the set $\mathcal{I}_{4^\circ}$, that generate the Poincaré sections marked in red in Fig. 2. We also introduce $p_1 = p - p^\star$, with $p^\star = -\left(\Omega^{(19)}/\omega^{(19)}\right) J^\star$, being $2\pi/\omega^{(19)}$ approximately equal to the period of the motion on the previously determined one-dimensional elliptic torus, while the angular velocity of the transverse (small) oscillations in its vicinity is close to $\omega^{(19)}$. We recall that the values of both $\omega^{(19)}$ and $\Omega^{(19)}$ appear in the expansion (59) of the Hamiltonian $\mathcal{H}^{(19)}$, that is provided at the end of the previous normalization algorithm. Moreover, we rename the angles (q, φ) as (q_1, q_2), respectively; then, we perform the two translations described just above, by expanding the new Hamiltonian $H^{(0)}(\boldsymbol{p}, \boldsymbol{q})$ up to degree 8 in the actions $\boldsymbol{p}$. By considering just the integrable approximations of $\mathcal{H}^{(19)}$ and $H^{(0)}$ (this means that the terms depending by the angles are temporarily neglected), one can easily realize that the energy constant $E^{(0)}$ corresponding to the new Hamiltonian is such that $E^{(0)} \simeq E$, because of the equation $\omega^{(19)} p^\star + \Omega^{(19)} J^\star = 0$ that is due to the definitions of the shift values $\left(p^\star, J^\star\right)$. Since $H^{(0)}(\boldsymbol{p}, \boldsymbol{q})$ is in a suitable form to apply the algorithm fully described in Sect. 2.3, we have performed 19 steps of such a computational procedure too, so producing $H^{(19)}(\boldsymbol{p}, \boldsymbol{q})$. During these computations, the Fourier expansions in q of all the Hamiltonians defined by the normalization algorithm have been truncated at a maximal trigonometric degree equal to 40. This choice allows to properly determine the generating functions $\chi_1^{(r)}$ and $\chi_2^{(r)}$ for the first 20 normalization steps.

It is convenient to define the norms of the generating functions as the sum of the absolute values of the coefficients appearing in their (finite) Taylor-Fourier expansions. In the left panel of Fig. 3, we report the plot of $\left\|\chi_2^{(r)}\right\|$ in a semi-log scale and as a function of the normalization step r, while we have decided to not include also $\left\|\chi_1^{(r)}\right\|$, because for every r it is definitely smaller than $\left\|\chi_2^{(r)}\right\|$. One can appreciate that the geometrical decrease of the generating functions is very sharp and regular; therefore, this shows that the normalization algorithm constructing the Kolmogorov normal form is convergent in a quite rapid way.

We can now check the quality of our results. Let us denote with $\mathcal{C}$ the canonical transformation we obtain by composing all the changes of coordinates we have discussed in the present Sect. 4.2. Therefore, we have that $(\boldsymbol{\xi}, \boldsymbol{\eta}) = \mathcal{C}(\boldsymbol{p}, \boldsymbol{q})$, where $(\boldsymbol{\xi}, \boldsymbol{\eta})$ are the canonical coordinates referring to the Hamiltonian secular model $H^{(\mathrm{sec})}$, that is defined in (75), while $(\boldsymbol{p}, \boldsymbol{q})$ are the action-angle variables that are introduced at the end of the previously described computational procedure. Inspired

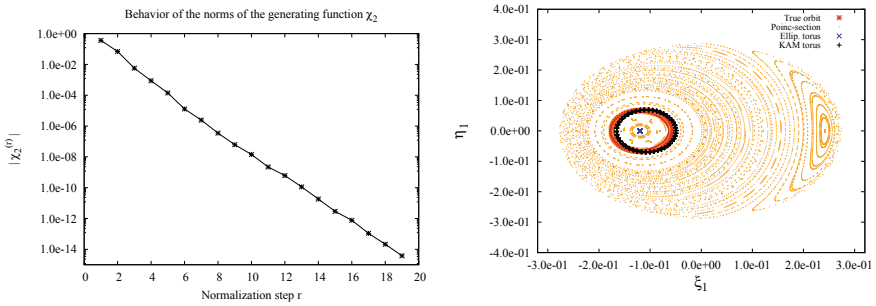


Fig. 3 On the left, study of the decrease of $\left\| \chi_2^{(r)} \right\|$ as a function of the normalization step r. On the right, comparisons between the Poincaré sections generated by two different initial conditions, that correspond to the set $\mathcal{I}_{4^\circ}$ and a point on the (nearly) invariant torus $\boldsymbol{p} = \boldsymbol{0}$ related to the Hamiltonian $H^{(19)}$, respectively. The former ones are marked in red as in the left panel of Fig. 2, while the latter ones are in black. The Poincaré sections are defined in the same way as those reported in Fig. 2; in particular, the dots plotted in blue there are located exactly in the same positions as those marked in orange here. The blue symbol $\times$ refers to the motion on the elliptic torus corresponding to the Hamiltonian $\mathcal{H}^{(19)}$

by the semi-analytic scheme (5), which provides a way to integrate the Hamilton equations, we start by computing $(\boldsymbol{\xi}(0), \boldsymbol{\eta}(0)) = \mathcal{C}(\boldsymbol{0}, \boldsymbol{0})$. Since $H^{(19)}$ is very close to be in Kolmogorov normal form and $H^{(\mathrm{sec})} \simeq H^{(19)}\big(\mathcal{C}(\boldsymbol{p}, \boldsymbol{q})\big)$ (the discrepancies are mainly due to the unavoidable truncations that are made on the expansions of the Hamiltonians), then $\big(\boldsymbol{\xi}(t), \boldsymbol{\eta}(t)\big) = \big(\mathcal{C}(\boldsymbol{0}, \boldsymbol{\omega}^{(19)} t)\big)$ provides a good approximation of the flow induced by $H^{(\mathrm{sec})}$. We also recall that the values of the angular velocity vector $\boldsymbol{\omega}^{(19)}$ appear in the expansion (21) of the Hamiltonian $H^{(19)}$. Computing the Poincaré sections of the motion law $\big(\mathcal{C}(\boldsymbol{0}, \boldsymbol{\omega}^{(19)} t)\big)$ is not very comfortable; therefore, it is convenient to refer to its approximation which is given by the numerical solution of the Hamilton equations for $H^{(\mathrm{sec})}$ starting form the initial conditions $(\boldsymbol{\xi}(0), \boldsymbol{\eta}(0)) = \mathcal{C}(\boldsymbol{0}, \boldsymbol{0})$. The Poincaré sections we have obtained in this way are plotted in black on the right panel of Fig. 3. They are in good agreement with the the Poincaré sections marked in red in both Figs. 2 and 3, that refer again to the flow induced by $H^{(\mathrm{sec})}$, but starting from the initial conditions related to the set $\mathcal{I}_{4^\circ}$. This confirms that we are able to obtain reliable approximations of the secular motions for extrasolar planetary systems, by using computational procedures based on the construction of suitable (Kolmogorov-like) normal forms.

Final Comments About Our Semi-analytic Results

Looking closely at the right panel of Fig. 3, one can observe that the Poincaré sections plotted in black goes from the part internal to the orbit in red to the external one and vice versa. This provides a clear indication that the energy level of the final KAM torus (that is $\simeq E^{(19)}$) is not very close to that of all the Poincaré sections plotted in Fig. 2 (being E its value). Indeed, the relative error $\big|E^{(19)} - E\big|/|E|$ is about 12 %. The agreement between the results produced by the purely numerical integrations or by adopting our semi-analytical approach can be strongly improved

by a suitable further refinement of our computational procedure. The description of such an extension goes beyond the scopes of the present work, but we stress that it can be done so as to ensure also that the condition on the coherence with the energy of the Poincaré sections, i.e.,

$$E^{(R_{\rm I})} = E \ , \tag{80}$$

is satisfied within a tolerance range that is acceptable for a numerical solution of the equation above, where $R_{\rm I}$ is the number of steps that are explicitly performed in order to construct the final Kolmogorov normal form. Here, we limit ourselves to anticipate some of the results that can be obtained by implementing that further refinement, in order to let the reader appreciate the power of this kind of methods. For what concerns the planetary system HD 4732 we already have studied the motions starting from the following sets of initial conditions: $\mathcal{I}_{2^\circ}$, $\mathcal{I}_{4^\circ}$, $\mathcal{I}_{6^\circ}$, ... $\mathcal{I}_{40^\circ}$. We can construct invariant KAM tori well approximating the orbits for all these cases, except those corresponding to the sets $\mathcal{I}_{32^\circ}$ and $\mathcal{I}_{34^\circ}$. We emphasize that these limitations are due to real dynamical phenomena. The Poincaré sections generated by those initial conditions clearly shows that between 34° and 36° there is the transition from the librations to the circulation regime, for what concerns the difference of the argument of the pericenters. Moreover, this kind of orbits are observed in stable situations up to initial values of the mutual inclinations that are about 40°, while for even larger angles there are robust configurations just inside the Lidov-Kozai resonance, which has different dynamical features (see [42]). As we have already mentioned above, we plan to describe these new results in a forthcoming work.

The evolution of the eccentricities plotted in the right panel of Fig. 2 clearly shows that their average value is larger than 0.1 for both the exoplanets orbiting around HD 4732. Therefore, the new approach that we have introduced in the present work behaves definitely better with respect to the previous one, which was described in [41] and was shown to be successful just for systems with exoplanetary eccentricities smaller than 0.1 . In our opinion the main source of improvement is due to the new strategy, because it combines the preliminary construction of the normal form for a suitable elliptic torus with the final one, which is performed in its vicinity for a KAM torus whose shape is a good approximation of the secular orbits. In order to mention another relevant success of our new approach, let us stress that in [9] we applied it also to the delicate case of a system including both the two largest exoplanets orbiting around υ Andromedæ A and the star itself.

Acknowledgements This work was partially supported by the project MIUR-PRIN 20178CJA2B "New frontiers of Celestial Mechanics: theory and applications". M.V. thanks the ASI Contract n. 2018-25-HH.0 (Scientific Activities for JUICE, C/D phase). Moreover, we extend our gratitude also to the MIUR Excellence Department Project awarded to the Department of Mathematics of the University of Rome "Tor Vergata" (CUP E83C18000100006), which made available the computational resources we exploited.

References

1. Arnold, V.I.: Proof of a theorem of A. N. Kolmogorov on the invariance of quasi–periodic motions under small perturbations of the Hamiltonian. Usp. Mat. Nauk. **18**, 13 (1963). Engl. transl. in: Russ. Math. Surv. **18**, 9 (1963)
2. Beaugé, C., Ferraz-Mello, S., Michtchenko, T.A.: Multi-planet extrasolar systems – detection and dynamics. Res. Astron. Astroph. **12**, 1044–1080 (2012). http://www.raa-journal.org/raa/index.php/raa/article/view/1199
3. Biasco, L., Chierchia, L., Valdinoci, E.: Elliptic two-dimensional invariant tori for the planetary three-body problem. Arch. Rational Mech. Anal. **170**, 91–135 (2003). https://link.springer.com/article/10.1007/s00205-003-0269-2
4. Biasco, L., Chierchia, L., Valdinoci, E.: N-dimensional elliptic invariant tori for the planar (N+1)-body problem. SIAM J. Math. Anal. **37**, 1560–1588 (2006). https://epubs.siam.org/doi/10.1137/S0036141004443646
5. Benettin, G., Galgani, L., Giorgilli, A., Strelcyn, J.M.: A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. Nuovo Cimento **79**, 201–223 (1984)
6. Berti, M., Biasco, L.: Branching of Cantor manifolds of elliptic tori and applications to PDEs. Comm. Math. Phys. **305**, 741–796 (2011). https://link.springer.com/article/10.1007/s00220-011-1264-3
7. Caracciolo, C.: Normal form for lower dimensional elliptic tori: convergence of a constructive algorithm. Math. Engin. **4**(6), 1–40 (2021). https://www.aimspress.com/article/doi/10.3934/mine.2022051
8. Caracciolo C., Locatelli, U.: Elliptic tori in FPU non-linear chains with a small number of nodes. Commun. Nonlinear Sc. Numer. Simulat. **97**, 105759 (2021). https://doi.org/10.1016/j.cnsns.2021.105759
9. Caracciolo, C., Locatelli, U., Sansottera, M., Volpi, M.: Librational KAM tori in the secular dynamics of the υ Andromedæ planetary system. Mon. Not. Royal Astron. Soc. **510**, 2147–2166 (2022). https://doi.org/10.1093/mnras/stab3514
10. Celletti, A., Chierchia, L.: KAM stability and Celestial Mechanics. Memoirs AMS **187**, 878 (2007). https://www.ams.org/books/memo/0878/
11. Chierchia, L.: Kolmogorov's 1954 paper on nearly-integrable Hamiltonian systems. Reg. Chaot. Dyn. **13**, 130–139 (2008). https://doi.org/10.1134/S1560354708020056
12. de la Llave, R., González, A., Jorba, À, Villanueva, J.: KAM theory without action-angle variables. Nonlinearity **18**, 855–895 (2005). https://iopscience.iop.org/article/10.1088/0951-7715/18/2/020
13. Eliasson L.H.: Perturbations of stable invariant tori for Hamiltonian systems. Ann. Scuola Norm. Sup. Pisa, Cl. Sci., IV Ser., **15**, 115–147 (1988)
14. Gröbner, W.: Die Lie-Reihen und Ihre Anwendungen. Springer, Berlin (1960). Italian transl.: Le serie di Lie e le loro applicazioni. Cremonese, Roma (1973)
15. Giorgilli, A.: Notes on exponential stability of Hamiltonian systems. In: Dynamical Systems, Part I. Pubbl. Cent. Ric. Mat. Ennio De Giorgi, Sc. Norm. Sup. Pisa, 87–198 (2003). https://www.springer.com/gp/book/9788876422942
16. Giorgilli, A., Delshams, A., Fontich, E., Galgani, L., Simó, C.: Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem. J. Diff. Equ. **77**, 167–198 (1989). https://www.sciencedirect.com/science/article/pii/0022039689901617
17. Giorgilli, A., Locatelli, U.: Kolmogorov theorem and classical perturbation theory. ZAMP **48**, 220–261 (1997). https://link.springer.com/article/10.1007/PL00001475
18. Giorgilli, A., Locatelli, U., Sansottera, M.: Kolmogorov and Nekhoroshev theory for the problem of three bodies. Cel. Mech. Dyn. Astr. **104**, 159–173 (2009). https://link.springer.com/article/10.1007/s10569-009-9192-7

19. Giorgilli, A., Locatelli, U., Sansottera, M.: On the convergence of an algorithm constructing the normal form for lower dimensional elliptic tori in planetary systems. Cel. Mech. Dyn. Astr. **119**, 397–424 (2014). https://doi.org/10.1007/s10569-014-9562-7
20. Giorgilli, A., Locatelli, U., Sansottera, M.: Secular dynamics of a planar model of the Sun-Jupiter-Saturn-Uranus system; effective stability in the light of Kolmogorov and Nekhoroshev theories. Regul. Chaotic Dyn. **22**, 54–77 (2017). https://doi.org/10.1134/S156035471701004X
21. Giorgilli, A., Sansottera, M.: Methods of algebraic manipulation in perturbation theory. In: Cincotta, P.M., Giordano, C.M., Efthymiopoulos, C. (eds.) "Chaos, Diffusion and Non-integrability in Hamiltonian Systems – Applications to Astronomy", Proceedings of the Third La Plata International School on Astronomy and Geophysics. Universidad Nacional de La Plata and Asociación Argentina de Astronomía Publishers, La Plata (2012)
22. Kolmogorov, A.N.: Preservation of conditionally periodic movements with small change in the Hamilton function. Dokl. Akad. Nauk SSSR **98**, 527–530 (1954). Engl. transl. in: Los Alamos Scientific Laboratory translation LA-TR-71-67; reprinted in: Lecture Notes in Physics **93**, 51–56. Springer (1979)
23. Laskar, J.: Secular evolution of the Solar System over 10 million years. Astron. Astroph. **198**, 341–362 (1988). http://articles.adsabs.harvard.edu/pdf/1988A%26A...198..341L
24. Laskar, J.: Les variables de Poincaré et le développement de la fonction perturbatrice. Groupe de travail sur la lecture des Méthodes nouvelles de la Mécanique Céleste, Notes scientifiques et techniques du Bureau des Longitudes S026 (1989). https://www.imcce.fr/content/medias/publications/publications-recherche/nst/docs/S026.pdf
25. Leontovich, A.M.: On the stability of the Lagrange periodic solutions for the reduced problem of three bodies. Soviet Math. Dokl. **3**, 425 (1962)
26. Libert, A.-S., Sansottera, M.: On the extension of the Laplace-Lagrange secular theory to order two in the masses for extrasolar systems. Cel. Mech. Dyn. Astr. **117**, 149–168 (2013). https://link.springer.com/article/10.1007%2Fs10569-013-9501-z
27. Locatelli U., Giorgilli, A.: Invariant tori in the secular motions of the three-body planetary systems. Cel. Mech. Dyn. Astr. **78**, 47–74 (2000). https://link.springer.com/article/10.1023/A:1011139523256
28. Locatelli U., Giorgilli, A.: Invariant tori in the Sun–Jupiter–Saturn system. Discr. Cont. Dyn. Sys. – B **7**, 377–398 (2007). https://www.aimsciences.org/article/doi/10.3934/dcdsb.2007.7.377
29. Melnikov, V.K.: On some cases of conservation of almost periodic motions with a small change of the Hamiltonian function. Dokl. Akad. Nauk SSSR **165**, 1245–1248 (1965)
30. Michtchenko, T.A., Malhotra, R.: Secular dynamics of the three-body problem: application to the υ Andromedæ planetary system. Icarus **168**, 237–248 (2004). https://www.sciencedirect.com/science/article/abs/pii/S0019103503004287?via%3Dihub
31. Morbidelli, A., Giorgilli, A.: Superexponential stability of KAM tori. J. Stat. Phys. **78**, 1607–1617 (1995). https://link.springer.com/article/10.1007/BF02180145
32. Mugrauer, M: Search for stellar companions of exoplanet host stars by exploring the second ESA-Gaia data release. Mon. Not. Royal Astron. Soc. **490**, 5088–5102 (2019). https://academic.oup.com/mnras/article/490/4/5088/5622591
33. Moser, J.: On invariant curves of area–preserving mappings of an annulus. Nachr. Akad. Wiss. Gött., Math. Phys. **1**, 1–20 (1962)
34. Poincaré, H.: *Les méthodes nouvelles de la Mécanique Céleste*, Gauthier–Villars, Paris (1892), reprinted by Blanchard (1987)
35. Pöschel, J.: Integrability of Hamiltonian systems on Cantor sets. Comm. Pure Appl. Math. **25**, 653–695 (1982)
36. Pöschel, J.: On elliptic lower dimensional tori in Hamiltonian sytems. Math. Z. **202**, 559–608 (1989). https://link.springer.com/article/10.1007/BF01221590
37. Sansottera, M., Danesi, V.: Kolmogorov variation: KAM with knobs (à la Kolmogorov). Submitted (2022)
38. Sansottera, M., Libert, A.-S.: Resonant Laplace-Lagrange theory for extrasolar systems in mean-motion resonance. Cel. Mech. Dyn. Astr. **131**, 38 (2019). https://link.springer.com/article/10.1007/s10569-019-9913-5

39. Sato, B., et al.: A double planetary system around the evolved intermediate-mass star HD 4732. Astroph. J. **762**, 9 (2013). https://doi.org/10.1088/0004-637X/762/1/9
40. Valvo, L., Locatelli, U.: Hamiltonian Control of Magnetic Field Lines: Computer Assisted Results Proving the Existence of KAM Barriers. J. Comput. Dyn. **9** (2022). https://www.aimsciences.org/article/doi/10.3934/jcd.2022002
41. Volpi, M., Locatelli, U., Sansottera, M.: A reverse KAM method to estimate unknown mutual inclinations in exoplanetary systems, Cel. Mech. Dyn. Astr. **130**, 36 (2018). https://link.springer.com/article/10.1007/s10569-018-9829-5
42. Volpi, M., Roisin A., Libert, A.-S.: On the 3D secular dynamics of radial-velocity-detected planetary systems, Astron. Astroph. **626**, A74 (2019). https://www.aanda.org/articles/aa/full_html/2019/06/aa34896-18/aa34896-18.html

A New Analysis of the Three-Body Problem

Jérôme Daquin, Sara Di Ruzza and Gabriella Pinzari

Abstract In the recent papers [5, 18], respectively, the existence of motions where the perihelions afford periodic oscillations about certain equilibria and the onset of a topological horseshoe have been proved. Such results have been obtained using, as neighbouring integrable system, the so-called two-centre (or *Euler*) problem and a suitable canonical setting proposed in [16, 17]. Here we review such results.

Keywords Two-centers problem · Three-body problem · Renormalizable integrability · Perihelion librations · Chaos

1 Overview

In the recent papers [5, 18] the existence, in the three-body problem (3BP), of motions which by no means can be regarded as "extending" in some way Keplerian motions has been proved. Indeed, the motions found in those papers can be better understood as continuations of the motions of the so-called *two-centre problem* (or *Euler problem*; 2CP from now on).

The motivation that pushed such researches was a new analysis of 2CP carried out in [17], combined with a remarkable property—which we called *renormalizable integrability*—pointed out in [16]. It relates the "simply averaged Newtonian potential" (see the precise definition below) and the function, which in this paper we shall refer to as *Euler integral*, that makes the 2CP integrable. Roughly, such property

J. Daquin
naXys Research Institute, University of Namur, Namur, Belgium
e-mail: jerome.daquin@unamur.be

S. Di Ruzza
Department of Mathematics, University of Palermo, Palermo, Italy
e-mail: sara.diruzza@unipa.it

G. Pinzari (✉)
Department of Mathematics, University of Padua, Padua, Italy
e-mail: gabriella.pinzari@math.unipd.it

G. Baù et al. (eds.), *New Frontiers of Celestial Mechanics: Theory and Applications*,
Springer Proceedings in Mathematics & Statistics 399,
https://doi.org/10.1007/978-3-031-13115-8_2

states that the averaged Newtonian potential and the Euler integral have the same motions, as they are one a function of the other. As, on the other hand, the motions of the Euler integral are, at least qualitatively, explicit, and the averaged Newtonian potential is a prominent part of the 3BP Hamiltonian, the papers [5, 18] gave partial answers to the natural question whether the motions of the Euler integral can be traced in 3BP. Let us introduce some mathematical tools in order to make our statements more precise.

In terms of Jacobi coordinates [10] the three-body problem Hamiltonian with masses 1, μ, κ is the translation-free function

$$\mathrm{H_J} = \frac{\|\mathbf{y}\|^2}{2}\left(1+\frac{1}{\mu}\right) + \frac{\|\mathbf{y}'\|^2}{2}\left(\frac{1}{1+\mu}+\frac{1}{\kappa}\right) - \frac{\mu}{\|\mathbf{x}\|} - \frac{\mu\kappa}{\|\mathbf{x}' - \frac{1}{1+\mu}\mathbf{x}\|} - \frac{\kappa}{\|\mathbf{x}' + \frac{\mu}{1+\mu}\mathbf{x}\|}\,.$$

Here, $(\mathbf{y}', \mathbf{y}, \mathbf{x}', \mathbf{x}) \in (\mathbb{R}^3)^4$ (or $(\mathbb{R}^2)^4$, in the planar case), $\|\cdot\|$ denotes Euclidean norm and the gravity constant has been taken equal to one, by a proper choice of the units system. We rescale impulses and positions

$$\mathbf{y} \to \frac{\mu}{1+\mu}\mathbf{y}\,, \quad \mathbf{x} \to (1+\mu)\mathbf{x}\,, \quad \mathbf{y}' \to \mu\beta\mathbf{y}'\,, \quad \mathbf{x}' \to \beta^{-1}\mathbf{x}', \tag{1}$$

multiply the Hamiltonian by $\frac{1+\mu}{\mu}$ (by a rescaling of time) and obtain

$$\mathrm{H_J} = \frac{\|\mathbf{y}\|^2}{2} - \frac{1}{\|\mathbf{x}\|} + \gamma\left(\frac{\|\mathbf{y}'\|^2}{2} - \frac{\overline{\beta}}{\beta+\overline{\beta}}\frac{1}{\|\mathbf{x}'-\beta\mathbf{x}\|} - \frac{\beta}{\beta+\overline{\beta}}\frac{1}{\|\mathbf{x}'+\overline{\beta}\mathbf{x}\|}\right), \tag{2}$$

with

$$\gamma = \frac{\kappa^3(1+\mu)^4}{\mu^3(1+\mu+\kappa)}\,, \quad \beta = \frac{\kappa^2(1+\mu)^2}{\mu^2(1+\mu+\kappa)}\,, \quad \overline{\beta} = \mu\beta\,. \tag{3}$$

Likewise, one might consider the problem written in the so-called 1-centric coordinates. In that case,

$$\mathrm{H_0} = \frac{\|\mathbf{y}\|^2}{2} - \frac{1}{\|\mathbf{x}\|} + \gamma\left(\frac{\|\mathbf{y}'\|^2}{2} - \frac{\beta}{\beta+\overline{\beta}}\frac{1}{\|\mathbf{x}'\|} - \frac{\overline{\beta}}{\beta+\overline{\beta}}\frac{1}{\|\mathbf{x}'-(\beta+\overline{\beta})\mathbf{x}\|}\right) + \overline{\beta}\,\mathbf{y}'\cdot\mathbf{y}, \tag{4}$$

with γ, β and $\overline{\beta}$ analogous to (3), up to replace the factors $(1+\mu+\kappa)$ with $(1+\kappa)$. Note that we are not assuming $\mu, \kappa \ll 1$ (in fact, in our applications, we shall make different choices), which means that Jacobi or 1-centric coordinates above are not necessarily centered at the most massive body. In order to simplify the analysis, we introduce a main assumption. Both the Hamiltonians $\mathrm{H_J}$ and $\mathrm{H_0}$ in (2) and (4) include

the Keplerian term

$$\mathrm{J}_0 := \frac{\|\mathbf{y}\|^2}{2} - \frac{1}{\|\mathbf{x}\|} = -\frac{1}{2\Lambda^2}\,. \tag{5}$$

We assume that J_0 is a "leading" term in such Hamiltonians. By averaging theory, this assumption allows us to replace (at the cost of a small error) H_J and H_0 with their respective ℓ-averages

$$\overline{\mathrm{H}}_i = -\frac{1}{2\Lambda^2} + \gamma\widehat{\mathrm{H}}_i \tag{6}$$

with $i = \mathrm{J}, 0$, where ℓ is the mean anomaly associated to (5), and[1]

$$\begin{aligned}
\widehat{\mathrm{H}}_\mathrm{J} &:= \frac{\|\mathbf{y}'\|^2}{2} - \frac{\overline{\beta}}{\beta+\overline{\beta}}\mathrm{U}_\beta - \frac{\beta}{\beta+\overline{\beta}}\mathrm{U}_{-\overline{\beta}} \\
\widehat{\mathrm{H}}_0 &:= \frac{\|\mathbf{y}'\|^2}{2} - \frac{\overline{\beta}}{\beta+\overline{\beta}}\mathrm{U}_{\beta+\overline{\beta}} - \frac{\beta}{\beta+\overline{\beta}}\frac{1}{\|\mathbf{x}'\|}
\end{aligned} \tag{7}$$

with

$$\mathrm{U}_\beta := \frac{1}{2\pi}\int_0^{2\pi} \frac{d\ell}{\|\mathbf{x}' - \beta\mathbf{x}(\ell)\|}. \tag{8}$$

In these formulae, the term $-\frac{1}{2\Lambda^2}$ will be referred to as "Keplerian term", while terms of the form $-\frac{1}{\|\mathbf{x}'-\beta\mathbf{x}\|}$ will be called "Newtonian potentials". Therefore, U_β will be called "averaged Newtonian potential". What we want to underline in that respect is that the averages (6) are "simple", i.e., computed with respect to only one mean anomaly. Most often, in the literature double averages are considered; e.g. [4, 7, 8, 11, 14, 15].

Whether and at which extent the Hamiltonians (6) are good approximations of (2), (4) is a demanding question, as, besides the mass parameters μ, κ, also the region of phase space which is being considered plays a crucial rôle. We limit ourselves to some heuristics, focusing, in particular, on the case considered in [5]. Here the Hamiltonian (2) has been investigated, with $\mu = 1 \gg \kappa$, or, equivalently, $\beta = \overline{\beta} \gg 1$ (see (19) for the precise values). Physically, this corresponds to a couple of asteroids with equal mass interacting with a star, with $\mathbf{x}$ being the relative distance of the twin asteroids, and $\mathbf{x}'$ the distance of the star from their center of mass. In that case, the region of phase space was chosen so that $\|\mathbf{x}'\| > \beta\|\mathbf{x}\|$, so that the two denominators of the Newtonian potentials do not vanish. Expanding such Newtonian potentials in powers of $\frac{\beta a}{\mathrm{r}}$, where $a = \Lambda^2$, $\mathrm{r} := \|\mathbf{x}'\|$, one sees that the lowest order terms depending on

[1] Remark that $\mathbf{y}(\ell)$ has vanishing ℓ-average so that the last term in (4) does not survive.

ℓ have size $\frac{\gamma\beta a}{r^2} \sim \frac{\kappa^3 a}{r^2}$ (as $\beta \sim \kappa$, $\gamma \sim \kappa^2$). So, such terms are negligible compared to the size $\frac{1}{a}$ of the Keplerian term, provided that $\frac{\kappa^{3/2}a}{r} \ll 1$.

We now turn to describe the main features of the Hamiltonians (6). Neglecting the Keplerian term, which is an inessential additive constant for $\overline{H}_i$ and reabsorbing the constant γ with a time change, we are led to look at the Hamiltonians $\widehat{H}_i$ in (7), which, from now on, will be our object of study. Without loss[2] of generality, we fix the constant action Λ to 1.

For definiteness and simplicity, we describe the setting in the case of the planar problem, in which case, after reducing the SO(2) symmetry, $\widehat{H}_i$ have 2 degrees-of-freedom; all the generalisations to the spatial problem being described in Sect. 2. To describe the coordinates we used, we denote as $\mathbb{E}$ the Keplerian ellipse generated by Hamiltonian (5), for negative values of the energy. Assume $\mathbb{E}$ is not a circle. Remark that, as the mean anomaly ℓ is averaged out, we loose any information concerning the position of $\mathbf{x}$ on $\mathbb{E}$, so we shall only need two couples of coordinates for determining the shape of $\mathbb{E}$ and the vectors $\mathbf{y}'$, $\mathbf{x}'$. These are:

- the "Delaunay couple" (G, g), where G is the Euclidean length of $\mathbf{x} \times \mathbf{y}$ and g detects the perihelion. We remark that g is measured with respect to $\mathbf{x}'$ (instead of with respect to a fixed direction), as the SO(2) reduction we use fixes a rotating frame which moves with $\mathbf{x}'$ (compare the formulae in (37));
- the "radial–polar couple"(R, r), where $r := \|\mathbf{x}'\|$ and $R := \frac{\mathbf{y}'\cdot\mathbf{x}'}{\|\mathbf{x}'\|}$.

We now describe what we mean by *renormalizable integrability* [16]. Note first that, in terms of the coordinates above, the functions $U_\beta(r, G, g)$ in (8) depend on (r, G, g) and remark the homogeneity property

$$U_\beta(r, G, g) = \beta^{-1} U(\beta^{-1} r, G, g) \qquad \text{where} \quad U := U_1 \,. \tag{9}$$

By *renormalizable integrability* we mean that there exists a function F of two arguments such that the function U in (9) verifies

$$U(r, G, g) = F(E_0(r, G, g), r) \,, \tag{10}$$

where

$$E_0(r, G, g) = G^2 + r\sqrt{1 - G^2}\cos g \,. \tag{11}$$

By (10), the level curves of E_0 are also level curves of U. On the other hand, the phase portrait of E_0 in the plane (g, G)—i.e., the family of curves

$$E_0(r, G, g) = G^2 + r\sqrt{1 - G^2}\cos g = \mathcal{E}, \tag{12}$$

[2] We can do this as the Hamiltonians H_J and H_0 rescale by a factor β^{-2} as $(\mathbf{y}', \mathbf{y}) \to \beta^{-1}(\mathbf{y}', \mathbf{y})$ and $(\mathbf{x}', \mathbf{x}) \to \beta^2(\mathbf{x}', \mathbf{x})$.

in the plane (g, G) accordingly to the different values of r—is completely explicit [17]. For $0 < r < 1$ or $1 < r < 2$ it includes two minima $(\pm\pi, 0)$ on the g-axis; two symmetric maxima on the G-axis and one saddle point at $(0, 0)$. When $r > 2$ the saddle point disappears and $(0, 0)$ turns to be a maximum. The phase portrait includes two separatrices in the case $0 < r < 2$; one separatrix in the case $r > 2$. These are the level set $\mathcal{S}_0(r)$ through the saddle, corresponding to $\mathcal{E} = r$, for $0 < r < 2$, and the level set $\mathcal{S}_1(r) = \{\mathcal{E} = 1\}$, for any r. Rotational motions in between $\mathcal{S}_0(r)$ and $\mathcal{S}_1(r)$, do exist only for $0 < r < 1$. The minima and the maxima are surrounded by librational motions and different motions (librations about different equilibria or rotations) are separated by $\mathcal{S}_0(r)$ and $\mathcal{S}_1(r)$. The reader is referred to Fig. 1 for further qualitative details about the portion of the phase space corresponding to $[-\pi, \pi] \times [-1, 1]$.

We call *perihelion librations* the librational motions about $(\pm\pi, 0)$ or $(0, 0)$. Their physical meaning is that the perihelion of $\mathbb{E}$ affords oscillations while $\mathbb{E}$, highly eccentric anytime, periodically flattens to a segment in correspondence of the times when G vanishes. After the flattening time, the sense of rotation on $\mathbb{E}$ is reversed (as G changes its sign). We remark that (see the next section for a discussion) the potential U is well defined along the level sets of E, with the exception of $\mathcal{S}_0(r)$, where U is singular. In particular, U remains regular for all $r > 2$.

Let $0 < \beta_* \leq \beta^*$ be defined via

$$\beta_* := \begin{cases} \dfrac{\beta\overline{\beta}}{\beta + \overline{\beta}} & \text{for } H_J \\ \overline{\beta} & \text{for } H_0 \end{cases} \qquad \beta^* := \begin{cases} \max\{\beta, \overline{\beta}\} & \text{for } H_J \\ \beta + \overline{\beta} & \text{for } H_0\,. \end{cases} \tag{13}$$

De-homogeneizating via (9), we see that, if $r > 2\beta^*$, then we fall in the third panel in Fig. 1 for any U_β's in (7) so that all of such potentials afford perihelion librations about $(0, 0)$ and $(\pm\pi, 0)$. The works [5, 18] deal precisely with this situation.

Before (and in order to) describing the purposes of such works, we informally discuss the rôle of the total angular momentum's length $C := \|\mathbf{x} \times \mathbf{y} + \mathbf{x}' \times \mathbf{y}'\|$. This quantity enters in (7) via kinetic term $\|\mathbf{y}'\|^2$, according to

$$\|\mathbf{y}'\|^2 = R^2 + \frac{(C - G)^2}{r^2}, \tag{14}$$

(as $|C - G|$ is the Euclidean length of $\mathbf{x}' \times \mathbf{y}'$, assuming that $\mathbf{x} \times \mathbf{y}$ and $\mathbf{x}' \times \mathbf{y}'$ are parallel). Combining (14) with an expansion

$$U_\beta(r, G, g) = -\frac{1}{r} + \frac{1}{r} \sum_{k \geq 1} u_\kappa(G, g) \left(\frac{\beta}{r}\right)^k,$$

of the U_β's in (7) in powers of r^{-1}, one can split the Hamiltonians $\widehat{H}_J$ and $\widehat{H}_0$ in (7) in two parts, which we call, respectively, *fast* and *slow*:

$$H_{\text{fast}} = \frac{R^2}{2} + \frac{C^2}{2r^2} - \frac{1}{r}, \quad H_{\text{slow}} := \widetilde{U}_{\beta,\overline{\beta}}(r, G, g) + \frac{-2CG + G^2}{2r^2}, \tag{15}$$

where $\widetilde{U}_{\beta,\overline{\beta}}$ collects terms of order $\frac{\beta^*}{r^2}$, or higher, hence, retains the symmetries and the equilibria of the U_β's discussed above. We fix, in phase space, a region of initial data where the terms in (15) verify (see [5] for an informal discussion)

$$\|H_{\text{fast}}\| \gg \|\widetilde{U}_{\beta,\overline{\beta}}\| \gg \|H_{\text{slow}} - \widetilde{U}_{\beta,\overline{\beta}}\| . \tag{16}$$

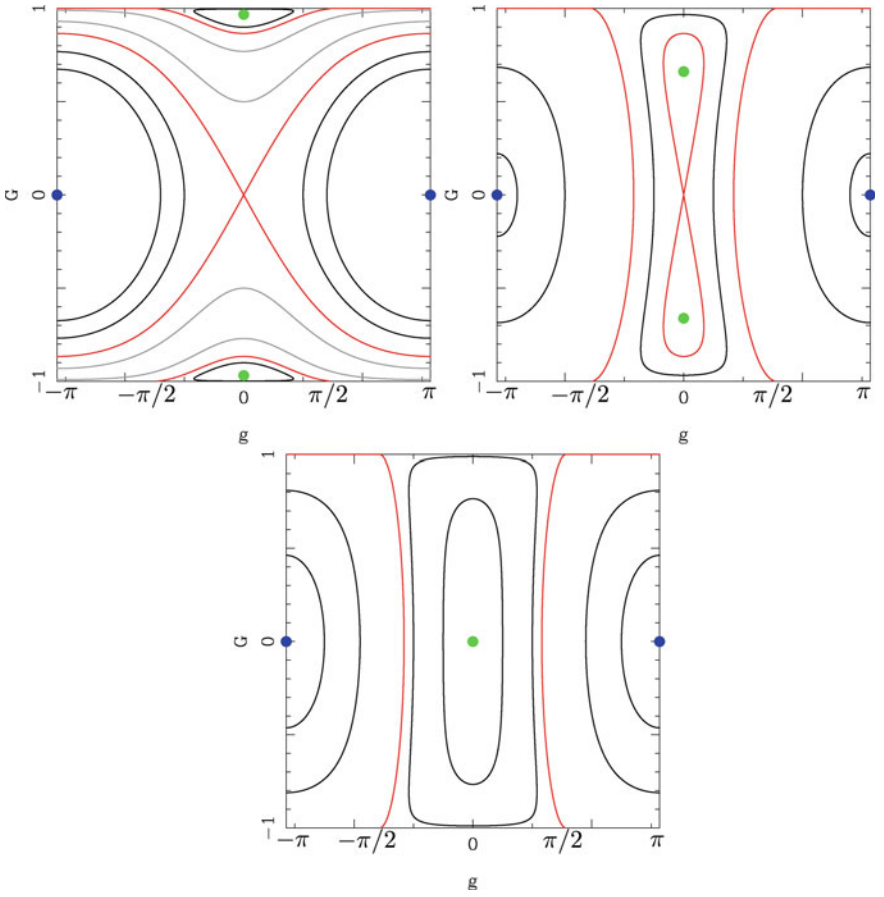


Fig. 1 Phase portraits of E_0 given by (11) in the plane (g,G) for $0 < r < 1$ (top left), $1 < r < 2$ (top right) and $r > 2$ (bottom). The points corresponding to minima of E_0 are labeled in blue, maxima appear in green and separatrices correspond to red curves

Then, the motions of R and r, mainly ruled H_{fast}, are faster than the ones of G and g, ruled by H_{slow}. If $C = 0$, the smallest term $H_{slow} - \widetilde{U}_{\beta,\overline{\beta}} = \frac{G^2}{2r^2}$ is even with respect to G, so H_{slow} retains the symmetries and equilibria of $\widetilde{U}_{\beta,\overline{\beta}}$ in Fig. 1, for the case $r > 2$. However, in this case, H_{fast} is unbounded below, so nothing prevents r to decrease below $2\beta^*$ and the scenario rapidly changes from (c) to (b) or (a). In this case, one has then to prove that perihelion librations occur in the full Hamiltonians (7) in such a short time that it prevents the scenario to change. The following result was obtained:

Theorem 1 ([18]) *Take, in* (7), $C = 0$. *Fix an arbitrary neighbourhood* U_0 *of* $(0, 0)$ *or of* $(0, \pi)$ *and an arbitrary neighbourhood* V_0 *of an unperturbed curve* $\gamma_0(t) = (G_0(t), g_0(t)) \in U_0$ *in Fig. 1. Then it is possible to find six numbers* $0 < c < 1$, $0 < \beta_- < \beta_+$, $0 < \alpha_- < \alpha_+$, $T > 0$, *such that, for any* $\beta_- < \beta_* \le \beta^* < \beta_+$ *the projections* $\Gamma_0(t) = (G(t), g(t))$ *of all the orbits* $\Gamma(t) = (R(t), G(t), r(t), g(t))$ *of* $\overline{H}_1$, $\overline{H}_2$ *with initial datum* $(R_0, r_0, G_0, g_0) \in [\frac{1}{\sqrt{c\alpha_+}}, \frac{1}{\sqrt{c\alpha_-}}] \times [c\alpha_-, \alpha_+] \times U_0$ *belong to* V_0 *for all* $0 \le t \le T$. *Moreover, the angle* $\gamma(t)$ *between the position ray of* $\Gamma_0(t)$ *and the* g*-axis affords a variation larger than* 2π *during the time* T.

The proof of Theorem 1 uses a new normal form theorem, together with the construction of a system of coordinates well adapted to perihelion librations, as reviewed in Sect. 3.

If $C \neq 0$, H_{fast} is bounded below, attaining its minimum at

$$R_0 = 0\,, \quad r_0 = C^2\,. \tag{17}$$

It is reasonable to expect that if the initial values of R and r are close to (17), they will remain there for some time and the motions of G and g will be close to be ruled by $H^0_{slow} := H_{slow}|_{r=r_0}$, which reads, to the lowest orders,

$$H^0_{slow}(G, g) = \frac{-2CG + G^2}{2r_0^2} - \beta^2 \frac{(5 - 3G^2)}{8r_0^3} - \beta^2 \frac{15(1 - G^2)}{8r_0^3} \cos 2g + O(r_0^{-5})\,. \tag{18}$$

On the other hand, the equilibria of $\widetilde{U}^0_{\beta,\overline{\beta}}$ have some chance of surviving in H^0_{slow} for small values of $|C|$, but the symmetries of $\widetilde{U}^0_{\beta,\overline{\beta}}$ do not persist. As an example, in Fig. 2, we report the phase portrait of H^0_{slow} for $\widehat{H}_J$, with

$$C = 25\,, \quad \beta = \overline{\beta} = 80\,. \tag{19}$$

We call *unperturbed motions* the motions obtained combining (17) with the motions in Fig. 2. The natural question now is whether and at which extent the motions of (7) may be regarded as perturbations of such unperturbed ones. The question was considered in [5], from the numerical point of view. Namely, in [5] the full motions

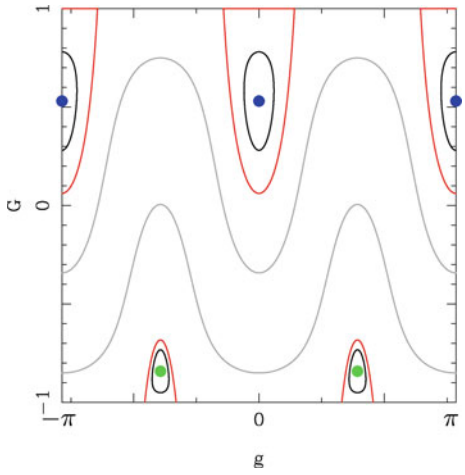


Fig. 2 Phase portrait of H^0_{slow} given in (18) with $C = 25$, $\beta = \overline{\beta} = 80$

of $\widehat{H}_J$ were analysed, with C, β and $\overline{\beta}$ chosen[3] as in (19), and the initial values of R and r close to (17). Numerical evidence of orbits continuing the unperturbed orbits above, interposed with zones of chaos, was obtained.

The paper is organised as follows:

1. In Sect. 2, we review recent results on the two-centre problem. We discuss first the existence of an invariant, referred as *Euler integral*, whose expression in the asymmetric setting is also given. Taking advantages of canonical coordinates lowering the number of degrees-of-freedom, coupled together with the *renormalizable integrability* property, the level sets of the averaged Newtonian potential are discussed in the planar case.
2. In Sect. 3, we outline the proof of Theorem 1 following [18]. The proof relies on normal form of Hamiltonians (7) free of small divisors, combined with an expression of the Euler integral suited for large values of r.
3. In Sect. 4, we further complement the understanding of the dynamics in the regime of large r: we retrace the steps of [5] in constructing explicitly an horseshoe orbit, therefore introducing the existence of symbolic dynamics. The methodology relies essentially on the construction of "boxes" stretching across one another under the action of a specific Poincaré mapping, and uses arguments of covering relations as introduced in [20].

[3] The quantities β, $\overline{\beta}$, γ, $\mathbf{y}'$, $\mathbf{x}'$, $\mathbf{y}$, $\mathbf{x}$, R, r, G, g, C of the present paper are related to β, $\overline{\beta}$, σ, y', x', y, x, R, r, G, g, C in [5] via the relations (with "here" , "there" standing for "in the present paper" and "in [5]", respectively) $\beta_{\text{here}} = (1+\mu)\beta_{\text{there}}, \overline{\beta}_{\text{here}} = (1+\mu)\overline{\beta}_{\text{there}}, \sigma(1+\mu)^2 = \gamma, \mathbf{y}' = \frac{\Lambda}{1+\mu} y'$, $\mathbf{x}' = \frac{1+\mu}{\Lambda^2} x'$, $\mathbf{y} = \Lambda y$, $\mathbf{x} = \frac{x}{\Lambda^2}$, $\mathrm{R}_{\text{here}} = \frac{\mathrm{R}_{\text{there}}\Lambda}{(1+\mu)}$, $\mathrm{r}_{\text{here}} = \frac{\mathrm{r}_{\text{there}}(1+\mu)}{\Lambda^2}$, $\mathrm{G}_{\text{here}} = \frac{\mathrm{G}_{\text{there}}}{\Lambda}$, $\mathrm{g}_{\text{here}} = \mathrm{g}_{\text{there}}$, $\mathrm{C}_{\text{here}} = \frac{\mathrm{C}_{\text{there}}}{\Lambda}$ where Λ, μ were chosen, in [5], 3.099 and 1, respectively. Note also the misprint in the definition of σ in [5, (1.4)], as the power of μ at the denominator should be 3 instead of 2. This misprint is inessential, as the number σ plays no rôle in [5].

2 Euler Problem Revisited

In this section we review the classical integration of the two-centre problem and complement it with considerations that will be useful to us in the next. The 2CP is the system, in $\mathbb{R}^3$ (or $\mathbb{R}^2$), of one particle interacting with two fixed masses via Newton Law. If $\pm\mathbf{v}_0 \in \mathbb{R}^3$ are the position coordinates of the centers, $\mathrm{m}_\pm$ their masses; $\mathbf{v}$, with $\mathbf{v} \neq \pm\mathbf{v}_0$, the position coordinate of the moving particle; $\mathbf{u} = \dot{\mathbf{v}}$ its velocity, and 1 its mass, the Hamiltonian of the system (*Euler Hamiltonian*) is

$$\mathrm{J} = \frac{\|\mathbf{u}\|^2}{2} - \frac{\mathrm{m}_+}{\|\mathbf{v}+\mathbf{v}_0\|} - \frac{\mathrm{m}_-}{\|\mathbf{v}-\mathbf{v}_0\|}\,. \tag{20}$$

Euler showed [13] that J exhibits 2 independent first integrals, in involution. One of these first integrals is the projection

$$\Theta = \mathbf{M}\cdot\frac{\mathbf{v}_0}{\|\mathbf{v}_0\|} \tag{21}$$

of the angular momentum $\mathbf{M} = \mathbf{v}\times\mathbf{u}$ of the particle along the direction $\mathbf{v}_0$. It is not specifically due to the Newtonian potential, but, rather, to its invariance by rotations around the axis $\mathbf{v}_0$. For example, it persists if the Newtonian potential is replaced with a α-homogeneous one. The existence of the following constant of motion, which we shall refer to as *Euler integral*:

$$\mathrm{E} = \|\mathbf{v}\times\mathbf{u}\|^2 + (\mathbf{v}_0\cdot\mathbf{u})^2 + 2\mathbf{v}\cdot\mathbf{v}_0\left(\frac{\mathrm{m}_+}{\|\mathbf{v}+\mathbf{v}_0\|} - \frac{\mathrm{m}_-}{\|\mathbf{v}-\mathbf{v}_0\|}\right) \tag{22}$$

is pretty specific of J. As observed in [3], in the limit of merging centers, i.e., $\mathbf{v}_0 = \mathbf{0}$, J reduces to the Kepler Hamiltonian (5), and E to the squared length of the angular momentum of the moving particle.

The formula in (22) is not easy[4] to be found in the literature, so we briefly discuss it.

After fixing a reference frame with the third axis in the direction of $\mathbf{v}_0$ and denoting as (v_1, v_2, v_3) the coordinates of $\mathbf{v}$ with respect to such frame, one introduces the so-called "elliptic coordinates"

$$\lambda = \frac{1}{2}\left(\frac{\mathrm{r}_+}{\mathrm{r}_0} + \frac{\mathrm{r}_-}{\mathrm{r}_0}\right)\,,\quad \beta = \frac{1}{2}\left(\frac{\mathrm{r}_+}{\mathrm{r}_0} - \frac{\mathrm{r}_-}{\mathrm{r}_0}\right)\,,\quad \omega := \arg\left(-v_2, v_1\right)\,, \tag{23}$$

where we have let, for short,

$$\mathrm{r}_0 := \|\mathbf{v}_0\|\,,\quad \mathrm{r}_\pm := \|\mathbf{v}\pm\mathbf{v}_0\|\,.$$

[4] See however [6] for a formula related to (22).

Regarding r_0 as a fixed external parameter and calling p_λ, p_β, p_ω the generalized momenta associated to λ, β and ω, it turns out that the Hamiltonian (20), written in the coordinates $(p_\lambda, p_\beta, \lambda, \beta)$ is independent of ω and has the expression

$$\mathrm{J}(p_\lambda, p_\beta, p_\omega, \lambda, \beta, \mathrm{r}_0) = \frac{1}{\lambda^2-\beta^2}\Big[\frac{p_\lambda^2(\lambda^2-1)}{2\mathrm{r}_0^2} + \frac{p_\beta^2(1-\beta^2)}{2\mathrm{r}_0^2} + \frac{p_\omega^2}{2\mathrm{r}_0^2}\left(\frac{1}{1-\beta^2} + \frac{1}{\lambda^2-1}\right) - \frac{(\mathrm{m}_+ + \mathrm{m}_-)\lambda}{\mathrm{r}_0^2} + \frac{(\mathrm{m}_+ - \mathrm{m}_-)\beta}{\mathrm{r}_0^2}\Big]. \tag{24}$$

It follows that the solution W of Hamilton–Jacobi equation

$$\mathrm{J}(W_\lambda, W_\beta, p_\omega, \lambda, \beta, \mathrm{r}_0) = h \tag{25}$$

can be searched of the form

$$W(\lambda, \beta, p_\omega, \mathrm{r}_0, h) = W^{(1)}(\lambda, p_\omega, \mathrm{r}_0, h) + W^{(2)}(\beta, p_\omega, \mathrm{r}_0, h)$$

and (25) separates completely as

$$\mathcal{F}^{(1)}(W_\lambda^{(1)}, \lambda, p_\omega, \mathrm{r}_0, h) + \mathcal{F}^{(2)}(W_\beta^{(2)}, \beta, p_\omega, \mathrm{r}_0, h) = 0 \tag{26}$$

with $\mathcal{F}^{(1)}$, $\mathcal{F}^{(2)}$ defined via (24)–(25).

The identity (26) implies that there must exist a function E, which we call *Euler integral*, depending on $(p_\omega, \mathrm{r}_0, h)$ only, such that

$$\mathcal{F}^{(1)}(p_\lambda, \lambda, p_\omega, \mathrm{r}_0, h) = -\mathcal{F}^{(2)}(p_\beta, \beta, p_\omega, \mathrm{r}_0, h) = \mathrm{E}(p_\omega, \mathrm{r}_0, h) \quad \forall\, (p_\lambda, p_\beta, \lambda, \beta)\,.$$

After elementary computations, one find that, in terms of the initial position–impulse coordinates, the Euler Integral

$$\mathrm{E} = \frac{1}{2}\left(\mathcal{F}^{(1)} - \mathcal{F}^{(2)}\right) \tag{27}$$

has the expression in (22), when written in the original coordinates.

The "asymmetric" case We are interested to find the expression of the Euler integral (22) when 2CP is written in the form

$$\mathrm{J} = \frac{\|\mathbf{y}\|^2}{2} - \frac{1}{\|\mathbf{x}\|} - \frac{\mathcal{M}'}{\|\mathbf{x}' - \mathbf{x}\|} \tag{28}$$

namely, when the two centres are in "asymmetric positions", $\mathbf{0}$, $\mathbf{x}'$. As we shall see, in that case we have

$$\mathrm{E} = \|\mathbf{M}\|^2 - \mathbf{x}' \cdot \mathbf{L} + \mathcal{M}' \frac{(\mathbf{x}' - \mathbf{x}) \cdot \mathbf{x}'}{\|\mathbf{x}' - \mathbf{x}\|} \tag{29}$$

where

$$\mathbf{M} := \mathbf{x} \times \mathbf{y}\,, \qquad \mathbf{L} := \mathbf{y} \times \mathbf{M} - \frac{\mathbf{x}}{\|\mathbf{x}\|} = \mathrm{e}\mathbf{P} \tag{30}$$

are the *angular momentum* and the *eccentricity vector* associated to the Kepler Hamiltonian (5) with e and $\mathbf{P}$ being the eccentricity and the perihelion direction ($\|\mathbf{P}\| = 1$). Notice that J reduces to a Kepler Hamiltonian in two cases: either for $\mathbf{x}' = \mathbf{0}$, in which case, as in the symmetric case above, E reduces to $\|\mathbf{M}\|^2$, or for $\mathcal{M}' = 0$. The latter case is more interesting to us, as J and E become, respectively, J_0 in (5) and

$$\mathrm{E}_0 = \|\mathbf{M}\|^2 - \mathbf{x}' \cdot \mathbf{L} \tag{31}$$

with E_0 being—as well expected—a combination of first integrals of J_0.

To prove (29)–(30), we change, canonically,

$$\mathbf{x}' = 2\mathbf{v}_0\,, \quad \mathbf{x} = \mathbf{v}_0 + \mathbf{v}\,, \quad \mathbf{y}' = \frac{1}{2}(\mathbf{u}_0 - \mathbf{u})\,, \quad \mathbf{y} = \mathbf{u}$$

(where $\mathbf{y}'$, $\mathbf{u}_0$ denote the generalized impulses conjugated to $\mathbf{x}'$, $\mathbf{v}_0$, respectively) we reach the Hamiltonian J in (20), with $\mathrm{m}_+ = 1$, $\mathrm{m}_- = \mathcal{M}'$. Turning back with the transformations, one sees that the function E in (22) takes the expression

$$\mathrm{E} := \left\| \left(\mathbf{x} - \frac{\mathbf{x}'}{2} \right) \times \mathbf{y} \right\|^2 + \frac{1}{4}(\mathbf{x}' \cdot \mathbf{y})^2 + \mathbf{x}' \cdot \left(\mathbf{x} - \frac{\mathbf{x}'}{2} \right) \left(\frac{1}{\|\mathbf{x}\|} - \frac{\mathcal{M}'}{\|\mathbf{x}' - \mathbf{x}\|} \right).$$

and we rewrite it as

$$\mathrm{E} = \mathrm{E}_0 + \mathrm{E}_1 + \mathrm{E}_2$$

with

$$\mathrm{E}_0 := \|\mathbf{M}\|^2 - \mathbf{x}' \cdot \mathbf{L}\,, \quad \mathrm{E}_1 := \mathcal{M}' \frac{(\mathbf{x}' - \mathbf{x}) \cdot \mathbf{x}'}{\|\mathbf{x}' - \mathbf{x}\|}$$
$$\mathrm{E}_2 := \frac{\|\mathbf{x}'\|^2}{2} \left(\frac{\|\mathbf{y}\|^2}{2} - \frac{1}{\|\mathbf{x}\|} - \frac{\mathcal{M}'}{\|\mathbf{x}' - \mathbf{x}\|} \right)$$

where $\mathbf{M}$, $\mathbf{L}$ are as in (30). Since E_2 is itself an integral for J, we can neglect it and rename

$$\mathrm{E} := \mathrm{E}_0 + \mathrm{E}_1 \tag{32}$$

the Euler integral to J. Namely,

$$\Big\{\mathrm{J}, \mathrm{E}\Big\} = 0\,. \tag{33}$$

A set of canonical coordinates which lets J *and* E *in 2 degrees-of-freedom* We describe a set of canonical coordinates, which we denote as $\mathcal{K}$, which we shall use for our analysis of the Euler Hamiltonian (28) and its integral E (32). This set of coordinates puts J and E in two degrees-of-freedom (represented by the couples (Λ, ℓ), (G, g) below), precisely like the classical ellipsoidal coordinates (23) do, both in the spatial and planar case.

We consider, in the region of $(\mathbf{y}, \mathbf{x})$ where J_0 in (5) takes negative values and the ellipse $\mathbb{E}(\mathbf{y}, \mathbf{x})$ it generates starting from any initial datum $(\mathbf{y}, \mathbf{x})$ in this region is not a circle. Denote as:

- a the *semi-major axis*;
- $\mathbf{P}$, with $\|\mathbf{P}\| = 1$, the direction of perihelion, assuming the ellipse is not a circle;
- ℓ: the mean anomaly, defined, mod 2π, as the area of the elliptic sector spanned by $\mathbf{x}$ from $\mathbf{P}$, normalized to 2π.

Finally,

- given three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$, with $\mathbf{u}, \mathbf{v} \perp \mathbf{w}$, we denote as $\alpha_{\mathbf{w}}(\mathbf{u}, \mathbf{v})$ the oriented angle from $\mathbf{u}$ to $\mathbf{v}$ relatively to the positive orientation established by $\mathbf{w}$.

We fix an arbitrary ("inertial") frame

$$\mathrm{F}_0 : \qquad \mathbf{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

in $\mathbb{R}^3$, and denote as

$$\mathbf{M} = \mathbf{x} \times \mathbf{y}\,, \quad \mathbf{M}' = \mathbf{x}' \times \mathbf{y}'\,, \quad \mathbf{C} = \mathbf{M}' + \mathbf{M}\,,$$

where "$\times$" denotes skew-product in $\mathbb{R}^3$. Observe the following relations

$$\mathbf{x}' \cdot \mathbf{C} = \mathbf{x}' \cdot \big(\mathbf{M} + \mathbf{M}'\big) = \mathbf{x}' \cdot \mathbf{M}\,, \qquad \mathbf{P} \cdot \mathbf{M} = 0\,, \quad \|\mathbf{P}\| = 1\,. \tag{34}$$

Assume that the "nodes"

$$\mathbf{n}_1 := \mathbf{k} \times \mathbf{C}\,, \quad \mathbf{n}_2 := \mathbf{C} \times \mathbf{x}'\,, \quad \mathbf{n}_3 := \mathbf{x}' \times \mathbf{M} \tag{35}$$

do not vanish. We define the coordinates

$$\mathcal{K} = (\mathrm{Z}, \mathrm{C}, \Theta, \mathrm{G}, \mathrm{R}, \Lambda, \zeta, g, \vartheta, \mathrm{g}, \mathrm{r}, \ell)$$

via the following formulae.

$$\begin{cases} \mathrm{Z} := \mathbf{C} \cdot \mathbf{k} \\ \mathrm{C} := \|\mathbf{C}\| \\ \mathrm{R} := \dfrac{\mathbf{y}' \cdot \mathbf{x}'}{\|\mathbf{x}'\|} \\ \Lambda = \sqrt{a} \\ \mathrm{G} := \|\mathbf{M}\| \\ \Theta := \dfrac{\mathbf{M} \cdot \mathbf{x}'}{\|\mathbf{x}'\|} \end{cases} \qquad \begin{cases} \mathrm{z} := \alpha_{\mathbf{k}}(\mathbf{i}, \mathbf{n}_1) \\ g := \alpha_{\mathbf{C}}(\mathbf{n}_1, \mathbf{n}_2) \\ \mathrm{r} := \|\mathbf{x}'\| \\ \ell := \text{meananomalyof}\mathbf{x} \text{ on } \mathbb{E} \\ \mathrm{g} := \alpha_{\mathbf{M}}(\mathbf{n}_3, \mathbf{M} \times \mathbf{P}) \\ \vartheta := \alpha_{\mathbf{x}'}(\mathbf{n}_2, \mathbf{n}_3) \end{cases} \tag{36}$$

The canonical character of $\mathcal{K}$ has been discussed in [17]. In the planar case, the coordinates (36) reduce to the 8 coordinates

$$\begin{cases} \mathrm{C} = \|\mathbf{x} \times \mathbf{y} + \mathbf{x}' \times \mathbf{y}'\| \\ \mathrm{G} = \|\mathbf{x} \times \mathbf{y}\| \\ \mathrm{R} = \dfrac{\mathbf{y}' \cdot \mathbf{x}'}{\|\mathbf{x}'\|} \\ \Lambda = \sqrt{a} \end{cases} \qquad \begin{cases} \gamma = \alpha_{\mathbf{k}}(\mathbf{i}, \mathbf{x}') + \dfrac{\pi}{2} \\ \mathrm{g} = \alpha_{\mathbf{k}}(\mathbf{x}', \mathbf{P}) + \pi \\ \mathrm{r} = \|\mathbf{x}'\| \\ \ell = \text{mean anomaly of } \mathbf{x} \text{ in } \mathbb{E} \end{cases} \tag{37}$$

Using the formulae in the previous section, we provide the expressions of J in (28) and E in (31) in terms of $\mathcal{K}$:

$$\begin{aligned} \mathrm{J}(\Lambda, \mathrm{G}, \Theta, \mathrm{r}, \ell, \mathrm{g}) &= -\frac{1}{2\Lambda^2} - \frac{\mathcal{M}'}{\sqrt{\mathrm{r}^2 + 2\mathrm{r}a\sqrt{1 - \frac{\Theta^2}{\mathrm{G}^2}}\mathrm{p} + a^2\varrho^2}} \\ &=: \mathrm{J}_0 + \mathrm{J}_1 \\ \mathrm{E}(\Lambda, \mathrm{G}, \Theta, \mathrm{r}, \ell, \mathrm{g}) &= \mathrm{G}^2 + \mathrm{r}\sqrt{1 - \frac{\Theta^2}{\mathrm{G}^2}}\sqrt{1 - \frac{\mathrm{G}^2}{\Lambda^2}}\cos\mathrm{g} \\ &\quad + \mathcal{M}'\mathrm{r}\frac{\mathrm{r} + a\sqrt{1 - \frac{\Theta^2}{\mathrm{G}^2}}\mathrm{p}}{\sqrt{\mathrm{r}^2 + 2\mathrm{r}a\sqrt{1 - \frac{\Theta^2}{\mathrm{G}^2}}\mathrm{p} + a^2\varrho^2}} \\ &=: \mathrm{E}_0 + \mathrm{E}_1 \end{aligned} \tag{38}$$

and, if $\xi = \xi(\Lambda, \mathrm{G}, \ell)$ is the *eccentric anomaly*, defined as the solution of *Kepler equation*

$$\xi - \mathrm{e}(\Lambda, \mathrm{G})\sin\xi = \ell \tag{39}$$

and $a = a(\Lambda)$ the *semi-major axis*; $\mathrm{e} = \mathrm{e}(\Lambda, \mathrm{G})$, the *eccentricity* of the ellipse, $\varrho = \varrho(\Lambda, \mathrm{G}, \ell)$, $\mathrm{p} = \mathrm{p}(\Lambda, \mathrm{G}, \ell, \mathrm{g})$ are defined as

$$
\begin{aligned}
a(\Lambda) &= \Lambda^2 \\
\mathrm{e}(\Lambda, \mathrm{G}) &:= \sqrt{1 - \frac{\mathrm{G}^2}{\Lambda^2}} \\
\varrho(\Lambda, \mathrm{G}, \ell) &:= 1 - \mathrm{e}(\Lambda, \mathrm{G}) \cos \xi(\Lambda, \mathrm{G}, \ell) \\
\mathrm{p}(\Lambda, \mathrm{G}, \ell, \mathrm{g}) &:= (\cos \xi(\Lambda, \mathrm{G}, \ell) - \mathrm{e}(\Lambda, \mathrm{G})) \cos \mathrm{g} - \frac{\mathrm{G}}{\Lambda} \sin \xi(\Lambda, \mathrm{G}, \ell) \sin \mathrm{g} \,.
\end{aligned} \tag{40}
$$

The angle

$$
\nu(\Lambda, \mathrm{G}, \ell) := \arg \left(\cos \xi(\Lambda, \mathrm{G}, \ell) - \mathrm{e}(\Lambda, \mathrm{G}), \frac{\mathrm{G}}{\Lambda} \sin \xi(\Lambda, \mathrm{G}, \ell) \right) \tag{41}
$$

is usually referred to as *true anomaly*, so one recognises that $\mathrm{p}(\Lambda, \mathrm{G}, \ell, \mathrm{g}) = \varrho \cos(\nu + \mathrm{g})$.

Observe that E and J in (38) do not depend on C, Z, ζ, γ, R, ϑ, while the Hamiltonians (7), do not depend on Z, ζ, γ, ℓ.

The details on the derivation of the formulae in (38) may be found in [17].

Renormalizable integrability In this section we review the property of *renormalizable integrability* pointed out in [16].

We consider the function U_β in (8) with $\beta = 1$, which is given by

$$
\mathrm{U}(\mathrm{r}, \Lambda, \Theta, \mathrm{G}, \mathrm{g}) = \frac{1}{2\pi} \int_0^{2\pi} \frac{d\ell}{\sqrt{\mathrm{r}^2 + 2ra\sqrt{1 - \frac{\Theta^2}{\mathrm{G}^2}}\mathrm{p} + a^2\varrho^2}} \tag{42}
$$

and the function

$$
\mathrm{E}_0 = \mathrm{G}^2 + \mathrm{r}\sqrt{1 - \frac{\Theta^2}{\mathrm{G}^2}}\sqrt{1 - \frac{\mathrm{G}^2}{\Lambda^2}} \cos \mathrm{g}
$$

in (38). These two functions have the following remarkable properties:

($\mathcal{P}_1$) they have one effective degree-of-freedom, as they depend on one conjugated couple of coordinates: the couple (G, g);

($\mathcal{P}_2$) they Poisson-commute:

$$
\left\{ \mathrm{U}, \ \mathrm{E}_0 \right\} = 0 \,. \tag{43}
$$

Relation (43) can be proved taking the ℓ-average of (33), and exploiting that J_0 depends only on Λ; see [16]. The following definition relies precisely with this situation.

Definition 1 ([16]) Let h, g be two functions of the form

$$
h(p, q, y, x) = \widehat{h}(\mathrm{I}(p, q), y, x) \,, \qquad g(p, q, y, x) = \widehat{g}(\mathrm{I}(p, q), y, x) \tag{44}
$$

where

$$(p, q, y, x) \in \mathcal{D} := \mathcal{B} \times U \tag{45}$$

with $U \subset \mathbb{R}^2$, $\mathcal{B} \subset \mathbb{R}^{2n}$ open and connected, $(p, q) = (p_1, \cdots, p_n, q_1, \cdots, q_n)$ conjugate coordinates with respect to the two-form $\omega = dy \wedge dx + \sum_{i=1}^{n} dp_i \wedge dq_i$ and $\mathrm{I}(p, q) = (\mathrm{I}_1(p, q), \cdots, \mathrm{I}_n(p, q))$, with

$$\mathrm{I}_i : \mathcal{B} \to \mathbb{R}, \qquad i = 1, \cdots, n$$

pairwise Poisson commuting:

$$\left\{\mathrm{I}_i, \mathrm{I}_j\right\} = 0 \qquad \forall\, 1 \leq i < j \leq n \qquad i = 1, \cdots, n\,. \tag{46}$$

We say that h is *renormalizably integrable via* g if there exists a function

$$\widetilde{h} : \qquad \mathrm{I}(\mathcal{B}) \times g(U) \to \mathbb{R}\,,$$

such that

$$h(p, q, y, x) = \widetilde{h}(\mathrm{I}(p, q), \widehat{g}(\mathrm{I}(p, q), y, x)) \tag{47}$$

for all $(p, q, y, x) \in \mathcal{D}$.

Proposition 1 ([16]) *If h is renormalizably integrable via g, then:*

(i) $\mathrm{I}_1, \cdots, \mathrm{I}_n$ *are first integrals to h and g;*
(ii) *h and g Poisson commute.*

Proposition 2 ([16]) U *is renormalizably integrable via* E_0. *Namely, there exists a function* F *such that*

$$\mathrm{U}(\mathrm{r}, \Lambda, \Theta, \mathrm{G}, \mathrm{g}) = \mathrm{F}\big(\mathrm{r}, \Lambda, \Theta, \mathrm{E}_0(\mathrm{r}, \Lambda, \Theta, \mathrm{G}, \mathrm{g})\big)\,.$$

The proof of Proposition 2 is based on $\mathcal{P}_1 \div \mathcal{P}_2$ above. Below, we list some consequences.

(i) If $F_{E_0} \neq 0$, the time laws of (G, g) under U or E_0 are basically (i.e., up to a change of time) the same;
(ii) Motions of E_0 corresponding to level sets for which $F_{E_0} = 0$ are fixed points curves to U ("frozen orbits"). In [16] we provided an example of frozen orbit of U in the spatial case, for $r \ll 1$;
(iii) U and E_0 have the same action–angle coordinates;
(iv) F may have several expressions, as well as U, which is defined via a quadrature. Two different representation formulae have been proposed in [16, 18].

In the next section, we investigate the dynamical properties of E_0 for the planar case ($\Theta = 0$).

The phase portrait of E_0 *in the planar case* Here we fix $\Lambda = 1, \Theta = 0$. For $r \in (0, 2)$, the function $E_0(g, G)$ has a minimum, a saddle and a maximum, respectively at

$$\mathbf{P}_- = (\pm\pi, 0)\,, \qquad \mathbf{P}_0 = (0, 0)\,, \qquad \mathbf{P}_+ = \left(0, \sqrt{1 - \frac{r^2}{4}}\right)$$

where it takes the values, respectively,

$$\mathcal{E}_- = -r\,, \qquad \mathcal{E}_0 = r\,, \quad \mathcal{E}_+ = 1 + \frac{r^2}{4}\,.$$

Thus, the level sets in (12) are non-empty only for

$$\mathcal{E} \in \left[-r, 1 + \frac{r^2}{4}\right]. \tag{48}$$

We denote as $\mathcal{S}_0$, the level set through the saddle $\mathbf{P}_0$. When $G = 1$, E_0 takes the value 1 for all g and we denote as $\mathcal{S}_1$ the level curve with $\mathcal{E} = 1$. The equations of $\mathcal{S}_0$, $\mathcal{S}_1$ are, respectively:

$$\begin{aligned} \mathcal{S}_0(r) &= \left\{(g, G) :\ G^2 + r\sqrt{1 - G^2}\cos g = r\right\}, \\ \mathcal{S}_1(r) &= \left\{G = \pm 1\right\} \cup \left\{G = \pm\sqrt{1 - r^2\cos^2 g}\right\}. \end{aligned} \tag{49}$$

$\mathcal{S}_1$ is composed of two branches, which will be referred to as "horizontal", "vertical", respectively, transversally intersecting at $(\pm\frac{\pi}{2}, 1)$, with g mod 2π. Note that, when $0 < r < 1$, the vertical branch is defined for all $g \in \mathbb{T}$; when $r > 1$, its domain in g is made of two disjoint neighbourhoods of $\pm\frac{\pi}{2}$.

When $r > 2$, the saddle $\mathbf{P}_0$ and its manifold $\mathcal{S}_0$ do not exist, $\mathbf{P}_- = (\pi, 0)$ is still a minimum, while the maximum becomes $\mathbf{P}_+ = (0, 0)$. The manifold $\mathcal{S}_1$ still exists, with the vertical branch closer and closer, as $r \to +\infty$, to the portion of straight $g = \pm\frac{\pi}{2}$ in the strip $-1 \leq G \leq 1$. In this case the admissible values for $\mathcal{E}$ are

$$\mathcal{E} \in [-\mathrm{r}, \mathrm{r}] \,.$$

It is worth mentioning [17] that, when $0 < \mathrm{r} < 2$, the motions generated by E_0 along $\mathcal{S}_0(\mathrm{r})$ can be explicitly computed, and are given by

$$\begin{cases} \mathrm{G}(t) = \frac{\sigma\Lambda}{\cosh \sigma\Lambda(t-t_0)}, \\ \mathrm{g}(t) = \pm\cos^{-1} \dfrac{1 - \frac{\alpha^2}{\cosh^2 \sigma\Lambda(t-t_0)}}{\sqrt{1 - \frac{\sigma^2}{\cosh^2 \sigma\Lambda(t-t_0)}}}, \end{cases}$$

where

$$\sigma^2 := \mathrm{r}(2-\mathrm{r}) \,, \quad \alpha^2 := 2 - \mathrm{r} \,, \quad \mathrm{r} \in (0,2) \,. \tag{50}$$

These motions—which have a remarkable similitude with the separatrix motions of the classical pendulum—are however meaningless for U, which is singular on $\mathcal{S}_0(\mathrm{r})$.

The scenario is depicted in Fig. 1 to which we refer for further qualitative details.

3 Perihelion Librations in the Three-Body Problem

In this section we review the results of [16–18].

As mentioned in the introduction, the proof of Theorem 1 is based on two ingredients: a normal form theory well designed around the Hamiltonians (7), where no non-resonance condition is required, and a set of action–angle-like coordinates which approximate well the natural action–angle coordinates of E_0 when r is large. In this section we briefly summarise the procedure. Full details may be found in [18].

A normal form theory without small divisors We describe a procedure for eliminating the angles[5] φ at high orders, given Hamiltonian of the form

$$\mathrm{H}(\mathbf{I}, \varphi, \mathbf{p}, \mathbf{q}, y, x) = \mathrm{h}(\mathbf{I}, \mathbf{J}(\mathbf{p}, \mathbf{q}), y) + f(\mathbf{I}, \varphi, \mathbf{p}, \mathbf{q}, y, x) \tag{51}$$

which we assume to be holomorphic on the neighbourhood

$$\mathbb{P}_{\rho,s,\delta,r,\xi} = \mathbb{I}_\rho \times \mathbb{T}^n_s \times \mathbb{B}_\delta \times \mathbb{Y}_r \times \mathbb{X}_\xi \supset \mathbb{P} = \mathbb{I} \times \mathbb{T}^n \times \mathbb{B} \times \mathbb{Y} \times \mathbb{X} \,,$$

for suitable $\rho, s, \delta, r, \xi > 0$ and

$$\mathbf{J}(\mathbf{p}, \mathbf{q}) = (p_1 q_1, \cdots, p_m q_m) \,.$$

[5] Note that the procedure described in this section does not seem to be related to [2, Sect. 6.4.4], for the lack of slow–fast couples.

Here, $\mathbb{I} \subset \mathbb{R}^n$, $\mathbb{B} \subset \mathbb{R}^{2m}$, $\mathbb{Y} \subset \mathbb{R}$, $\mathbb{X} \subset \mathbb{R}$ are open and connected; $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ is the standard torus, and we have used the common notation $A_r := \bigcup_{x \in A} B_r(x)$, where $B_r(x)$ is the complex open ball centered in x with radius r.

We denote as $\mathcal{O}_{\rho,s,\delta,r,\xi}$ the set of complex holomorphic functions

$$\phi : \quad \mathbb{P}_{\hat{\rho},\hat{s},\hat{\delta},\hat{r},\hat{\xi}} \to \mathbb{C}$$

for some $\hat{\rho} > \rho$, $\hat{s} > s$, $\hat{\delta} > \delta$, $\hat{r} > r$, $\hat{\xi} > \xi$, equipped with the norm

$$\|\phi\|_{\rho,s,\delta,r,\xi} := \sum_{k,h,j} \|\phi_{khj}\|_{\rho,r,\xi} e^{s|k|} \delta^{h+j}$$

where $\phi_{khj}(\mathbf{I}, y, x)$ are the coefficients of the Taylor–Fourier expansion[6]

$$\phi = \sum_{k,h,j} \phi_{khj}(\mathbf{I}, y, x) e^{\mathrm{i}ks} \mathbf{p}^h \mathbf{q}^j \,, \quad \|\phi\|_{\rho,r,\xi} := \sup_{\mathbb{I}_\rho \times \mathbb{Y}_r \times \mathbb{X}_\xi} |\phi(\mathbf{I}, y, x)| \,.$$

If ϕ is independent of x, we simply write $\|\phi\|_{\rho,r}$ for $\|\phi\|_{\rho,r,\xi}$. If $\phi \in \mathcal{O}_{\rho,s,\delta,r,\xi}$, we define its "off-average" $\widetilde{\phi}$ and "average" $\overline{\phi}$ as

$$\widetilde{\phi} := \sum_{\substack{k,h,j: \\ (k,h-j) \neq (0,0)}} \phi_{khj}(\mathbf{I}, y, x) e^{\mathrm{i}ks} \mathbf{p}^h \mathbf{q}^j$$

$$\overline{\phi} := \phi - \widetilde{\phi} = \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} \Pi_{\mathbf{pq}} \phi(\mathbf{I}, \boldsymbol{\varphi}, \mathbf{J}(\mathbf{p}, \mathbf{q}), y, x) d\boldsymbol{\varphi} \,,$$

with

$$\Pi_{\mathbf{pq}} \phi(\mathbf{I}, \boldsymbol{\varphi}, \mathbf{J}(\mathbf{p}, \mathbf{q}), y, x) := \sum_{k,h} \phi_{khh}(\mathbf{I}, y, x) e^{\mathrm{i}ks} \mathbf{p}^h \mathbf{q}^h$$

We decompose

$$\mathcal{O}_{\rho,s,\delta,r,\xi} = \mathcal{Z}_{\rho,s,\delta,r,\xi} \oplus \mathcal{N}_{\rho,s,\delta,r,\xi} \,.$$

where $\mathcal{Z}_{\rho,s,\delta,r,\xi}, \mathcal{N}_{\rho,s,\delta,r,\xi}$ are the "zero-average" and the "normal" classes

$$\mathcal{Z}_{\rho,s,\delta,r,\xi} := \{\phi \in \mathcal{O}_{\rho,s,\delta,r,\xi} : \quad \phi = \widetilde{\phi}\} = \{\phi \in \mathcal{O}_{\rho,s,\delta,r,\xi} : \quad \overline{\phi} = 0\} \tag{52}$$

$$\mathcal{N}_{\rho,s,\delta,r,\xi} := \{\phi \in \mathcal{O}_{\rho,s,\delta,r,\xi} : \quad \phi = \overline{\phi}\} = \{\phi \in \mathcal{O}_{\rho,s,\delta,r,\xi} : \quad \widetilde{\phi} = 0\} \,. \tag{53}$$

respectively. We finally let $\omega_{y,\mathbf{I},\mathbf{J}} := \partial_{y,\mathbf{I},\mathbf{J}} \mathrm{h}$.

[6] We denote as $\mathbf{x}^h := x_1^{h_1} \cdots x_n^{h_n}$, where $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$ and $h = (h_1, \cdots, h_n) \in \mathbb{N}^n$.

In the following result, no non-resonance condition is required on the frequencies $\omega_{\mathbf{I}}$, which, as a matter of fact, might also be zero.

Theorem 2 ([18]) *For any n, m, there exists a number $\mathrm{c}_{n,m} \geq 1$ such that, for any $N \in \mathbb{N}$ such that the following inequalities are satisfied*

$$\begin{aligned} &4N\mathcal{X}\left\|\operatorname{Im}\frac{\omega_{\mathbf{I}}}{\omega_y}\right\|_{\rho,r} < s\,, \quad 4N\mathcal{X}\left\|\frac{\omega_{\mathbf{J}}}{\omega_y}\right\|_{\rho,r} < 1 \\ &\widetilde{\mathrm{c}}_{n,m} N\frac{\mathcal{X}}{\mathrm{d}}\,\|f\|_{\rho,s,\delta,r,\xi}\left\|\frac{1}{\omega_y}\right\|_{\rho,s,\delta,r,\xi} < 1 \end{aligned} \tag{54}$$

with $\mathrm{d} := \min\left\{\rho s, r\xi, \delta^2\right\}$, $\mathcal{X} := \sup\left\{|x| :\ x \in \mathbb{X}_\xi\right\}$, *one can find an operator*

$$\Psi_* :\quad \mathcal{O}_{\rho,s,\delta,r,\xi} \to \mathcal{O}_{1/3(\rho,s,\delta,r,\xi)} \tag{55}$$

which carries H *to*

$$\mathrm{H}_* = \mathrm{h} + g_* + f_*$$

where $g_* \in \mathcal{N}_{1/3(\rho,s,\delta,r,\xi)}$, $f_* \in \mathcal{O}_{1/3(\rho,s,\delta,r,\xi)}$ and, moreover, the following inequalities hold

$$\begin{aligned} &\|g_* - \overline{f}\|_{1/3(\rho,s,\delta,r,\xi)} \leq 162\widetilde{\mathrm{c}}_{n,m}\frac{\mathcal{X}}{\mathrm{d}}\left\|\frac{\widetilde{f}}{\omega_y}\right\|_{\rho,s,\delta,r,\xi}\|f\|_{\rho,s,\delta,r,\xi} \\ &\|f_*\|_{1/3(\rho,s,\delta,r,\xi)} \leq \frac{1}{2^{N+1}}\|f\|_{\rho,s,\delta,r,\xi}\,. \end{aligned} \tag{56}$$

The transformation Ψ_* can be obtained as a composition of time-one Hamiltonian flows, and satisfies the following. If

$$(\mathbf{I}, \boldsymbol{\varphi}, \mathbf{p}, \mathbf{q}, y, x) := \Psi_*(\mathbf{I}_*, \boldsymbol{\varphi}_*, \mathbf{p}_*, \mathbf{q}_*, \mathrm{R}_*, \mathrm{r}_*)$$

the following uniform bounds hold:

$$\begin{aligned} &\mathrm{d}\max\left\{\frac{|\mathbf{I}-\mathbf{I}_*|}{\rho}, \frac{|\boldsymbol{\varphi}-\boldsymbol{\varphi}_*|}{s}, \frac{|\mathbf{p}-\mathbf{p}_*|}{\delta}, \frac{|\mathbf{q}-\mathbf{q}_*|}{\delta}, \frac{|y-y_*|}{r}, \frac{|x-x_*|}{\xi}\right\} \\ &\leq \max\left\{s|\mathbf{I}-\mathbf{I}_*|,\ \rho|\boldsymbol{\varphi}-\boldsymbol{\varphi}_*|,\ \delta|\mathbf{p}-\mathbf{p}_*|,\ \delta|\mathbf{q}-\mathbf{q}_*|,\ \xi|y-y_*|,\ r|x-x_*|\right\} \\ &\leq 19\,\mathcal{X}\left\|\frac{f}{\omega_y}\right\|_{\rho,s,\delta,r,\xi}. \end{aligned} \tag{57}$$

Hints on the proof of Theorem 2 may be found in Appendix A.

Asymptotic action–angle coordinates The explicit construction of the action–angle coordinates for E_0 for any value of r and Θ exhibits elliptic integrals. This is true even in the case $\Theta = 0$, in which the phase portrait is, as discussed, explicit. As we are interested to the case that r is large, we adopt the "approximate" solution of integrating only the leading part of E. Namely, we replace Eq. (12) with

$$\sqrt{1 - G^2} \cos g = \widetilde{\mathcal{E}} \,. \tag{58}$$

We show that, for this case, the action–angle coordinates, denoted as $(\mathcal{G}, \gamma)$, are given by

$$\mathcal{G} = \widetilde{\mathcal{E}} \,, \quad \gamma = \tau \tag{59}$$

where τ is the time the flows employs to reach the value (G, g) on the level set $\widetilde{\mathcal{E}}$, starting from $(\sqrt{1 - \widetilde{\mathcal{E}}^2}, 0)$ $((\sqrt{1 - \widetilde{\mathcal{E}}^2}, \pi))$. Namely, for the Hamiltonian (58), the action–angle coordinates coincide with the energy–time coordinates. Indeed, by (58), the action variable can be taken to be

$$\mathcal{G}(\widetilde{\mathcal{E}}) = \begin{cases} -1 + \dfrac{1}{\pi} \displaystyle\int_{-\arccos|\widetilde{\mathcal{E}}|}^{\arccos|\widetilde{\mathcal{E}}|} \sqrt{1 - \frac{\widetilde{\mathcal{E}}^2}{\cos^2 g}} dg & -1 < \widetilde{\mathcal{E}} < 0 \\ 1 - \dfrac{1}{\pi} \displaystyle\int_{-\arccos\widetilde{\mathcal{E}}}^{\arccos\widetilde{\mathcal{E}}} \sqrt{1 - \frac{\widetilde{\mathcal{E}}^2}{\cos^2 g}} dg & 0 < \widetilde{\mathcal{E}} < 1 \,. \end{cases}$$

We have defined $\mathcal{G}(\widetilde{\mathcal{E}})$ so that $\mathcal{G}(0) = 0$. Then the period of the orbit is given by

$$\mathcal{T}(\widetilde{\mathcal{E}}) = 2\pi \mathcal{G}_{\widetilde{\mathcal{E}}}(\widetilde{\mathcal{E}}) \,.$$

With the change of variable

$$w = \frac{|\widetilde{\mathcal{E}}|}{\sqrt{1 - \widetilde{\mathcal{E}}^2}} \tan g \,, \tag{60}$$

we obtain

$$\mathcal{T}(\widetilde{\mathcal{E}}) = 4|\widetilde{\mathcal{E}}| \int_0^{\arccos|\widetilde{\mathcal{E}}|} \frac{1}{\cos^2 g} \frac{dg}{\sqrt{1 - \frac{\widetilde{\mathcal{E}}^2}{\cos^2 g}}} = 4 \int_0^1 \frac{dw}{\sqrt{1 - w^2}} = 2\pi$$

which implies (59).

Looking at the (multi-valued) generating function

$$S(\mathcal{G}, \mathrm{g}) = \int_{P_0(\mathcal{G})}^{P_{\mathrm{g}}(\mathcal{G})} \sqrt{1 - \frac{\mathcal{G}^2}{\cos^2 \mathrm{g}'}} d\mathrm{g}'$$

(where, as it is standard to do [1], the integral is computed along the $\mathcal{G}$th level set, from $P_0(\mathcal{G}) := (-\arccos \mathcal{G}, 0)$ to a prefixed point $P_{\mathrm{g}}(\mathcal{G}) = (\mathrm{g}, \cdot)$ of the level set, so as to make $S(\mathcal{G}, \cdot)$ continuous) we obtain the transformation of coordinates

$$\begin{cases} \mathrm{G} = \sqrt{1 - \mathcal{G}^2} \cos \gamma \\ \mathrm{g} = -\tan^{-1} \left(\dfrac{1}{\mathcal{G}} \sqrt{1 - \mathcal{G}^2} \sin \gamma \right) + k\pi \\ \text{with} \quad k = \begin{cases} 0 \ \text{if} \ 0 < \mathcal{G} < 1 \\ 1 \ \text{if} \ -1 < \mathcal{G} < 0 \,. \end{cases} \end{cases} \tag{61}$$

Then, using the coordinates $(\mathcal{G}, \gamma)$, one obtains the expression

$$\mathrm{E}_0 = \mathcal{G}\,\mathrm{r} + (1 - \mathcal{G}^2) \cos^2 \gamma$$

which will be used in the next section.

The proof of Theorem 1 is a direct application of Proposition 2. As such, a careful evaluation of the involved quantities is needed. Those evaluations are completely explicit in [18]. For the purpose of this review, we report below the main ideas while we skip most of computational details. The reader who is interested in them might consult [18].

Sketch of proof of Theorem 1 For definiteness, we sketch the proof of Theorem 1 for $(0, 0)$. The proof for $(0, \pi)$ is similar. For the purposes of this proof, we let $\widehat{\mathrm{H}}_1 := \widehat{\mathrm{H}}_{\mathrm{J}}$ and $\widehat{\mathrm{H}}_2 := \widehat{\mathrm{H}}_0$, where $\widehat{\mathrm{H}}_{\mathrm{J}}$ and $\widehat{\mathrm{H}}_{\mathrm{J}}$ are as in (7). It is convenient to rewrite the functions $\widehat{\mathrm{H}}_i$ as

$$\begin{aligned} \widehat{\mathrm{H}}_1(\mathrm{R}, \mathrm{G}, \mathrm{r}, \mathrm{g}) = {} & \left(\frac{\mathrm{R}^2}{2} - \frac{1}{\mathrm{r}} \right) + \frac{\mathrm{G}^2}{2\mathrm{r}^2} - \frac{\overline{\beta}}{\beta + \overline{\beta}} \frac{1}{\mathrm{r}} \left(\widehat{\mathrm{F}}_{\beta\varepsilon(\mathrm{r})} \left(\widehat{\mathrm{E}}_{\beta\varepsilon(\mathrm{r})}(\mathrm{G}, \mathrm{g}) \right) - 1 \right) \\ & - \frac{\beta}{\beta + \overline{\beta}} \frac{1}{\mathrm{r}} \left(\widehat{\mathrm{F}}_{-\overline{\beta}\varepsilon(\mathrm{r})} \left(\widehat{\mathrm{E}}_{-\overline{\beta}\varepsilon(\mathrm{r})}(\mathrm{G}, \mathrm{g}) \right) - 1 \right) \\ \widehat{\mathrm{H}}_2(\mathrm{R}, \mathrm{G}, \mathrm{r}, \mathrm{g}) = {} & \left(\frac{\mathrm{R}^2}{2} - \frac{1}{\mathrm{r}} \right) + \frac{\mathrm{G}^2}{2\mathrm{r}^2} - \frac{\overline{\beta}}{\beta + \overline{\beta}} \frac{1}{\mathrm{r}} \left(\widehat{\mathrm{F}}_{(\beta+\overline{\beta})\varepsilon(\mathrm{r})} \left(\widehat{\mathrm{E}}_{(\beta+\overline{\beta})\varepsilon(\mathrm{r})}(\mathrm{G}, \mathrm{g}) \right) - 1 \right) \end{aligned}$$

where

$$\varepsilon(\mathrm{r}) := \frac{1}{\mathrm{r}}, \quad \widehat{\mathrm{E}}_\varepsilon(\mathrm{G}, \mathrm{g}) := \varepsilon \mathrm{E}_0(\varepsilon^{-1}, \mathrm{G}, \mathrm{g}), \quad \widehat{\mathrm{F}}_\varepsilon(t) := \varepsilon^{-1} \mathrm{F}(\varepsilon^{-1}, \varepsilon^{-1} t).$$

We next change coordinates via the canonical changes

$$\mathcal{C}_1 : \ (\mathcal{G}, \gamma) \to (\mathrm{G}, \mathrm{g}), \quad \mathcal{C}_2 : \ (y, x) \to (\mathrm{R}, \mathrm{r})$$

where $\mathcal{C}_1$ is defined as in (61), with $k = 0$, while $\mathcal{C}_2$ is

$$\begin{cases} \mathrm{R}(y,x) = \dfrac{1}{y}\sqrt{\dfrac{\cos\xi'(x)+1}{1-\cos\xi'(x)}} \\ \mathrm{r}(y,x) = y^2(1-\cos\xi'(x)) \end{cases} \tag{62}$$

where $\xi'(x)$ solves

$$\xi' - \sin\xi' = x\,. \tag{63}$$

$\mathcal{C}_2$ has been chosen so that

$$\left(\frac{\mathrm{R}^2}{2} - \frac{1}{\mathrm{r}}\right) \circ \mathcal{C}_2 = -\frac{1}{2y^2}\,.$$

Using the new coordinates, we have

$$\begin{aligned} \widehat{\mathrm{H}}_1 = &-\frac{1}{2y^2} + \frac{1}{\mathrm{r}(y,x)}\left(\varepsilon(y,x)\frac{(1-\mathcal{G}^2)}{2}\cos^2\gamma - \frac{\overline{\beta}}{\beta+\overline{\beta}}\left(\widehat{\mathrm{F}}_{\beta\varepsilon(y,x)}\left(\widehat{\mathrm{E}}_{\beta\varepsilon(y,x)}(\mathcal{G},\gamma)\right)-1\right)\right. \\ &\left. -\frac{\beta}{\beta+\overline{\beta}}\left(\widehat{\mathrm{F}}_{-\overline{\beta}\varepsilon(y,x)}\left(\widehat{\mathrm{E}}_{-\overline{\beta}\varepsilon(y,x)}(\mathcal{G},\gamma)\right)-1\right)\right) \\ \widehat{\mathrm{H}}_2 = &-\frac{1}{2y^2} + \frac{1}{\mathrm{r}(y,x)}\left(\varepsilon(y,x)\frac{(1-\mathcal{G}^2)}{2}\cos^2\gamma\right. \\ &\left. -\frac{\overline{\beta}}{\beta+\overline{\beta}}\frac{1}{\mathrm{r}(y,x)}\left(\widehat{\mathrm{F}}_{(\beta+\overline{\beta})\varepsilon(y,x)}\left(\widehat{\mathrm{E}}_{(\beta+\overline{\beta})\varepsilon(y,x)}(\mathcal{G},\gamma)\right)-1\right)\right) \end{aligned}$$

having abusively denoted as $\varepsilon(y,x) := \varepsilon\big(\mathrm{r}(y,x)\big)$ and

$$\widehat{\mathrm{E}}_\varepsilon(\mathcal{G},\gamma) := \mathcal{G} + \varepsilon\left(1-\mathcal{G}^2\right)\cos^2\gamma\,. \tag{64}$$

A domain where we shall check holomorphy for $\widehat{\mathrm{H}}_i$ is chosen as

$$\mathbb{D}_{\delta,s_0,\sqrt{\alpha_-},\sqrt{\varepsilon_0}} := \mathbb{Y}_{\sqrt{\alpha_-}} \times \mathbb{X}_{\sqrt{\varepsilon_0}} \times \mathbb{G}_\delta \times \mathbb{T}_{s_0} \tag{65}$$

where

$$\begin{aligned} &\mathbb{Y} := \left\{y \in \mathbb{R} :\ 2\sqrt{\alpha_-} < y < \sqrt{\alpha_+}\right\}, \quad \mathbb{X} := \left\{x \in \mathbb{R} :\ |x-\pi| \le \pi - 2\sqrt{\varepsilon_0}\right\} \\ &\mathbb{G} := \left\{\mathcal{G} \in \mathbb{R} :\ 1-\delta < \mathcal{G} < 1\right\} \end{aligned} \tag{66}$$

where $0 < \alpha_- < \frac{\alpha_+}{4}$, $0 < \varepsilon_0 < \frac{\pi^2}{4}$, $0 < \delta < 1$. If $c_0 > 0$ is such that for any $0 < \varepsilon_0 < 1$ and for any $x \in \mathbb{X}_{\sqrt{\varepsilon_0}}$, Eq. (63) has a unique solution $\xi'(x)$ which depends

analytically on x and and verifies

$$|1 - \cos \xi'(x)| \geq c_0 \varepsilon_0 \tag{67}$$

(the existence of such a number c_0 is well known) and if the following inequalities are satisfied

$$\begin{aligned} &0 < \delta \leq \frac{1}{4}, \qquad C^*(s_0)\delta < 1 \qquad C^*(s_0) := 16 \left(\sup_{\mathbb{T}_{s_0}} |\sin \gamma| \right)^2 \\ &\alpha_- \varepsilon_0 > \frac{4\beta^*}{c_0} \end{aligned} \tag{68}$$

with β^* as in (13), then $\widehat{\mathrm{H}}_i$ are holomorphic in the domain (65). The proof is based on the explicit evaluation of the function $\widehat{\mathrm{F}}_\varepsilon(t)$ for complex values of its arguments, the accomplishment of which is obtained using the explicit expression of $\widehat{\mathrm{F}}_\varepsilon(t)$: see [18, Proposition 3.1 and Proposition 3.3].

We aim to apply Theorem 2, with $\mathbf{I} = \mathcal{G}$, $\boldsymbol{\varphi} = \gamma$, (y, x) as in (62), $\mathrm{h}(y) = -\frac{1}{2y^2}$ and, finally

$$f(\mathcal{G}, \gamma, y, x) = \begin{cases} \frac{1}{\mathrm{r}(y,x)} \left(\varepsilon(y,x) \frac{(1-\mathcal{G}^2)}{2} \cos^2 \gamma - \frac{\overline{\beta}}{\beta+\overline{\beta}} \left(\widehat{\mathrm{F}}_{\beta\varepsilon(y,x)} \left(\widehat{\mathrm{E}}_{\beta\varepsilon(y,x)}(\mathcal{G}, \gamma) \right) - 1 \right) \right. \\ \left. - \frac{\beta}{\beta+\overline{\beta}} \left(\widehat{\mathrm{F}}_{-\overline{\beta}\varepsilon(y,x)} \left(\widehat{\mathrm{E}}_{-\overline{\beta}\varepsilon(y,x)}(\mathcal{G}, \gamma) \right) - 1 \right) \right) \quad i = 1 \\ \frac{1}{\mathrm{r}(y,x)} \left(\varepsilon(y,x) \frac{(1-\mathcal{G}^2)}{2} \cos^2 \gamma \right. \\ \left. - \frac{\overline{\beta}}{\beta+\overline{\beta}} \frac{1}{\mathrm{r}(y,x)} \left(\widehat{\mathrm{F}}_{(\beta+\overline{\beta})\varepsilon(y,x)} \left(\widehat{\mathrm{E}}_{(\beta+\overline{\beta})\varepsilon(y,x)}(\mathcal{G}, \gamma) \right) - 1 \right) \right) \\ \quad i = 2 \end{cases} \tag{69}$$

As h does not depend on $\mathcal{G}$ and the coordinates $\mathbf{p}$, $\mathbf{q}$ do not exist, in order to apply Theorem 2, only the last condition in (54) needs to be verified. Direct computations show that such condition is verified provided that $N = [N_0] - 1$, where

$$\frac{1}{N_0} := C_* \max \left\{ \frac{\beta_*}{c_0^2 \varepsilon_0^2 \delta s_0} \sqrt{\frac{1}{\alpha_-}}, \ \frac{\beta_*}{c_0^2 \varepsilon_0^{\frac{5}{2}}} \frac{1}{\alpha_-} \right\} \frac{\alpha_+^{3/2}}{\alpha_-^{3/2}}$$

with β_* as in (13) and C_* is independent of s_0. Assuming also that

$$N_0^{-1} < \frac{c_0^2 \varepsilon_0^2 \alpha_-^2}{2\alpha_+^2} \tag{70}$$

we have, in particular, $N_0 > 2$. We denote as

$$\widehat{\mathrm{H}}_* = \mathrm{h}(y_*) + g_*(y_*, x_*, \mathcal{G}_*) + f_*(\mathcal{G}_*, \gamma_*, y_*, x_*) \tag{71}$$

the Hamiltonian obtained after the application of Theorem 2 where, g_*, f_* satisfy the following bounds:

$$\|g_* - \overline{f}\| \leq 2\Delta\,, \quad \|g_*\| \leq 2^{-N}\Delta$$

with $\overline{f}(y_*, x_*, \mathcal{G}_*)$ the γ_*-average of $f(y_*, x_*, \mathcal{G}_*, \gamma_*)$ and

$$\Delta := C_* \frac{m_0^2 a \beta_*}{c_0^2 \varepsilon_0^2 \alpha_-^2}$$

is an upper bound to $\|f\|$ above. Let now $\Gamma_*(t) = (\mathcal{G}_*(t), \gamma_*(t), y_*(t), x_*(t))$ be a solution of $\widehat{\mathrm{H}}_*$ with initial datum $\Gamma_*(0) = (\mathcal{G}_*(0), \gamma_*(0), y_*(0), x_*(0)) \in \mathbb{D}$ and verifying

$$\begin{aligned} &|\mathcal{G}_*(0) - 1| \leq \frac{\delta}{2}\,, \quad 2\sqrt{m_0^3 \alpha_-} \leq |y_*(0)| \leq \frac{\sqrt{m_0^3 \alpha_-} + \sqrt{m_0^3 \alpha_+}}{2} \\ &x_*(0) = \pi\,. \end{aligned} \tag{72}$$

We look for a time $T > 0$ such that $\Gamma_*(t) \in \mathbb{D}$ for all $0 \leq t \leq T$. Then T can be taken to be

$$T = \min\left\{ \sqrt{\alpha_-^3}\,,\ \frac{\sqrt{\alpha_- \varepsilon_0}}{\Delta}\,,\ 2^{N_0} \frac{s_0 \delta}{\Delta} \right\} \tag{73}$$

as this choice easily allows to check

$$\begin{aligned} &|y_*(t) - y_*(0)| \leq |y_*(0)| - \sqrt{\alpha_-}\,, \quad |\mathcal{G}_*(t) - \mathcal{G}_*(0)| \leq \frac{2^{-(N+1)}\Delta t}{s_0} \leq \frac{\delta}{2} \\ &|x_*(t) - x_*(0)| \leq \pi - \sqrt{\varepsilon_0} \end{aligned}$$

for all $|t| \leq T$. In addition, at the time $t = T$, one has

$$\begin{aligned} |\gamma_*(T) - \gamma_*(0)| \ &\geq\ c^\circ \min\left\{ \beta_* \sqrt{\frac{\alpha_-^3}{\alpha_+^4}}\,,\ \frac{c_0^2 \varepsilon_0^{5/2} \alpha_-^2}{\alpha_+^2} \sqrt{\alpha_-}\,,\ \frac{c_0^2 \varepsilon_0^2 \alpha_-^2}{\alpha_+^2} 2^{N_0} s_0 \delta \right\} \\ &=: \frac{3\pi}{\eta} \end{aligned}$$

with c° independent of α_-, α_+, β, $\overline{\beta}$, δ, ε_0 and s_0. We then see that $|\gamma_*(T) - \gamma_*(0)|$ is lower bounded by 3π as soon as

$$\eta < 1\,. \tag{74}$$

The last step of the proof consists of proving that inequalities (68), (70) and (74) may be simultaneously satisfied. This is discussed in [18, Remark 5.1]. □

4 Chaos in a Binary Asteroid System

This section describes the main steps of [5] in constructing explicitly a topological horseshoe; henceforth providing evidences of the existence of symbolic dynamics. The construction essentially relies on the introduction of a two-dimensional Poincaré map from which invariants are computed. Following arguments presented in [20], the introduction of ad hoc sets and the computation of their images provide the self-covering relationships needed to conclude.

We fix $\beta = \overline{\beta} = 80$, which corresponds to take $\mu = 1$ and $\kappa \sim 40$; see (3). We interpret the Hamiltonian $\widehat{\mathrm{H}}_{\mathrm{J}}$ with this choice of parameters as governing the (averaged out after many periods of the reference asteroid) motions of a binary asteroid system interacting with a massive body, with the Jacobi reduction referred at one of the two twin asteroids. The parity triggered by the equality $\beta = \overline{\beta}$ reflects on the Taylor–Fourier coefficients of the expansion

$$\widehat{\mathrm{H}}_{\mathrm{J}}(\mathrm{R}, \mathrm{G}, \mathrm{r}, \mathrm{g}) = \frac{\mathrm{R}^2}{2} + \frac{(\mathrm{C}-\mathrm{G})^2}{2\mathrm{r}^2} - \frac{1}{\mathrm{r}} + \frac{1}{\mathrm{r}} \sum_{\nu=1}^{\infty} q_\nu(\mathrm{G}, \mathrm{g}) \left(\frac{\beta}{\mathrm{r}}\right)^\nu \tag{75}$$

accordingly to

$$q_\nu(\mathrm{G}, \mathrm{g}) = \begin{cases} \displaystyle\sum_{p=0}^{\nu/2} \tilde{q}_p(\mathrm{G}) \cos 2p\,\mathrm{g} & \text{if } \nu \text{ is even} \\ 0 & \text{otherwise}\,. \end{cases} \tag{76}$$

In our numeric implementations, we truncated the infinite sum in (75) up to a certain order $\nu_{\max}$, chosen so that the results did not vary increasing it again. Moreover, as the coefficients $q_\nu(\mathrm{G}, \cdot)$ in (76) are π-periodic, without loss of generality, we restricted $\mathrm{g} \in \mathbb{T}/2 \sim [0, \pi)$. We describe the steps we followed in our numerical analysis, recalling the reader to use Footnote 3 to relate the values in (77) and (78) with the homonymous ones in [5].

Construction of a 2-dimensional Poincaré map The motions of (75) evolve on 3-dimensional manifolds $\mathcal{M}_c$ labeled by the constant value c of the energy. The structure of $\widehat{\mathrm{H}}_{\mathrm{J}}$ allows to reduce the coordinate R after fixing c and hence to identify $\mathcal{M}_c$ as the 3-dimensional space of triples $\{(\mathrm{r}, \mathrm{G}, \mathrm{g})\}$. The dimension can be further reduced to 2 considering a plane Σ through a given $P_* = (\mathrm{r}_*, \mathrm{G}_*, \mathrm{g}_*)$ and perpendicular to the velocity vector $V_* = (v_{\mathrm{r}}^*, v_{\mathrm{G}}^*, v_{\mathrm{g}}^*)$ of the the orbit through P_*. This leads us to construct a Poincaré map, which we define as follows. We start by defining two operators l and π consisting in "lifting" the initial two-dimensional seed $z = (\mathrm{G}, \mathrm{g})$ to the four-dimensional space $(\mathrm{R}, \mathrm{G}, \mathrm{r}, \mathrm{g})$ and "projecting" it back to plan after the action of the flow-map $\Phi^t_{\widehat{\mathrm{H}}_{\mathrm{J}}}$ during the first return time τ. The lift operator reconstructs the four-dimensional state vector from a seed on $D \times \mathbb{T}/2$, where the domain D of the variable G is a compact subset of the form $[-1, 1]$. For a suitable $(\mathcal{A}, A) \subset \mathbb{R}^2 \times \mathbb{R}^2$, its definition reads

$$\begin{aligned} l : \quad & D \times \mathbb{T}/2 \supset \mathcal{A} \to D \times \mathbb{T}/2 \times A \\ & z \mapsto \tilde{z} = l(z), \end{aligned}$$

where $\tilde{z} = (\mathrm{G}, \mathrm{g}, \mathrm{R}, \mathrm{r})$ satisfies the two following conditions:

1. *Planarity condition.* The triplet $(\mathrm{r}, \mathrm{G}, \mathrm{g})$ belongs to the plane Σ, i.e. r solves the algebraic condition $v_{\mathrm{r}}^*(\mathrm{r} - \mathrm{r}_*) + v_{\mathrm{G}}^*(\mathrm{G} - \mathrm{G}_*) + v_{\mathrm{g}}^*(\mathrm{g} - \mathrm{g}_*) = 0$.
2. *Energetic condition.* The component R solves the energetic condition $\widehat{\mathrm{H}}_{\mathrm{J}}(\mathrm{R}, \mathrm{G}, \mathrm{r}, \mathrm{g}) = c$.

The projector π is the projection onto the first two components of the vector,

$$\begin{aligned} \pi : \quad & D \times \mathbb{T}/2 \times A \to D \times \mathbb{T}/2 \\ & \tilde{z} = (z_1, z_2, z_3, z_4) \mapsto \pi(z) = (z_1, z_2)\,. \end{aligned}$$

The Poincaré mapping is therefore defined and constructed as

$$\begin{aligned} P : \quad & D \times \mathbb{T}/2 \to D \times \mathbb{T}/2 \\ & z \mapsto z' = P(z) = \left(\pi \circ \Phi^{\tau(z)}_{\widehat{\mathrm{H}}_{\mathrm{J}}} \circ l\right)(z)\,. \end{aligned}$$

The mapping is nothing else than a "snapshots" of the whole flow at specific return time τ. It should be noted that the successive (first) return time is in general function of the current seed (initial condition or current state), i.e. $\tau = \tau(\tilde{z})$, formally defined (if it exists) as

$$\tau(z) = \inf\left\{t \in \mathbb{R}_+, \left(\mathrm{r}(t), \mathrm{G}(t), \mathrm{g}(t)\right) \in \Sigma\right\},$$

where $\left(\mathrm{r}(t), \mathrm{G}(t), \mathrm{g}(t)\right)$ is obtained through the flow.

With C = 24.394, we fixed the initial values

$$\mathrm{R}_* = -0.0060\,, \quad \mathrm{G}_* = -0.804\,, \quad \mathrm{r}_* = 652.256\,, \quad \mathrm{g}_* = 1.4524\ \mathrm{rad} \tag{77}$$

and we obtained the results plotted in the first panel of Fig. 3. We invite the reader to compare this figure with the unperturbed phase portrait of Fig. 2. In particular, due to the non-integrability of the problem, chaotic zones appear, mostly distributed for positive values of G. This chaos was the object of our next investigations, as discussed below.

Hyperbolic fixed points and heteroclinic intersections Equilibrium points of the mapping P (i.e. periodic orbits of the Hamiltonian system (75)) have been found using a Newton algorithm with initial guesses distributed on a resolved grid of initial conditions in $D \times \mathbb{T}/2$. We found more than 20 fixed points x_*. The eigensystems associated to the fixed points have been computed to determine the local stability properties. The result of the analysis is displayed on Fig. 3 along with the following convention: hyperbolic fixed points appear as red crosses, elliptical points are marked with blue circles.

The local stable manifold associated to an hyperbolic point $x_\star$,

$$\mathcal{W}^s_{\text{loc.}}(x_\star) = \left\{ x \mid \|P^n(x) - x_\star\| \to 0,\ n \in \mathbb{N}_+, n \to \infty \right\},$$

can be grown by computing the images of a fundamental domain $I \subset E_s(x_\star)$, $E_s(x_\star)$ being the stable eigenspace associated to the saddle point $x_\star$. The local unstable manifolds $\mathcal{W}^u_{\text{loc.}}(x_\star)$ were similarly computed, but changing the sign of the time integration. See Fig. 3.

Covering relations Let us introduce some notations. Let N be a compact set contained in $\mathbb{R}^2$ and $u(N) = s(N) = 1$ being, respectively, the *exit* and *entry dimension* (two real numbers such that their sum is equal to the dimension of the space containing N); let $c_N : \mathbb{R}^2 \to \mathbb{R}^2$ be an homeomorphism such that $c_N(N) = [-1, 1]^2$; let $N_c = [-1, 1]^2$, $N_c^- = \{-1, 1\} \times [-1, 1]$, $N_c^+ = [-1, 1] \times \{-1, 1\}$; then, the two set $N^- = c_N^{-1}(N_c^-)$ and $N^+ = c_N^{-1}(N_c^+)$ are, respectively, the *exit set* and the *entry set*. In the case of dimension 2, they are topologically a sum of two disjoint intervals. The quadruple $(N, u(N), s(N), c_N)$ is called a *h-set* and N is called *support* of the h-set. Finally, let $S(N)^l_c = (-\infty, -1) \times \mathbb{R}$, $S(N)^r_c = (1, \infty) \times \mathbb{R}$, and $S(N)^l = c_N^{-1}(S(N)^l_c)$, $S(N)^r = c_N^{-1}(S(N)^r_c)$ be, respectively, the left and the right side of N. The general definition of covering relation can be found in [9]. Here we provide a simplified notion, suited to the case that N is two-dimensional, based on [20, Theorem 16].

Definition 2 Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a continuous map and N and M the supports of two h-sets. We say that M f-covers N and we denote it by $M \overset{f}{\Longrightarrow} N$ if:

(1) $\exists q_0 \in [-1, 1]$ such that $f(c_N([-1, 1] \times \{q_0\})) \subset \text{int}(S(N)^l \bigcup N \bigcup S(N)^r)$,
(2) $f(M^-) \bigcap N = \emptyset$,
(3) $f(M) \bigcap N^+ = \emptyset$.

Conditions (2) and (3) are called, respectively, *exit* and *entry condition*.

The notions of covering (including self-covering) relations are useful in defining *topological horseshoe* [9, 20].

Definition 3 Let N_1 and N_2 be the supports of two disjoint h-sets in $\mathbb{R}^2$. A continuous map $f : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a *topological horseshoe* for N_1 and N_2 if

$$N_1 \overset{f}{\Longrightarrow} N_1\,, \quad N_1 \overset{f}{\Longrightarrow} N_2\,, \quad N_2 \overset{f}{\Longrightarrow} N_1\,, \quad N_2 \overset{f}{\Longrightarrow} N_2\,.$$

Topological horseshoes are associated to symbolic dynamics as discussed in [9, Theorem 2] and in [20, Theorem 18].

The topological horseshoe Based on the couple of hyperbolic fixed points

$$\begin{cases} q_1 = (\mathrm{g}_1, \mathrm{G}_1) = (0.203945459, 0.665706), \\ q_2 = (\mathrm{g}_2, \mathrm{G}_2) = (0.278077917, 0.714484), \end{cases} \tag{78}$$

we define two sets $N_1, N_2 \subset \mathbb{R}^2$ which are supports of two h-sets as follows:

$$\begin{cases} N_1 = q_1 + A_1 v_1^s + B_1 v_1^u, \\ N_2 = q_2 + A_2 v_2^s + B_2 v_2^u, \end{cases}$$

where v_1^s, v_1^u, v_2^s, v_2^u are the stable and the unstable eigenvectors related to q_1, q_2, respectively, and the A_1, A_2, B_1 and B_2 are numbers suitably chosen in a grid of values. Then the following covering relations are numerically detected

$$N_1 \overset{P}{\Longrightarrow} N_1\,, \quad N_1 \overset{P}{\Longrightarrow} N_2\,, \quad N_2 \overset{P}{\Longrightarrow} N_1\,, \quad N_2 \overset{P}{\Longrightarrow} N_2\,,$$

proving the numerical evidence of a topological horseshoe, i.e., existence of symbolic dynamics for P. The obtained horseshoe associated to q_1 and q_2 with the aforementioned parameters is illustrated in Fig. 3.

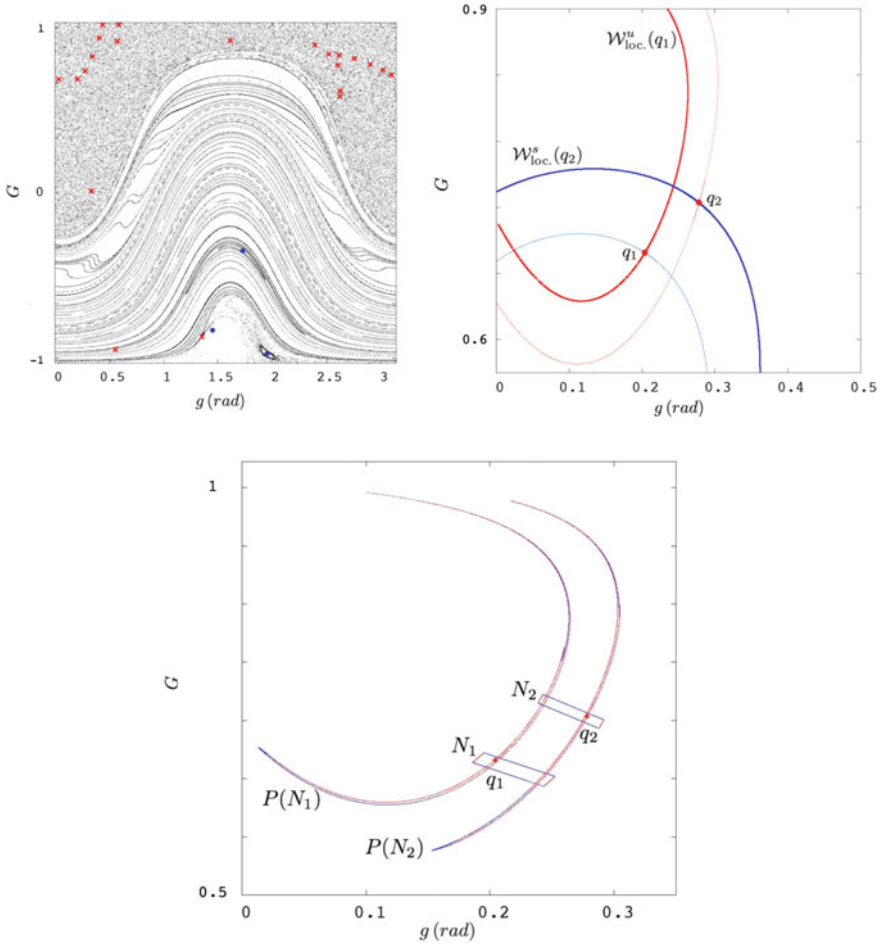


Fig. 3 Composite panels illustrating our main steps in constructing the topological horseshoe. (Top left) The continuous flow is reduced to a 2-dimensional mapping by introducing a suitable Poincaré map P. The phase space contains both elliptic (blue) and hyperbolic (red) fixed-points. (Top right) Finite pieces of the stable and unstable manifolds might be constructed from the knowledge of the eigensystem derived from the linearisation DP. (Bottom) Carefully chosen sets and their images under P provide the covering relations and imply existence of symbolic dynamics. See text for details

Acknowledgements The authors acknowledge the European Research Council (Grant 677793 Stable and Chaotic Motions in the Planetary Problem) for supporting them during the completion of the results described in the paper; warmly thank the organisers of I-CELMECH Training School that held in Milan in winter 2020 for their interest and especially U. Locatelli for a highlighting discussion.

Appendix 1: Outline of the Proof of Theorem 2

In this section we provide some technical details of the proof of Theorem 2. For the full proof we refer to [18].

The proof is by recursion. We assume that, at a certain step, we have a system of the form

$$\mathrm{H}(\mathbf{I}, \boldsymbol{\varphi}, \mathbf{J}(\mathbf{p}, \mathbf{q}), y) = \mathrm{h}(\mathbf{I}, \mathbf{J}(\mathbf{p}, \mathbf{q}), y) + g(\mathbf{I}, \mathbf{J}(\mathbf{p}, \mathbf{q}), y, x) + f(\mathbf{I}, \boldsymbol{\varphi}, \mathbf{J}(\mathbf{p}, \mathbf{q}), y, x) \quad (79)$$

where $f \in \mathcal{O}_{\rho,s,\delta,r,\xi}$, $g \in \mathcal{N}_{\rho,s,\delta,r,\xi}$. At the first step, just take $g \equiv 0$.

After splitting f on its Taylor–Fourier basis

$$f = \sum_{k,h,j} f_{khj}(\mathbf{I}, y, x) e^{\mathrm{i}k\cdot\boldsymbol{\varphi}} \mathbf{p}^h \mathbf{q}^j .$$

one looks for a time-1 map

$$\Phi = e^{\mathcal{L}_\phi} = \sum_{k=0}^{\infty} \frac{\mathcal{L}_\phi^k}{k!} \qquad \mathcal{L}_\phi(f) := \{\phi,\ f\}$$

generated by a small Hamiltonian ϕ which will be taken in the class $\mathcal{Z}_{\rho,s,\delta,r,\xi}$ in (52). Here,

$$\{\phi,\ f\} := \sum_{i=1}^{n} (\partial_{\mathbf{I}_i}\phi\partial_{\boldsymbol{\varphi}_i} f - \partial_{\mathbf{I}_i} f \partial_{\boldsymbol{\varphi}_i}\phi) + \sum_{i=1}^{m} (\partial_{\mathbf{p}_i}\phi\partial_{\mathbf{q}_i} f - \partial_{\mathbf{p}_i} f \partial_{\mathbf{q}_i}\phi) + (\partial_y\phi\partial_x f - \partial_y f \partial_x\phi)$$

denotes the Poisson parentheses of ϕ and f. One lets

$$\phi = \sum_{\substack{(k,h,j):\\ (k,h-j)\neq(0,0)}} \phi_{khj}(\mathbf{I}, y, x) e^{\mathrm{i}k\cdot\boldsymbol{\varphi}} \mathbf{p}^h \mathbf{q}^j . \quad (80)$$

The operation

$$\phi \to \{\phi, \mathrm{h}\}$$

acts diagonally on the monomials in the expansion (80), carrying

$$\phi_{khj} \to -\big(\omega_y \partial_x \phi_{khj} + \lambda_{khj}\phi_{khj}\big)\,, \quad \text{with} \quad \lambda_{khj} := (h-j)\cdot\omega_{\mathbf{J}} + \mathrm{i}k\cdot\omega_{\mathbf{I}}\,. \tag{81}$$

Therefore, one defines

$$\{\phi, \mathrm{h}\} =: -D_\omega \phi\,.$$

The formal application of $\Phi = e^{\mathcal{L}_\phi}$ yields:

$$e^{\mathcal{L}_\phi}\mathrm{H} = e^{\mathcal{L}_\phi}(\mathrm{h}+g+f) = \mathrm{h} + g - D_\omega\phi + f + \Phi_2(\mathrm{h}) + \Phi_1(g) + \Phi_1(f) \tag{82}$$

where the $\Phi_h := \Phi_h := \sum_{j\geq h} \frac{\mathcal{L}_\phi^j}{j!}$'s are the tails of $e^{\mathcal{L}_\phi}$.

Next, one requires that the residual term $-D_\omega\phi + f$ lies in the class $\mathcal{N}_{\rho,s,\delta,r,\xi}$ in (53)

$$(-D_\omega\phi + f) \in \mathcal{N}_{\rho,s,\delta,r,\xi} \tag{83}$$

for ϕ.

Since we have chosen $\phi \in \mathcal{Z}_{\rho,s,\delta,r,\xi}$, by (81), we have that also $D_\omega\phi \in \mathcal{Z}_{\rho,s,\delta,r,\xi}$. So, Eq. (83) becomes

$$-D_\omega\phi + \widetilde{f} = 0\,.$$

In terms of the Taylor–Fourier modes, the equation becomes

$$\omega_y\partial_x\phi_{khj} + \lambda_{khj}\phi_{khj} = f_{khj} \qquad \forall\,(k,h,j):\ (k,h-j)\neq(0,0)\,. \tag{84}$$

In the standard situation, one typically proceeds to solve such equation via Fourier series:

$$f_{khj}(\mathbf{I},y,x) = \sum_\ell f_{khj\ell}(\mathbf{I},y)e^{\mathrm{i}\ell x}\,, \qquad \phi_{khj}(\mathbf{I},y,x) = \sum_\ell \phi_{khj\ell}(\mathbf{I},y)e^{\mathrm{i}\ell x}$$

so as to find $\phi_{khj\ell} = \dfrac{f_{khj\ell}}{\mu_{khj\ell}}$ with the usual denominators $\mu_{khj\ell} := \lambda_{khj} + \mathrm{i}\ell\omega_y$ which one requires not to vanish via, e.g., a "diophantine inequality" to be held for all (k,h,j,ℓ) with $(k,h-j)\neq(0,0)$. In this standard case, there is not much freedom in the choice of ϕ. In fact, such solution is determined up to solutions of the homogenous equation

$$D_\omega\phi_0 = 0 \tag{85}$$

which, in view of the Diophantine condition, has the only trivial solution $\phi_0 \equiv 0$. *The situation is different if f is not periodic in x, or ϕ is not needed so*. In such a case, it is possible to find a solution of (84), corresponding to a non-trivial solution of (85), where small divisors do not appear. This is

$$\phi_{khj}(\mathbf{I}, y, x) = \begin{cases} \dfrac{1}{\omega_y} \displaystyle\int_0^x f_{khj}(\mathbf{I}, y, \tau) e^{\frac{\lambda_{khj}}{\omega_y}(\tau - x)} d\tau & \text{if } (k, h-j) \neq (0,0) \\ 0 & \text{otherwise.} \end{cases} \tag{86}$$

Multiplying by $e^{ik\varphi}$ and summing over k, h and j, we obtain

$$\phi(\mathbf{I}, \boldsymbol{\varphi}, p, q, y, x) = \frac{1}{\omega_y} \int_0^x \widetilde{f}\left(\mathbf{I}, \boldsymbol{\varphi} + \frac{\omega_{\mathbf{I}}}{\omega_y}(\tau - x), p e^{\frac{\omega_{\mathbf{J}}}{\omega_y}(\tau - x)}, q e^{-\frac{\omega_{\mathbf{J}}}{\omega_y}(\tau - x)}, y, \tau\right) d\tau \,.$$

In [18] it is proved that, under the assumptions (54), this function can be used to obtain a convergent time-one map and that the construction can be iterated so as to provide the proof of Theorem 2. The construction of the iterations and the proof of its convergence is obtained adapting the techniques of [19] to the present case.

References

1. Arnold, V.I.: Methods, Mathematical, of Classical Mechanics. Graduate Texts in Mathematics., vol. 60, 2nd edn. Springer, New York (1989). Translated from the Russian by K. Vogtmann and A. Weinstein
2. Arnold, V.I., Kozlov, V.V., Neishtadt, A.I.: Mathematical aspects of classical and celestial mechanics. Encyclopaedia of Mathematical Sciences, vol. 3, 3rd edn. Springer, Berlin (2006). [Dynamical systems. III], Translated from the Russian original by E. Khukhro
3. Bekov, A.A., Omarov, T.B.: Integrable cases of the Hamilton-Jacobi equation and some non-steady problems of celestial mechanics. Sov. Astron. **22**, 366–370 (1978)
4. Chierchia, L., Pinzari, G.: The planetary N-body problem: symplectic foliation, reductions and invariant tori. Invent. Math. **186**(1), 1–77 (2011)
5. Di Ruzza, S., Daquin, J., Pinzari, G.: Symbolic dynamics in a binary asteroid system. Commun. Nonlinear Sci. Numer. Simul. **91**, 105414 (2020)
6. Dullin, H.R., Montgomery, R.: Syzygies in the two center problem. Nonlinearity **29**(4), 1212–1237 (2016)
7. Féjoz, J.: Démonstration du 'théorème d'Arnold' sur la stabilité du système planétaire (d'après Herman). Ergod. Theory Dynam. Syst. **24**(5), 1521–1582 (2004)
8. Féjoz, J., Guardia, M.: Secular instability in the three-body problem. Arch. Ration. Mech. Anal. **221**(1), 335–362 (2016)
9. Gierzkiewicz, A., Zgliczyński, P.: A computer-assisted proof of symbolic dynamics in hyperion's rotation. Celest. Mech. Dyn. Astron. **131**(7), 33 (2019)
10. Giorgilli, A.: Appunti di Meccanica Celeste (2008). http://www.mat.unimi.it/users/antonio/meccel/Meccel_5.pdf
11. Giorgilli, A., Locatelli, U., Sansottera, M.: Secular dynamics of a planar model of the sun-jupiter-saturn-uranus system; effective stability in the light of Kolmogorov and Nekhoroshev theories. Regul. Chaotic Dyn. **22**(1), 54–77 (2017)

12. Guzzo, M., Lega, E.: Evolution of the tangent vectors and localization of the stable and unstable manifolds of hyperbolic orbits by Fast Lyapunov Indicators. SIAM J. Appl. Math. **74**(4), 1058–1086 (2014)
13. Jacobi, C.G.J.: Jacobi's lectures on dynamics. Texts and Readings in Mathematics, vol. 51, revised edn. Hindustan Book Agency, New Delhi (2009). Delivered at the University of Königsberg in the winter semester 1842–1843 and according to the notes prepared by C. W. Brockardt, Edited by A. Clebsch, Translated from the original German by K. Balagangadharan, Translation edited by Biswarup Banerjee
14. Laskar, J., Robutel, P.: Stability of the planetary three-body problem. I. Expansion of the planetary Hamiltonian. Celest. Mech. Dyn. Astron. **62**(3), 193–217 (1995)
15. Pinzari, G.: On the Kolmogorov set for many–body problems. Ph.D. thesis, Università Roma Tre, April (2009)
16. Pinzari, G.: A first integral to the partially averaged newtonian potential of the three-body problem. Celest. Mech. Dyn. Astron. **131**(5), 22 (2019)
17. Pinzari, G.: Euler integral and perihelion librations. Discret Contin. Dyn. Syst. **40**(12), 6919–6943 (2020)
18. Pinzari, G.: Perihelion librations in the secular three-body problem. J. Nonlinear Sci. **30**(4), 1771–1808 (2020)
19. Pöschel, J.: Nekhoroshev estimates for quasi-convex Hamiltonian systems. Math. Z. **213**(2), 187–216 (1993)
20. Zgliczynski, P., Gidea, M.: Covering relations for multidimensional dynamical systems. J. Differ. Equ. **202**(1), 32–58 (2004)

KAM Theory for Some Dissipative Systems

Renato Calleja, Alessandra Celletti and Rafael de la Llave

Abstract Dissipative systems play a very important role in several physical models, most notably in Celestial Mechanics, where the dissipation drives the motion of natural and artificial satellites, leading them to migration of orbits, resonant states, etc. Hence, there is the need to develop theories that ensure the existence of structures such as invariant tori or periodic orbits, and device efficient computational methods. In this work we study the existence of invariant tori for those dissipative systems known as *conformally symplectic* systems, which have the property that they transform the symplectic form into a multiple of itself. To give explicit examples of conformally symplectic systems, we will present two different models: a discrete system known as the standard map and a continuous system known as the spin-orbit problem. In both cases we will consider the conservative and dissipative versions, that will help to highlight the differences between the symplectic and conformally symplectic dynamics. For such dissipative systems we will present a KAM theorem in an a-posteriori format, originally developed in [44] for the symplectic case: assume we start with an approximate solution satisfying a suitable non-degeneracy condition, then we can find a true solution nearby. The theorem does not assume that the system is close to integrable. This method provides a very efficient algorithm which provides rigorous estimates close to optimal. Indeed, the method gives a criterion (the Sobolev

R. Calleja (✉)
Department of Mathematics and Mechanics, IIMAS, National Autonomous University of Mexico (UNAM), Apdo. Postal 20-126, C.P., 04510 Mexico D.F., Mexico
e-mail: calleja@mym.iimas.unam.mx
URL: https://mym.iimas.unam.mx/renato

A. Celletti
Department of Mathematics, University of Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italy
e-mail: celletti@mat.uniroma2.it
URL: http://www.mat.uniroma2.it/celletti

R. de la Llave
School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, GA 30332-1160, USA
e-mail: rafael.delallave@math.gatech.edu
URL: https://people.math.gatech.edurll6

G. Baù et al. (eds.), *New Frontiers of Celestial Mechanics: Theory and Applications*, Springer Proceedings in Mathematics & Statistics 399,
https://doi.org/10.1007/978-3-031-13115-8_3

blow up criterion) that allows to compute numerically the breakdown of invariant tori. We will review this method as well as an extension of J. Greene's method. Computing close to the breakdown, allows to discover new mathematical phenomena, such as the *bundle collapse mechanism*. We will also provide a short survey of the present state of KAM estimates for the existence of invariant tori in the conservative and dissipative standard maps and spin-orbit problems.

Keywords KAM theory · Invariant tori · Dissipative systems · Conformally symplectic systems

1 Introduction

Dissipative dynamical systems play a fundamental role in shaping the motions of physical problems. The role of dissipative forces in Celestial Mechanics is often of less importance with respect to the conservative forces, which are mainly given by the gravitational attraction between celestial bodies. Nevertheless dissipative forces are present at any size and time scale and their effects accumulate over time, so that even if some of them are negligible on a scale of centuries, they might be dominant in a scale of a million of years.

A partial list of dissipative forces includes tidal forces, Stokes drag, Poynting-Robertson effect, Yarkowski/YORP effects, atmospheric drag. These forces act on bodies of different dimensions, namely planets, satellites, spacecraft, dust particles, and in different epochs of the Solar system from the dynamics within the interplanetary nebula at the early stage of formation of the Solar system, to present times. For example, the effect of the Earth's atmosphere on the orbital lifetime of artificial satellites, happens in practical scales of time. It becomes therefore important to understand invariant structures (e.g., periodic orbits and quasi-periodic motions on invariant tori) in dissipative systems.

The definition of *dissipative system* is not uniform in the literature. Here we will adopt that a dissipative system has the property that the phase space volume contracts during the time evolution of the flow. In this work we will be concerned with a special class of dissipative systems known as *conformally symplectic* systems, see Definitions 1 and 2. These systems enjoy the property that the flow or the map transform the symplectic form into a multiple of itself.

Conformally symplectic systems have appeared in many applications (see, e.g., [6, 18, 19, 28, 51, 115]) or have been studied because they are geometrically natural objects [1, 5, 109].

For applications to Celestial Mechanics, an important source of conformally symplectic systems is that of a mechanical system with friction proportional to the velocity. This is the case of the so-called spin-orbit problem in Celestial Mechanics [33, 34, 42, 114], which will be presented in Sect. 2.3. It describes the motion of an oblate satellite around a central planet, under some simplifying assumptions like that the orbit of the satellite is Keplerian and that the spin-axis is perpendicular to the orbital

plane. When the satellite is assumed to be rigid, the problem is conservative, while when the satellite is assumed to be non-rigid, the problem is affected by a tidal torque. The dissipative part of the spin-orbit problem depends upon two parameters: the dissipative constant, which is a function of the physical properties of the satellite, and a drift term, which depends on the (Keplerian) eccentricity of the orbit. A discrete analogue of the spin-orbit problem is the dissipative standard map [35]. In Sect. 2.1 we will review conservative and dissipative versions of the standard map.

Indeed, the presence of a drift term is fundamental in conformally symplectic systems: while in the conservative case one can find an invariant torus with fixed frequency by adjusting the initial conditions, in the dissipative case it is not possible to just tune the initial conditions to obtain a quasi-periodic solution of a fixed frequency. One needs to adjust a drift parameter to find an invariant torus with preassigned frequency (for some appropriate choice of initial conditions).

We stress that adding a dissipation to a Hamiltonian system is a very singular perturbation: the Hamiltonian admits quasi-periodic solutions with many frequencies, while a system with positive dissipation leads to attractors with only one quasi-periodic solution. To obtain attractors with a fixed frequency, one needs to adjust the drift parameters.

The existence of invariant tori is the subject of the celebrated Kolmogorov–Arnold–Moser (KAM) theory ([2, 78, 90], see also [3, 43, 53, 65, 91, 96, 112]) which, in its original formulation, proved the persistence of invariant tori in nearly–integrable Hamiltonian systems. The theory can be developed under two main assumptions:

- the frequency vector must satisfy a Diophantine condition (to deal with the so-called small divisors problem),
- a non–degeneracy condition must be satisfied (to ensure the solution of the cohomological equations providing the approximate solutions).

Also, geometric properties of the system play an important role. Notably, the original results were developed for Hamiltonian systems, but this has been greatly extended.

A KAM theory for non-Hamiltonian systems with adjustment of parameters was developed in the remarkable and pioneer paper [92], and later in [10, 11, 41, 86, 101]. A KAM theory for conformally symplectic systems with adjustment of parameters was developed in [21] using the so-called *automatic reducibility* method introduced in [44]. The paper [21] produces an a-posteriori result. A-posteriori means that the existence of an approximate solution, which satisfies an invariance equation up to a small error, ensures the existence of a true solution of the invariance equation, provided some non-degeneracy conditions and smallness conditions on parameters are satisfied.

The automatic reducibility proofs of KAM theorem provide very efficient and stable algorithms to construct invariant tori in the symplectic [56, 62, 63] and conformally symplectic [18, 19, 26, 28, 64] case. The a-posteriori format guarantees that these solutions are correct. Indeed, it was proved that the algorithm leads to a continuation method in parameters that, given enough resources, reaches arbitrarily close the boundary of the set of parameters for which the solution exists. In [20] it

was found numerically that the tori—which are normally hyperbolic– break down because the stable bundle becomes close to the tangent, even if the stable Lyapunov exponent (which is given by the conformal symplectic constant) remains away from zero.

1.1 Consequences of the A-Posteriori Method for Conformally Symplectic Systems

The results presented in these notes are part of a more systematic program of providing KAM theorems in an a-posteriori format with many consequences that, for completeness, we shortly review below.

• Regularity results.

The a-posteriori format, leads automatically to many regularity results: deducing finitely differentiable results from analytic ones, bootstrap of regularity, Whitney dependence on the frequency. We will not even mention these regularity results, but we point that in the conformally symplectic case, we can obtain several rather striking geometric results. The conformally symplectic systems are very rigid. A classic result that plays a role is the *paring rule* of Lyapunov exponents. The conformal geometric structure restricts severely the Lyapunov exponents that can appear [50, 115].

• Rigidity of neighborhoods of tori.

In [22] it is shown that the dynamics in a neighborbood of a Lagrangian torus is conjugate to a rotation and a linear contraction. In particular, the only invariant in a neighborhood is the rotation and all the tori with the same rotation are analytically conjugate in a neighborbood.

• Greene's method.

An analogue of Greene's method ([59]) to compute the analyticity breakdown is given in [27], which presents a partial justification of the method. It is proved that when the invariant attractor exists, then one can predict the eigenvalues of the periodic orbits approximating the torus for parameter values close to those of the attractor.

• Whiskered tori.

In [24, 25] one can find a theory of whiskered tori in conformally symplectic systems. This theory involves interactions of dynamics and geometry. The theory allows—there are examples—that the stable and unstable bundles are trivial, but somewhat surprisingly, concludes that the center bundles have to be trivial.

• The singular limit of zero dissipation.

We showed in [23] that, if one fixes the frequency ω, one can choose the drift parameter μ as a function of the perturbation in a smooth way: $\mu = \mu(\omega, \varepsilon) \equiv \mu_\varepsilon(\omega)$. Note however that $\mu_0(\omega) = 0$, but for $\varepsilon > 0$, the function μ_ε is invertible so that the function $\mu_\varepsilon(\omega)$ is a smooth function with a limit as $\varepsilon \to 0$. Nevertheless, the sets of ω that appear have a complicated behaviour (devil's staircase). Hence, the floating frequency KAM methods, e.g. [2, 39, 95], have difficulty dealing with this limit.

One of the advantages of the a-posteriori theorems is that they can validate approximate solutions, no matter how they are obtained. We have already mentioned the validation of numerical computations. It turns out that one can also validate formal asymptotic expansions and obtain estimates on the domains of existence of the tori in the singular limit [23]. This limit has also been studied numerically [12, 13], leading to the conjecture that the Lindstedt series are Gevrey. A proof of the conjecture is given in [14].

• Breakdown of the rotational tori.

One of the consequences of the conformal symplectic geometry is the "pairing rule" for exponents [115]. Hence the tori, which have a dynamics which is a rotation, must have normal exponents coinciding with λ. The tori are normally hyperbolic attractors. Notice that the loss of hyperbolicity cannot happen because of the exponents break down. This leads to the mechanism of bundle collapse that was discovered in [20] and will be discussed in more detail in Sect. 6.

1.2 Organization of the Paper

The work is organized as follows. In Sect. 2 we present the conservative and dissipative standard maps and spin-orbit problems. Conformally symplectic systems and Diophantine vectors are introduced in Sect. 3. The definition of invariant tori and the statement of the KAM theorem for conformally symplectic systems is given in Sect. 4. Two numerical methods for the computation of the breakdown threshold of invariant attractors is presented in Sect. 5. The relation between the collision of invariant bundles and the breakdown of the tori is described in Sect. 6. Applications of KAM estimates to the conservative/dissipative standard maps and spin-orbit problems are briefly recalled in Sect. 7.

2 Conservative/Dissipative Standard Maps and Spin-Orbit Problems

In this Section we present two models, a discrete and a continuous one, that will help to have a qualitative understanding of the main features of conservative and dissipative systems. The first example is a discrete model, known as the *standard map* (see Sects. 2.1 and 2.2). The continuous example is a physical model, known as the *spin-orbit problem*, which is closely related to the standard map (see Sect. 2.3). In both cases we present their conservative and dissipative formulations.

2.1 The Conservative Standard Map

The standard map is a discrete model introduced by Chirikov in [40], which has been widely studied to understand several features of dynamical systems, such as regular motions, chaotic dynamics, breakdown of invariant tori, existence of periodic orbits, etc. The standard map is a 2-dimensional discrete system in the variables $(y, x) \in \mathbb{R} \times \mathbb{T}$, which is described by the formulas:

$$\begin{aligned} y' &= y + \varepsilon\, V(x) \\ x' &= x + y' \,, \end{aligned} \tag{1}$$

where $\varepsilon > 0$ is called the *perturbing parameter* and $V = V(x)$ is an analytic function.

A wide number of articles and books in the literature (see, e.g., [60, 83]) deals with the classical (Chirikov) standard map [40] obtained setting $V(x) = \sin x$ in (1).

Instead of (1) we can use an equivalent notation and write the standard map assigning an integer index to each iterate:

$$\begin{aligned} y_{j+1} &= y_j + \varepsilon V(x_j) \\ x_{j+1} &= x_j + y_{j+1} = x_j + y_j + \varepsilon V(x_j) \qquad \text{for } \; j \geq 0 \,. \end{aligned} \tag{2}$$

We can easily verify that the standard map (2) satisfies the following properties, that will be useful for further discussion.

(A) The standard map is integrable for $\varepsilon = 0$. In fact, for $\varepsilon = 0$ one gets the formulas:

$$\begin{aligned} y_{j+1} &= y_j = y_0 \\ x_{j+1} &= x_j + y_{j+1} = x_j + y_j = x_0 + j y_0 \qquad \text{for } j \geq 0 \,, \end{aligned}$$

which show that the mapping is integrable, since y_j is constant and x_j increases by y_0. For $\varepsilon \neq 0$ but small, the map is nearly-integrable.

(B) The standard map is conservative, since the determinant of its Jacobian is equal to one:

$$\det \begin{pmatrix} \frac{\partial y_{j+1}}{\partial y_j} & \frac{\partial y_{j+1}}{\partial x_j} \\ \frac{\partial x_{j+1}}{\partial y_j} & \frac{\partial x_{j+1}}{\partial x_j} \end{pmatrix} = \det \begin{pmatrix} 1 & \varepsilon V_x(x_j) \\ 1 & 1 + \varepsilon V_x(x_j) \end{pmatrix} = 1 \,.$$

(C) The standard map satisfies the so-called twist property, which amounts to requiring that for a constant $c \in \mathbb{R}$, using the formulation (1):

$$\left| \frac{\partial x'}{\partial y} \right| \geq c > 0$$

for all $(y, x) \in \mathbb{R} \times \mathbb{T}$. From (1) we have that the twist property is trivially satisfied, since

$$\frac{\partial x'}{\partial y} = 1 \, .$$

The twist property is not satisfied when considering a *slight* modification of (1), yielding a discrete system which is known as the *non-twist* standard map (see, e.g., [47, 48]). This mapping is described by the equations:

$$\begin{aligned} y' &= y + \varepsilon \, V(x) \qquad\qquad y \in \mathbb{R} \, , \; x \in \mathbb{T} \\ x' &= x + a(1 - y'^2) \end{aligned}$$

with $a \in \mathbb{R}$. In this case, the twist condition is violated along a curve in the (y, x) plane.

Systems violating the twist condition appear in Celestial Mechanics, for example in the critical inclination for the motion near an oblate planet [79]. One of the advantages of the KAM results we will establish is that we do not need to assume global non-degeneracy conditions on the map, but rather some properties of the approximate solution. We just need to assume that a $d \times d$ matrix is invertible. The matrix is an explicit algebraic expression on derivatives of the approximate solution and averages.

Figure 1 shows the graph of the iterates of the standard map for several values of the perturbing parameter and for several initial conditions in each plot.

From the upper left plot of Fig. 1, we see that for $\varepsilon = 0$ the system is integrable; the initial conditions has been chosen to give *rotational* quasi–periodic curves (lying on straight lines).

When we switch-on the perturbation, even for small values as $\varepsilon = 0.1$, the system becomes non–integrable. It is easy to check that there exists a stable equilibrium point at $(\pi, 0)$ and an unstable one at $(0, 0)$. The quasi–periodic (KAM) curves are distorted with respect to the integrable case and the stable point $(\pi, 0)$ is surrounded by elliptic *librational* islands. The amplitude of the islands increases as ε gets larger, as it is shown for $\varepsilon = 0.4$ where we also notice the appearance of minor resonances. Chaotic dynamics is clearly present for $\varepsilon = 0.7$ around the unstable equilibrium point, while the number of rotational quasi–periodic curves decreases when increasing the perturbing parameter. In particular, for $\varepsilon = 0.9$ we see large chaotic regions, a few quasi–periodic curves, new islands around higher–order periodic orbits. Finally, for $\varepsilon = 1$ we have only chaotic and librational motions, while quasi–periodic curves disappear.

As we will mention in Sect. 7, there is a wide literature on KAM applications to the standard map to prove the existence of invariant rotational tori with fixed frequency, see [29, 46, 56].

The example we have presented in this Section shows a marked difference with respect to the model that will be presented in Sect. 2.2, thus witnessing the divergence of the dynamical behaviour between conservative and dissipative dynamical systems. This difference is clearly demonstrated by the dynamics associated to the conservative

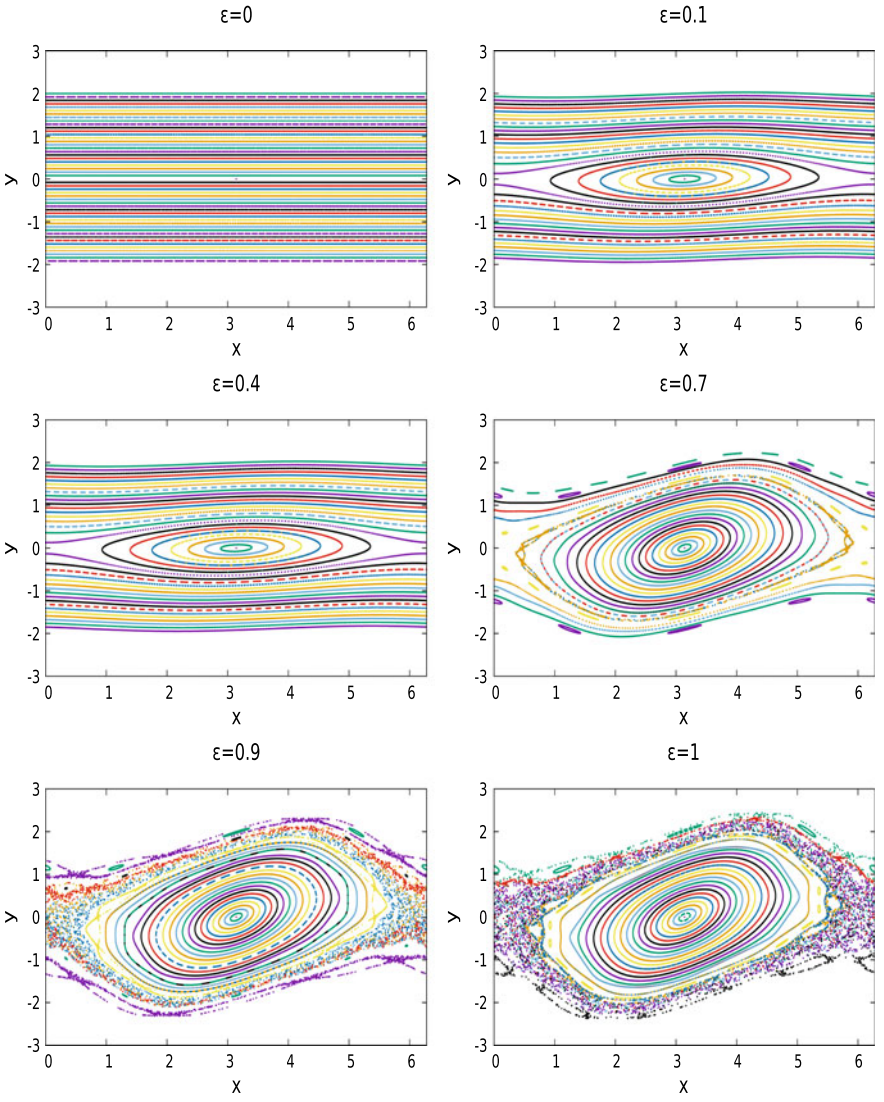


Fig. 1 Graphs of the conservative standard map for different values of the perturbing parameter and different initial conditions

and dissipative standard maps, as well as by that of more complex systems, like the conservative and dissipative spin-orbit problems, which will be described in Sect. 2.3.

2.2 The Dissipative Standard Map

The dissipative standard map is obtained from (1) adding two parameters: a dissipative parameter $0 < \lambda < 1$ and a drift parameter μ. For $(y, x) \in \mathbb{R} \times \mathbb{T}$, the equations describing the dissipative standard map are the following:

$$\begin{aligned} y' &= \lambda y + \mu + \varepsilon\, V(x) \\ x' &= x + y' \,, \end{aligned} \tag{3}$$

where $\lambda, \varepsilon \in \mathbb{R}_+$, $\mu \in \mathbb{R}$. We remark that we obtain the conservative standard map when $\lambda = 1$ and $\mu = 0$. We also remark that the Jacobian of the mapping (3) is equal to λ, which gives a measure of the rate of contraction or expansion of the area of the phase space. There are several results related to the existence of attractors in the dissipative standard map; a partial list of papers is the following: [7, 8, 55, 70, 72, 104, 110, 113, 116]. Rigorous mathematical works on strange attractors for dissipative 2-D maps with twist are [82, 84, 111].

It is also important to stress that for $\varepsilon = 0$ the trajectory $\{y \equiv \frac{\mu}{1-\lambda}\} \times \mathbb{T}$, or equivalently $\{\omega \equiv \frac{\mu}{1-\lambda}\} \times \mathbb{T}$, is invariant. In fact, for $\varepsilon = 0$ we have $y' = \lambda y + \mu$ and since we are looking for an invariant object, we need to have $y' = y$. Hence, we must solve the equation

$$y = \lambda y + \mu \,. \tag{4}$$

On the other hand, the frequency ω associated to the standard map is, by definition, given by

$$\omega = \lim_{j \to \infty} \frac{x_j}{j} \,,$$

which yields $\omega = y$. Combining this last result with (4), we obtain

$$\omega = \frac{\mu}{1 - \lambda} \,,$$

showing the strong relation between the frequency and the drift, which cannot be chosen independently. In particular, if we fix the frequency (as it will be required in the KAM theorem of Sect. 4.2), then we need to tune properly the drift parameter μ. This is a substantial difference with respect to the conservative case; dissipative dynamical systems will require a procedure to prove KAM theory differently than in the conservative case.

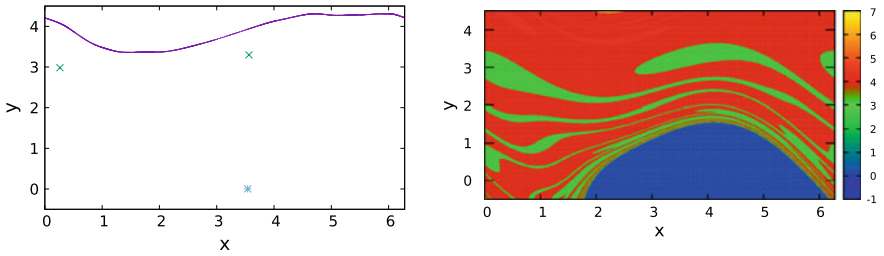


Fig. 2 Left: Attractors for the dissipative standard map with $\varepsilon = 0.9$, $\lambda = 0.91$, $\mu = 2\pi(1 - \lambda)(\frac{\sqrt{5}-1}{2})$. Right: The corresponding basins of attraction using a color scale providing the frequency

The dynamics associated to the dissipative standard map admits (see Fig. 2) attracting periodic orbits and invariant curve attractors; for different parameters and initial conditions, there appear also strange attractors which have an intricate geometrical structure [83, 111]: introducing a suitable definition of dimension, the strange attractors are shown to have, for some parameter values, a non–integer dimension (namely a *fractal* dimension). We will not consider these cases and concentrate on those in which the attractor is a one-dimensional smooth torus and the motion is smoothly conjugate to a rotation.

We remark that, due to the dissipative character of the map, there might exist at most one invariant curve attractor, while there might be more coexisting periodic orbits (see Fig. 2 and [55]) or strange attractors.

The existence and breakdown of smooth invariant tori in the dissipative standard map have been recently studied in [26] (see also [20]).

Each of the attractors of Fig. 2 is characterized by an associated *basin of attraction*, which is composed by the set of initial conditions (x_0, y_0) whose evolution ends on the given attractor. Figure 2, right, shows the basins of attraction for the case in Fig. 2, left; they have been obtained taking a grid of 500×500 initial conditions and looking at their evolution after having performed a number of preliminary iterations.

We want to stress that the role of the drift parameter μ is of paramount importance in dissipative systems, since an inappropriate choice might prevent to find a specific attractor. An example is given in Fig. 3, where we look for the torus with frequency equal to the golden ratio multiplied by the factor 2π, namely $\omega = 2\pi\,\frac{\sqrt{5}-1}{2} \simeq 3.8832$, for the dissipative standard map with $\varepsilon = 0.1$, $\lambda = 0.9$. The upper left panel shows that taking $\mu = 0$, the solution spirals on the point attractor at $(\pi, 0)$; taking $\mu = 0.1$ (Fig. 3, upper right panel) leads to an attractor which has frequency different than ω, while the right choice corresponds to $\mu = 0.0617984$ as in the left bottom panel of Fig. 3. We present in Fig. 3, bottom right panel, the behaviour of the drift as a function of the dissipative parameter λ, which shows that μ tends to zero in the limit of the conservative case, as it is expected.

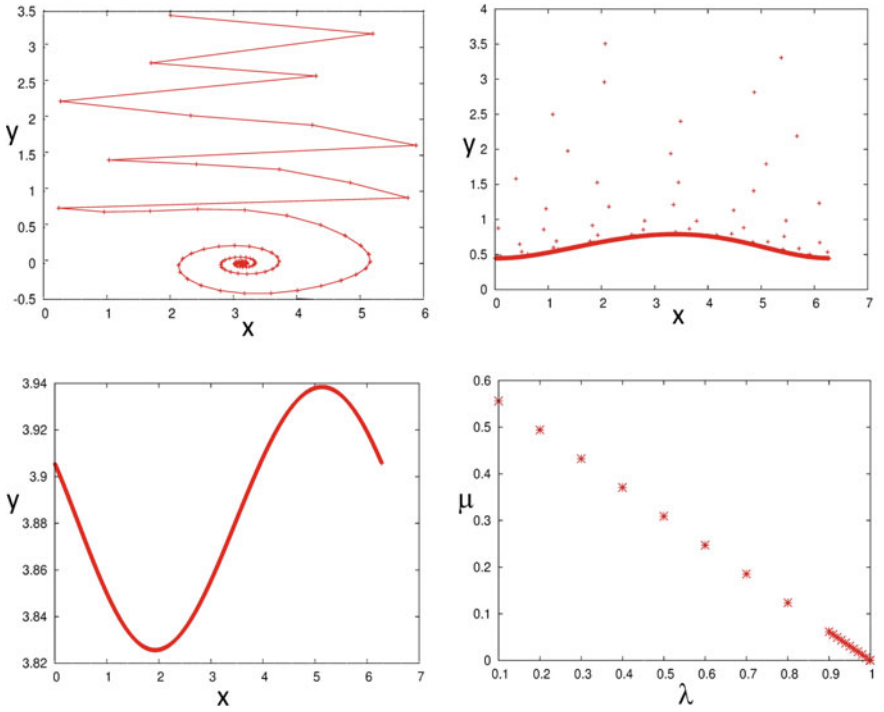


Fig. 3 The dissipative standard map for different values of the drift: $\mu = 0$ upper left, $\mu = 0.1$ upper right, $\mu = 0.0617984$ bottom left. Graph of μ versus λ, bottom right

The twist condition for the dissipative standard map is a condition that now involves the parameters. A non-twist version of the dissipative standard map is the following map,

$$\begin{aligned} y' &= \lambda y + \varepsilon\, V(x) \\ x' &= x + (y' - a)^2 + \mu\,. \end{aligned} \tag{5}$$

In Fig. 4, we notice that this map has parameter values where the rotation number does not change in a monotone direction when we change the parameter a. See [17] for a study of the invariant circles of the map (5).

2.3 The Spin-Orbit Problems

An interesting example of a continuous system which shows the main dynamical features of regular and chaotic invariant objects is the so-called spin-orbit problem in Celestial Mechanics. The conservative version of the model is based upon the

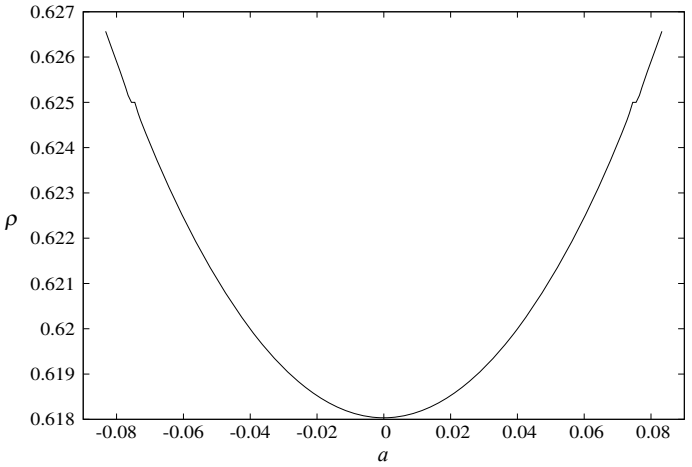


Fig. 4 Rotation number ρ in the map (5) w.r.t. the parameter a. Reproduced from [17]

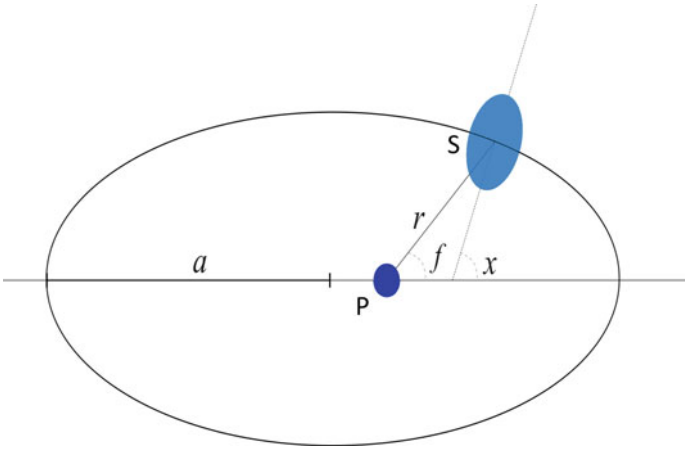


Fig. 5 The spin-orbit problem

following assumptions. We consider a triaxial satellite, say $\mathcal{S}$, with principal moments of inertia $I_1 < I_2 < I_3$. We assume that the satellite moves on a Keplerian orbit around a central planet, say $\mathcal{P}$, while it rotates around a spin–axis perpendicular to the orbit plane and coinciding with its shortest physical axis.

We take a reference system centered in the planet and with the horizontal axis coinciding with the direction of the semimajor axis. We denote by r the orbital radius, by f the true anomaly, while we denote by x the angle between the semimajor axis and the direction of the longest axis of the ellipsoidal satellite (see Fig. 5).

The equation of motion describing the conservative spin-orbit problem is

$$\ddot{x} + \varepsilon \left(\frac{a}{r}\right)^3 \sin(2x - 2f) = 0 \,, \tag{6}$$

where $\varepsilon = \frac{3}{2}\frac{I_2 - I_1}{I_3}$ is a parameter which measures the equatorial flattening of the satellite. Equation (6) is associated to the one-dimensional, time-dependent Hamiltonian function:

$$\mathcal{H}(y, x, t) = \frac{y^2}{2} - \frac{\varepsilon}{2}\Big(\frac{a}{r(t)}\Big)^3 \cos(2x - 2f(t)) \,, \tag{7}$$

where $y = \dot{x}$. Due to the assumptions of the model, the quantities r and f are known functions of the time, being the solution of Kepler's problem which determines the elliptical orbit of the satellite. They depend on the orbital eccentricity, which plays the role of an additional parameter.

It is important to observe that:

- the Hamiltonian (7) is integrable whenever $\varepsilon = 0$, namely the satellite has equatorial symmetry with $I_1 = I_2$;
- the Hamiltonian (7) is integrable when the eccentricity is equal to zero, since the orbit becomes circular, namely $r = a$ and f coincides with the mean anomaly, which is proportional to time.

The existence and breakdown of invariant tori in the conservative spin-orbit problem have been investigated in [33, 34].

We remark that Hamilton's equations associated to (7) are given by

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\varepsilon\Big(\frac{a}{r(t)}\Big)^3 \sin(2x - 2f(t)) \,. \end{aligned} \tag{8}$$

Integrating (8) with a modified Euler's method with time-step h, we obtain a discrete system which retains the many features of the conservative standard map when taking the solution on the Poincaré map at time intervals multiple of 2π:

$$\begin{aligned} y_{n+1} &= y_n - \varepsilon\Big(\frac{a}{r_n}\Big)^3 \sin(2x_n - 2f_n)\, h \\ x_{n+1} &= x_n + y_{n+1}\, h \\ t_{n+1} &= t_n + h \,, \end{aligned}$$

where $r_n = r(t_n)$, $f_n = f(t_n)$.

The dissipative spin-orbit problem is obtained by taking into account that the satellite is non rigid and therefore it is subject to a tidal torque. The equation of motion including a model for the tidal torque can be written as

$$\ddot{x} + \varepsilon \left(\frac{a}{r}\right)^3 \sin(2x - 2f) = -K_d \left[L(e,t)\dot{x} - N(e,t)\right], \tag{9}$$

where

$$L(e,t) = \frac{a^6}{r^6}, \qquad N(e,t) = \frac{a^6}{r^6}\,\dot{f}$$

(see, e.g., [35, 93]). The coefficient K_d is called the dissipative constant, and depends on the physical and orbital features of the body:

$$K_d = 3n\frac{k_2}{\xi Q}\left(\frac{R_e}{a}\right)^3 \frac{M}{m},$$

where n is the mean motion, k_2 is the so-called Love number (depending on the structure of the satellite), Q is called the quality factor (it compares the frequency of oscillation of the system to the rate of dissipation of the energy), ξ is a structure constant such that $I_3 = \xi m R_e^2$ with R_e the equatorial radius, M is the mass of the planet, m is the mass of the satellite. For bodies like the Moon or Mercury, realistic values are $\varepsilon = 10^{-4}$ and $K_d = 10^{-8}$.

The expression for the tidal torque can be simplified by assuming (as, e.g., in [42]) that the dynamics is ruled by the averages of $L(e,t)$ and $N(e,t)$ over one orbital period. The averaged quantities are given by

$$\overline{L}(e) = \frac{1}{(1-e^2)^{\frac{9}{2}}}\left(1 + 3e^2 + \frac{3}{8}e^4\right),$$
$$\overline{N}(e) = \frac{1}{(1-e^2)^6}\left(1 + \frac{15}{2}e^2 + \frac{45}{8}e^4 + \frac{5}{16}e^6\right).$$

Hence, we obtain the following equation of motion in the averaged case:

$$\ddot{x} + \varepsilon \left(\frac{a}{r}\right)^3 \sin(2x - 2f) = -K_d \left(\overline{L}(e)\dot{x} - \overline{N}(e)\right). \tag{10}$$

We can refer to the quantity $\lambda = K_d \overline{L}(e)$ as the dissipative parameter and to $\mu = \frac{\overline{N}(e)}{\overline{L}(e)}$ as the drift parameter.

Let us write (10) in canonical form as

$$\dot{x} = y$$
$$\dot{y} = -\varepsilon\left(\frac{a}{r(t)}\right)^3 \sin(2x - 2f(t)) - \lambda(y - \mu). \tag{11}$$

Similarly to the conservative case, the integration of (11) with a modified Euler's method with time-step h, leads to a discrete system similar to the dissipative standard map with dissipative and drift parameters, when taking the solution on the Poincaré map at time intervals multiple of 2π:

$$
\begin{aligned}
y_{n+1} &= (1-\lambda h) y_n + \lambda \mu h - \varepsilon \Big(\frac{a}{r_n}\Big)^3 \sin(2x_n - 2f_n)\, h \\
x_{n+1} &= x_n + y_{n+1}\, h \\
t_{n+1} &= t_n + h\ .
\end{aligned}
$$

As we will mention in Sect. 7, the existence and breakdown of invariant attractors in the dissipative spin-orbit problem have been studied in [18, 19, 28] through an application of KAM theory for conformally symplectic systems and through suitable numerical methods.

3 Conformally Symplectic Systems and Diophantine Vectors

In this Section we give the definition of conformally symplectic systems for maps and flows (see Sect. 3.1) and we introduce the set of Diophantine vectors for discrete and continuous systems (see Sect. 3.2).

3.1 Discrete and Continuous Conformally Symplectic Systems

An important class of dissipative dynamical systems is given by the *conformally symplectic systems;* the dissipative standard map is an example of a conformally symplectic discrete system, while the dissipative spin-orbit problem is an example of a conformally symplectic continuous system.

Before giving the formal definition, let us say that conformally symplectic systems are characterized by the property that the map or the flow transforms the symplectic form into a multiple of itself (see Definitions 1 and 2 below). Beside the examples mentioned before, we stress that conformally symplectic models can be found in different fields, e.g. the Euler-Lagrange equations of exponentially discounted systems ([6], typically found in finance, when inflation is present and one needs to minimize the cost in present money) or Gaussian thermostats ([51, 115], namely mechanical systems with forcing and a thermostating term based on the Gauss Least Constraint Principle for nonholonomic constraints).

Let us start to introduce the notion of $2n$-dimensional conformally symplectic maps. Let $\mathcal{M} = U \times \mathbb{T}^n$ be the phase space with $U \subseteq \mathbb{R}^n$ an open, simply connected domain with smooth boundary; the phase space $\mathcal{M}$ is endowed with the standard scalar product and a symplectic form Ω, represented by a matrix J at the point $\underline{z}$ acting on vectors $\underline{u}, \underline{v} \in \mathbb{R}^n$ as $\Omega(\underline{u}, \underline{v}) = (\underline{u}, J(\underline{z})\underline{v})$ with $(\cdot, \cdot)$ denoting the scalar product. Note that the matrix J depends not only on the symplectic form but on the metric considered.

Definition 1 A diffeomorphism f on $\mathcal{M}$ is *conformally symplectic, if there exists a function* $\lambda : \mathcal{M} \to \mathbb{R}$ such that, denoting by f^* the pull–back of f, we have:

$$f^*\Omega = \lambda\Omega \ . \tag{12}$$

We remark that for $n = 1$ any diffeomorphism is conformally symplectic with λ depending on the coordinates, namely one can take $\lambda(x) = det\,(Df\,(x))$ or $\lambda(x) = -det\,(Df\,(x))$. Instead, for $n \geq 2$ one obtains that λ is a constant. In fact, taking the exterior derivative of $f^*\Omega = \lambda\Omega$, one obtains:

$$d(f^*\Omega) = f^*\, d\Omega = 0 \ = \ d\lambda \wedge \Omega + \lambda \wedge d\Omega = d\lambda \wedge \Omega \ ,$$

that gives $d\lambda = 0$; since the manifold is connected, then λ is equal to a constant.

We also remark that for $\lambda = 1$ in (12) we recover the symplectic case.

Let us give some explicit examples which might help to clarify the meaning of Definition 1. First, we notice that we can re-formulate the notion of conformally symplectic by saying that the diffeomorphism f is conformally symplectic if

$$Df^T\ J\ Df = \lambda\ J \ , \tag{13}$$

where the superscript T denotes transposition. In fact, from (12) we have:

$$\begin{aligned} f^*\Omega = \lambda\Omega &\Leftrightarrow \Omega(Df\ \underline{u},\, Df\ \underline{v}) = \lambda\ \Omega(\underline{u},\, \underline{v}) \\ &\Leftrightarrow (Df\ \underline{u},\, J\ Df\ \underline{v}) = \lambda\ (\underline{u},\, J\ \underline{v}) \\ &\Leftrightarrow (u,\, Df^T\ J\ Df\ \underline{v}) = (\underline{u},\, \lambda\, J\ \underline{v}) \\ &\Leftrightarrow Df^T\ J\ Df = \lambda\ J \ . \end{aligned}$$

An example of a conformally symplectic diffeomorphism is given by the dissipative standard map. Recalling (3), we have that (13) is satisfied, as shown below:

$$\begin{pmatrix} \lambda & \lambda \\ \varepsilon V_x & 1 + \varepsilon V_x \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \lambda & \varepsilon V_x \\ \lambda & 1 + \varepsilon V_x \end{pmatrix} = \begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix} \ = \lambda\, J \ .$$

An example of a map which does not satisfy the conformally symplectic condition (13) is given by the following 4-dimensional dissipative standard map with conformal factors λ_1, λ_2 with $\lambda_1 \neq \lambda_2$:

$$\begin{aligned} y_1' &= \lambda_1 y_1 + \mu_1 + \varepsilon V_1(x_1, x_2) \\ y_2' &= \lambda_2 y_2 + \mu_2 + \varepsilon V_2(x_1, x_2) \\ x_1' &= x_1 + y_1' \\ x_2' &= x_2 + y_2' \ . \end{aligned}$$

In fact, even for $\varepsilon = 0$, we obtain that (13) is not satisfied:

$$Df^T\, J\, Df = \begin{pmatrix} 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \\ -\lambda_1 & 0 & 0 & 0 \\ 0 & -\lambda_2 & 0 & 0 \end{pmatrix} \neq \lambda \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \lambda\, J\ .$$

To conclude, we give the definition of conformally symplectic systems for continuous dynamical systems.

Definition 2 We say that a vector field X is a conformally symplectic flow if, denoting by L_X the Lie derivative, there exists a function $\lambda : \mathbb{R}^{2n} \to \mathbb{R}$ such that

$$L_X \Omega = \lambda \Omega\ .$$

In analogy to the definition of conformally symplectic maps, we remark that the time t-flow Φ_t satisfies the relation

$$(\Phi_t)^* \Omega = e^{\lambda t} \Omega\ .$$

3.2 Diophantine Vectors for Maps and Flows

In this section we give the definition of Diophantine vectors for maps and flows and we briefly recall the main properties of Diophantine vectors. We start by giving the definition for maps.

Definition 3 We say that the vector $\underline{\omega} \in \mathbb{R}^n$ satisfies the Diophantine condition, if for a constant $C > 0$ and an exponent $\tau > 0$, one has

$$\left| \frac{\underline{\omega} \cdot \underline{q}}{2\pi} - p \right|^{-1} \leq C |\underline{q}|^\tau\ , \qquad p \in \mathbb{Z}\ , \quad \underline{q} \in \mathbb{Z}^n \backslash \{0\}\ .$$

In the case of flows we have the following definition.

Definition 4 We say that the vector $\underline{\omega} \in \mathbb{R}^n$ satisfies the Diophantine condition, if for a Diophantine constant $C > 0$ and a Diophantine exponent $\tau > 0$, one has:

$$|\underline{\omega} \cdot \underline{k}|^{-1} \leq C |\underline{k}|^\tau\ , \qquad \underline{k} \in \mathbb{Z}^n \backslash \{0\}\ .$$

We conclude this section by listing below some important properties of Diophantine vectors.

(i) Let us denote by $\mathcal{D}(C, \tau)$ the set of Diophantine vectors satisfying Definition 4. Then, the size of the set of Diophantine vectors $\mathcal{D}(C, \tau)$ increases as C or τ increases. The set of vectors that satisfy this condition for some C, τ is of full Lebesgue measure in $\mathbb{R}^n$.

(ii) There are no Diophantine vectors in $\mathbb{R}^n$ with $\tau < n - 1$.

(iii) The set of Diophantine vectors with $\tau = n - 1$ in $\mathbb{R}^n$ has zero Lebesgue measure, but it is everywhere dense.

(iv) For $\tau > n - 1$, almost every vector in $\mathbb{R}^n$ is τ-Diophantine, namely the complement has zero Lebesgue measure, although it is everywhere dense.

4 Invariant Tori and KAM Theory for Conformally Symplectic Systems

In this Section we provide the definition of KAM (rotational) invariant tori (for maps and flows) (see Sect. 4.1); the statement of the KAM theorem for conformally symplectic maps is given in Sect. 4.2, whose proof is briefly recalled in Sect. 4.3. The proof can be translated into a very efficient KAM algorithm (see [21]), which is at the basis of different results: the derivation of numerical methods to compute the breakdown threshold (Sect. 5), the investigation of the breakdown mechanism (Sect. 6), the implementations to specific models (see Sect. 7).

4.1 Invariant KAM Tori

We start by giving the definition of *conditionally periodic and quasi-periodic* motions.

Definition 5 A conditionally periodic motion is represented by a function $t \mapsto f(\omega_1 t, \ldots, \omega_n t)$, where $f(x_1, \ldots, x_n)$ is periodic in all variables; the vector $\underline{\omega} = (\omega_1, \ldots, \omega_n)$ is called frequency.

A quasi-periodic motion is a conditionally periodic motion with incommensurable frequencies.

Next we give the following definition of *invariant torus.*

Definition 6 An invariant torus is a manifold diffeomorphic to the standard torus $\mathbb{T}^n$, that gets mapped into itself by the evolution.

We remark that any trajectory on an invariant torus carrying quasi-periodic motions is dense on the torus. We conclude by giving the definition of (rotational) KAM torus for maps and flows. This definition is based on the invariance equation (14) below, whose solution will be the centerpiece of the KAM theorem presented in Sect. 4.2.

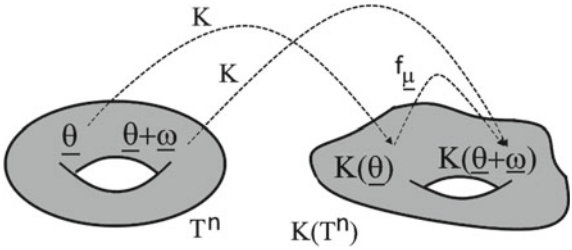


Fig. 6 Geometric interpretation of the invariance equation $f_{\underline{\mu}} \circ K(\underline{\theta}) = K(\underline{\theta} + \underline{\omega})$ in the unknowns $K, \underline{\mu}$

Definition 7 Let $\mathcal{M} \subseteq \mathbb{R}^n \times \mathbb{T}^n$ be a symplectic manifold and let $f : \mathcal{M} \to \mathcal{M}$ be a symplectic map. A KAM torus with frequency $\underline{\omega} \in \mathcal{D}(C, \tau)$ is an n–dimensional invariant torus described parametrically by an embedding $K : \mathbb{T}^n \to \mathcal{M}$, which is the solutions of the invariance equation:

$$f \circ K(\underline{\theta}) = K(\underline{\theta} + \underline{\omega}) \,. \tag{14}$$

For a family $f_{\underline{\mu}}$ of conformally symplectic diffeomorphisms depending on a real parameter vector $\underline{\mu}$, a KAM attractor with frequency $\underline{\omega} \in \mathcal{D}(C, \tau)$ is an n–dimensional invariant torus described parametrically by an embedding $K : \mathbb{T}^n \to \mathcal{M}$ and a drift $\underline{\mu}$, which are the solutions of the invariance equation:

$$f_{\underline{\mu}} \circ K(\underline{\theta}) = K(\underline{\theta} + \underline{\omega}) \,. \tag{15}$$

For conformally symplectic vector fields $X_{\underline{\mu}}$, the invariance equation is given by

$$X_{\underline{\mu}} \circ K(\underline{\theta}) = (\underline{\omega} \cdot \partial_{\underline{\theta}})\, K(\underline{\theta}) \,.$$

We remark that for symplectic systems the invariance equation (14) contains as only unknown the embedding K, while for conformally symplectic systems the invariance equation (15) contains as unknowns both the embedding K and the drift term $\underline{\mu}$.

A graphical representation of the invariance equation (14) is given in Fig. 6.

Although the theory that will be presented in the next Sections apply both to maps and flows, for simplicity of exposition we will limit to the presentation of KAM theory for maps. We refer to [21] for the details concerning continuous systems.

4.2 Conformally Symplectic KAM Theorem

We will try to answer a specific question, which is formulated below, by means of a suitable statement of the KAM theorem; the question is motivated by many applications in several models of Celestial Mechanics, which are often described by nearly-integrable systems. This is why we set the following question in the framework

of nearly-integrable systems, although the formulation of KAM theory does not need that the system is close to integrable (compare with [44]).

Assume that a given integrable dynamical system admits an invariant torus run by a quasi-periodic motion with frequency $\underline{\omega}$ (e.g., think at Kepler's 2-body problem). Consider a perturbation of the integrable system (e.g., the restricted 3-body problem, which is described by the 2-body problem with a perturbation proportional to the primaries' mass ratio). The main question that we want to raise in the framework of KAM theory for nearly-integrable systems is the following: does the perturbed system still admits an invariant torus run by a quasi-periodic motion with the same frequency as the unperturbed system? The answer is given by the celebrated KAM theory ([2, 78, 90]), which can be implemented under very general assumptions, precisely a non-degeneracy condition of the unperturbed system and a Diophantine condition on the frequency.

We remark that invariant tori are *Lagrangian:* if f is a symplectic map and K satisfies the invariance Eq. (14), then

$$K^* \Omega = 0 \, .$$

The same holds for a conformally symplectic map f_μ, when $|\lambda| \neq 1$ and K satisfies the invariance Eq. (15). If f is symplectic and $\underline{\omega}$ is irrational, then the torus is Lagrangian, i.e. with maximal dimension and isotropic (namely, the symplectic form restricted on the manifold is zero, which implies that each tangent space is an isotropic subspace of the ambient manifold's tangent space).

Next step is to consider a nearly-integrable dynamical system affected by a dissipative force, so that the overall system is conformally symplectic (an example is given by the spin-orbit problem with tidal torque). We assume that the integrable symplectic system admits an invariant torus with Diophantine frequency; the question becomes whether the non-integrable system with dissipation still admits, for suitable values of the drift parameter, an invariant attractor run by a quasi-periodic motion with the same frequency of the unperturbed system. The answer is given by the KAM theorem for conformally symplectic systems as given by Theorem 1 (see [21]).

Since we will be interested to give explicit estimates in specific model problems, we introduce the following norms for analytic and differentiable functions.

Definition 8 *Analytic norm.* Given $\rho > 0$, we define the complex extension of the torus, say $\mathbb{T}^n_\rho$, as the set

$$\mathbb{T}^n_\rho = \{\underline{\theta} \in \mathbb{C}^n/(2\pi\mathbb{Z})^n : \ \mathrm{Re}(\underline{\theta}) \in \mathbb{T}^n, \ |\mathrm{Im}(\theta_j)| \leq \rho \, , \quad j = 1, ..., n\} \, ;$$

we denote by $\mathcal{A}_\rho$ the set of analytic functions in $Int(\mathbb{T}^n_\rho)$ with the norm

$$\| f \|_\rho = \sup_{\underline{\theta} \in \mathbb{T}^n_\rho} |f(\underline{\theta})| \, .$$

Sobolev norm. For a function $f = f(\underline{z})$ expanded in Fourier series as $f(\underline{z}) = \sum_{\underline{k}\in\mathbb{Z}^n} \widehat{f}_{\underline{k}}\, e^{2\pi i \underline{k}\cdot\underline{z}}$ for an integer $m > 0$, we define the space H^m as

$$H^m = \left\{ f : \mathbb{T}^n \to \mathbb{C} \ : \ \|f\|_m^2 \equiv \sum_{\underline{k}\in\mathbb{Z}^n} |\, \widehat{f}_{\underline{k}}\, |^2 (1 + |\underline{k}|^2)^m < \infty \right\} .$$

Borrowing the statement from [21], we give below the formulation of the KAM theorem for conformally symplectic systems (see [44] for the statement for symplectic systems). We give the statement for maps, although the results can be formulated also for systems with continuous time (flows). Indeed in [21] one can find a construction that shows that the results for maps imply the results for flows as well as direct proofs of the results for flows.

Theorem 1 *Let $\underline{\omega} \in \mathcal{D}(C, \tau)$, $f_{\underline{\mu}} : \mathbb{R}^n \times \mathbb{T}^n \to \mathbb{R}^n \times \mathbb{T}^n$ be a conformally symplectic diffeomorphism, and let $(K, \underline{\mu})$ be an approximate solution of the invariance equation (15) with error term E:*

$$f_{\underline{\mu}} \circ K(\underline{\theta}) - K(\underline{\theta} + \underline{\omega}) = E(\underline{\theta}) .$$

Let N be the quantity

$$N(\underline{\theta}) = (DK(\underline{\theta})^T DK(\underline{\theta}))^{-1} \tag{16}$$

and let $M(\underline{\theta})$ be the $2n \times 2n$ matrix defined by

$$M(\underline{\theta}) = [DK(\underline{\theta}) \mid J(K(\underline{\theta}))^{-1}\, DK(\underline{\theta}) N(\underline{\theta})] .$$

Let $P(\underline{\theta})$ be defined as

$$P(\underline{\theta}) \equiv DK(\underline{\theta}) N(\underline{\theta}) ;$$

let $A(\underline{\theta}) \equiv \lambda\, \mathrm{Id}$ and let $S(\underline{\theta})$ be

$$S(\underline{\theta}) \equiv P(\underline{\theta} + \underline{\omega})^T Df_{\underline{\mu}} \circ K(\underline{\theta}) J^{-1} \circ K(\underline{\theta}) P(\underline{\theta}) - N(\underline{\theta} + \underline{\omega})^T \gamma(\underline{\theta} + \underline{\omega}) N(\underline{\theta} + \underline{\omega}) A(\underline{\theta}) \tag{17}$$

with

$$\gamma(\underline{\theta}) \equiv DK(\underline{\theta})^T J^{-1} \circ K(\underline{\theta}) DK(\underline{\theta}) .$$

Assume that the following non–degeneracy condition is satisfied:

$$\det \begin{pmatrix} \langle S \rangle & \langle SB^0 \rangle + \langle \widetilde{A}_1 \rangle \\ (\lambda - 1)\mathrm{Id} & \langle \widetilde{A}_2 \rangle \end{pmatrix} \neq 0 \tag{18}$$

with $\widetilde{A}_1$, $\widetilde{A}_2$ the first and second n columns of $\tilde{A} = M^{-1}(\underline{\theta} + \underline{\omega}) D_{\underline{\mu}} f_{\underline{\mu}} \circ K$, $B^0 = B - \langle B \rangle$ is the solution of $\lambda B^0(\underline{\theta}) - B^0(\underline{\theta} + \underline{\omega}) = -(\widetilde{A}_2)^0(\underline{\theta})$.

For $\rho > 0$, let $0 < \delta < \frac{\rho}{2}$; if the solution is sufficiently approximate, namely

$$\|E\|_\rho \leq C_3\, C^{-4}\, \delta^{4\tau}$$

for a suitable constant $C_3 > 0$, then there exists an exact solution $(K_e, \underline{\mu}_e)$, such that

$$\|K_e - K\|_{\rho-2\delta} \leq C_4\, C^2\, \delta^{-2\tau}\, \|E\|_\rho\,, \quad |\underline{\mu}_e - \underline{\mu}| \leq C_5\, \|E\|_\rho$$

with suitable constants $C_4, C_5 > 0$.

Remark 1 It is useful to make a remark on the non–degeneracy condition (18), when applied to the conservative and dissipative standard maps ((1) and (3)). For the conservative standard map, the non–degeneracy condition is typically the so-called *twist* condition, which can be written as

$$\frac{\partial x'}{\partial y} \neq 0\,, \tag{19}$$

implying that the lift of the map transforms any vertical line always on the same side.

Instead, for the dissipative standard map, that we modify adding a generic dependence on the drift through a function $p = p(\mu)$, $\mu \in \mathbb{R}$, say

$$\begin{aligned} y' &= \lambda y + p(\mu) + \varepsilon\, V(x) \\ x' &= x + y'\,, \end{aligned}$$

then the non–degeneracy condition involves the twist condition and a non–degeneracy condition with respect to to the parameters, namely:

$$\frac{\partial x'}{\partial y} \neq 0\,, \qquad \frac{dp(\mu)}{d\mu} \neq 0\,. \tag{20}$$

We remark, however, that (19) and (20) involve global properties of the system, while (18) is a condition involving just the approximate solution, so that (18) may be applied in situations where (19) and (20) fail.

The proof of Theorem 1 is given in [21] through the a-posteriori approach developed in [44] and making use of an adjustment of parameters (see [9, 92]): assume we can find an approximate solution $(K, \underline{\mu})$ of the invariance equation, satisfying a non-degeneracy condition, then we can find a true solution $(K_e, \underline{\mu}_e)$ close to $(K, \underline{\mu})$, such that $\|K_e - K\|$, $|\underline{\mu}_e - \underline{\mu}|$ is small. A sketch of the proof of Theorem 1 is presented in Sect. 4.3.

We conclude this Section by remarking that the a-posteriori approach presents several advantages, among which:

- (i) it can be developed in any coordinate frame and not necessarily in action-angle variables. In many practical problems, the action-angle variables are difficult to compute and involve complex singularities.
  Of course, once we have the existence of the torus, we can construct action angle variables. Hence, compared to more standard results, the accomplishment of the method is that the existence of action variables and the quasi-integrability is moved from the hypothesis to the conclusions. This is useful in practice since the hypothesis is hard to verify in applications.
- (ii) The system is not assumed to be nearly integrable.
- (iii) Instead of constructing a sequence of coordinate transformations on shrinking domains as in the perturbation approach, one computes suitable corrections to the embedding and the drift.
  The computation of the embeddings requires to work only with variables of n dimensions whereas transformation theory requires to work with variables in $2n$ dimensions. The complexity of representing functions grows exponentially – with a large exponent – with the dimension. The composition of two functions has rather awkward analytic and numerical properties.
- (iv) The non-degeneracy assumptions are not global properties of the map, but are rather properties of the considered approximate solution.
- (v) One does not need to justify how the approximate solution was obtained. In particular, one can take as approximate solution the result of numerical calculations or a formal expansion.
  Verifying the hypothesis in a numerical approximation requires just a finite number of calculations. Even if this number is too large to do by hand, it could be moderate to do with a computer (e.g., a few hours on a common laptop). If these can be done taking care of roundoff and truncation errors, this may lead to a computer assisted proof.
  One can also verify the hypothesis easily in a numerical expansion.

4.3 A Sketch of the Proof of the KAM Theorem

The proof of Theorem 1 can be summarized as composed by five main steps:

- Step 1: starting from an approximate solution, write the linearization of the invariance equation.
- Step 2: by a Newton's method find a quadratically smaller approximation.
- Step 3: under a non–degeneracy condition, solve the cohomological equation that allows to find the new approximation.
- Step 4: iterate the procedure and show its convergence.
- Step 5: prove that the solution is locally unique.

We briefly describe such steps as follows.

4.3.1 Step 1: Approximate Solution and Linearization

Let $(K, \underline{\mu})$ be an approximate solution satisfying

$$f_{\underline{\mu}} \circ K(\underline{\theta}) - K(\underline{\theta} + \underline{\omega}) = E(\underline{\theta}) \,. \tag{21}$$

Using the Lagrangian property $K^*\Omega = 0$ written in coordinates, namely

$$DK^T(\underline{\theta})\ J \circ K(\underline{\theta})\ DK(\underline{\theta}) = 0 \,,$$

we get that the tangent space is given by

$$Range\Big(DK(\underline{\theta})\Big) \oplus Range\Big(V(\underline{\theta})\Big) \tag{22}$$

with N as in (16) and

$$V(\underline{\theta}) = J^{-1} \circ K(\underline{\theta})\ DK(\underline{\theta}) N(\underline{\theta}) \,.$$

Define the quantity

$$M(\underline{\theta}) = [DK(\underline{\theta}) \mid V(\underline{\theta})] \,. \tag{23}$$

Then, we have the following result.

Lemma 1 *Up to a remainder R, we have the following relation:*

$$Df_{\underline{\mu}} \circ K(\underline{\theta})\ M(\underline{\theta}) = M(\underline{\theta} + \underline{\omega}) \begin{pmatrix} \mathrm{Id} & S(\underline{\theta}) \\ 0 & \lambda \mathrm{Id} \end{pmatrix} + R(\underline{\theta}) \,.$$

Proof Recalling the definition of M in (23), we have that taking the derivative of

$$f_{\underline{\mu}} \circ K(\underline{\theta}) = K(\underline{\theta} + \underline{\omega}) + E(\underline{\theta}) \,,$$

one obtains the relation

$$Df_{\underline{\mu}} \circ K(\underline{\theta})\ DK(\underline{\theta}) = DK(\underline{\theta} + \underline{\omega}) + DE(\underline{\theta}) \,.$$

Due to (22), one obtains:

$$Df_{\underline{\mu}} \circ K(\underline{\theta})\ V(\underline{\theta}) = DK(\underline{\theta} + \underline{\omega})\ S(\underline{\theta}) + V(\underline{\theta} + \underline{\omega})\ \lambda\ \mathrm{Id} + h.o.t.$$

with S as in (17).

4.3.2 Step 2: Determine a New Approximation.

Let the new approximation $(K', \underline{\mu}')$ be defined as $K' = K + MW$, $\underline{\mu}' = \underline{\mu} + \underline{\sigma}$. Let E' be the error associated to $(K', \underline{\mu}')$:

$$f_{\underline{\mu}'} \circ K'(\underline{\theta}) - K'(\underline{\theta} + \underline{\omega}) = E'(\underline{\theta}) \,. \tag{24}$$

Expanding (24) in Taylor series, we get

$$\begin{aligned} &f_{\underline{\mu}} \circ K(\underline{\theta}) + Df_{\underline{\mu}} \circ K(\underline{\theta})\, M(\underline{\theta}) W(\underline{\theta}) + D_{\underline{\mu}} f_{\underline{\mu}} \circ K(\underline{\theta}) \underline{\sigma} \\ &\quad - K(\underline{\theta} + \underline{\omega}) - M(\underline{\theta} + \underline{\omega})\, W(\underline{\theta} + \underline{\omega}) + h.o.t. = E'(\underline{\theta}) \,. \end{aligned}$$

Recalling (21), the new error E' is quadratically smaller provided the following relation holds:

$$Df_{\underline{\mu}} \circ K(\underline{\theta})\, M(\underline{\theta}) W(\underline{\theta}) - M(\underline{\theta} + \underline{\omega})\, W(\underline{\theta} + \underline{\omega}) + D_{\underline{\mu}} f_{\underline{\mu}} \circ K(\underline{\theta}) \underline{\sigma} = -E(\underline{\theta}) \,. \tag{25}$$

Combining (25) and Lemma 1, we have:

$$Df_{\underline{\mu}} \circ K(\underline{\theta})\, M(\underline{\theta}) = M(\underline{\theta} + \underline{\omega}) \begin{pmatrix} \mathrm{Id} & S(\underline{\theta}) \\ 0 & \lambda \mathrm{Id} \end{pmatrix} + R(\underline{\theta}) \,.$$

This allows to get the following equations for $W = (W_1, W_2)$ and $\underline{\sigma}$

$$M(\underline{\theta} + \underline{\omega}) \begin{pmatrix} \mathrm{Id} & S(\underline{\theta}) \\ 0 & \lambda \mathrm{Id} \end{pmatrix} W(\underline{\theta}) - M(\underline{\theta} + \underline{\omega})\, W(\underline{\theta} + \underline{\omega}) = -E(\underline{\theta}) - D_{\underline{\mu}} f_{\underline{\mu}} \circ K(\underline{\theta}) \underline{\sigma}$$

that we are going to make more explicit. Multiplying by $M(\underline{\theta} + \underline{\omega})^{-1}$ and writing $W = (W_1, W_2)$, one gets that the previous equation is equivalent to:

$$\begin{pmatrix} \mathrm{Id} & S(\underline{\theta}) \\ 0 & \lambda \mathrm{Id} \end{pmatrix} \begin{pmatrix} W_1(\underline{\theta}) \\ W_2(\underline{\theta}) \end{pmatrix} - \begin{pmatrix} W_1(\underline{\theta} + \underline{\omega}) \\ W_2(\underline{\theta} + \underline{\omega}) \end{pmatrix} = \begin{pmatrix} -\tilde{E}_1(\underline{\theta}) - \tilde{A}_1(\underline{\theta}) \underline{\sigma} \\ -\tilde{E}_2(\underline{\theta}) - \tilde{A}_2(\underline{\theta}) \underline{\sigma} \end{pmatrix} \tag{26}$$

with $\tilde{E}_j(\underline{\theta}) = -(M(\underline{\theta} + \underline{\omega})^{-1} E)_j$, $\tilde{A}_j(\underline{\theta}) = (M(\underline{\theta} + \underline{\omega})^{-1} D_{\underline{\mu}} f_{\underline{\mu}} \circ K)_j$. Writing (26) in components, we obtain:

$$\begin{aligned} W_1(\underline{\theta}) - W_1(\underline{\theta} + \underline{\omega}) &= -\widetilde{E}_1(\underline{\theta}) - S(\underline{\theta}) W_2(\underline{\theta}) - \widetilde{A}_1(\underline{\theta})\, \underline{\sigma} \\ \lambda W_2(\underline{\theta}) - W_2(\underline{\theta} + \underline{\omega}) &= -\widetilde{E}_2(\underline{\theta}) - \widetilde{A}_2(\underline{\theta})\, \underline{\sigma} \,. \end{aligned} \tag{27}$$

The cohomological equations (27) allow to find the corrections W_1, W_2 and $\underline{\sigma}$, as sketched in the next step.

4.3.3 Step 3: Solve the Cohomological Equations

To determine the new approximation, we need to solve equations (27), which are equations with constant coefficients for W_1, W_2 and σ for known S, $\widetilde{E} \equiv (\widetilde{E}_1, \widetilde{E}_2)$, $\widetilde{A} \equiv [\widetilde{A}_1 | \widetilde{A}_2]$.

The first equation in (27) is a standard small divisor equation, which can be solved under the Diophantine condition on the frequency, so to bound the small divisors.

For $|\lambda| \neq 1$ and for all real vectors $\underline{\omega}$, it is possible to solve the second equation in (27) by an elementary contraction mapping argument.

We remark that, using Cauchy estimates for the cohomological equations (27), we can bound $\|W_1\|_{\rho-\delta}$ and $\|W_2\|_{\rho-\delta}$ by $\|E\|_\rho$.

To solve the cohomological equations, we proceed as follows. Take the averages of each equation in (27) and use the non–degeneracy condition to determine $\langle W_2 \rangle$, $\underline{\sigma}$ by solving the equation

$$\begin{pmatrix} \langle S \rangle & \langle SB^0 \rangle + \langle \widetilde{A}_1 \rangle \\ (\lambda - 1)\mathrm{Id} & \langle \widetilde{A}_2 \rangle \end{pmatrix} \begin{pmatrix} \langle W_2 \rangle \\ \underline{\sigma} \end{pmatrix} = \begin{pmatrix} -\langle S\tilde{B}^0 \rangle - \langle \widetilde{E}_1 \rangle \\ -\langle \widetilde{E}_2 \rangle \end{pmatrix} ,$$

where we have split W_2 as $W_2 = \langle W_2 \rangle + B^0 + \tilde{B}^0 \underline{\sigma}$.

Next, we need to solve the second equation in (27) for W_2, which is an equation of the form $\lambda W_2(\underline{\theta}) - W_2(\underline{\theta} + \underline{\omega}) = Q_2(\underline{\theta})$ with Q_2 known. Such equation is always solvable for any $|\lambda| \neq 1$ by a contraction mapping argument, using that $\lambda W_2(\underline{\theta}) - W_2(\underline{\theta} + \underline{\omega}) = \sum_{\underline{k}} \widehat{W}_{2,\underline{k}}\, e^{i\underline{k}\cdot\underline{\theta}}(\lambda - e^{i\underline{k}\cdot\underline{\omega}})$.

Finally, we solve the first equation in (27) for W_1, which amounts to solving an equation of the form $W_1(\underline{\theta}) - W_1(\underline{\theta} + \underline{\omega}) = Q_1(\underline{\theta})$ with Q_1 known. It involves small (zero) divisors, since for $\underline{k} = \underline{0}$ one has $1 - e^{i\underline{k}\cdot\underline{\omega}} = 0$. The left hand side of the first equation in (27) can be expanded as

$$W_1(\underline{\theta}) - W_1(\underline{\theta} + \underline{\omega}) = \sum_{\underline{k} \in \mathbb{Z}^n \setminus \{\underline{0}\}} \widehat{W}_{1,\underline{k}}\, e^{i\underline{k}\cdot\underline{\theta}}(1 - e^{i\underline{k}\cdot\underline{\omega}}) .$$

To get a bound for the solution of (27), we need the following result.

Proposition 1 *Let $Z = Z(\underline{\theta})$ be a function with zero average and such that $Z \in \mathcal{A}_\rho$ or $Z \in H^m$. Let $\underline{\omega} \in D(C, \tau)$. Assume that the function $U = U(\underline{\theta})$ satisfies*

$$\lambda U(\underline{\theta}) - U(\underline{\theta} + \underline{\omega}) = Z(\underline{\theta}) .$$

Then, if $\lambda \neq 1$, $|\lambda| \in [A, A^{-1}]$ for $0 < A < 1$, we have that

$$\|U(\underline{\theta})\|_{\rho-\delta} \leq C\delta^{-\tau}\|Z\|_\rho .$$

We refer to [21, 102] for the proof of Proposition 1.

4.3.4 Step 4: Convergence of the Iterative Step.

The solution described in Step 3, allows to state that the invariance equation is satisfied with an error quadratically smaller, i.e.

$$\|E'\|_{\rho-\delta} \leq C_8 \delta^{-2\tau} \|E\|_{\rho}^2 .$$

The procedure at Step 3 can be iterated to get a sequence of approximate solutions, say $\{K_j, \underline{\mu}_j\}$. Its convergence is obtained through an abstract implicit function theorem, alternating the iteration with carefully chosen smoothing operators defined in a scale of Banach spaces (which can be either analytic functions or Sobolev spaces).

4.3.5 Step 5: Local Uniqueness

Under smallness conditions, one can prove that, if there exist two solutions $(K_a, \underline{\mu}_a)$, $(K_b, \underline{\mu}_b)$, then there exists $\underline{\psi} \in \mathbb{R}^n$ such that

$$K_b(\underline{\theta}) = K_a(\underline{\theta} + \underline{\psi}) \qquad \text{and} \qquad \underline{\mu}_a = \underline{\mu}_b .$$

We remark that in the analytic case, the smoothing is obtained by rescaling the size of the strip on which the analytic functions are defined at each step, given that the domains where they are defined shrink by a given amount. Then, for the sequence of solutions $\{K_j, \underline{\mu}_j\}$, one can take the analyticity domain parameters ρ_h and the shrinking parameters δ_h as

$$\rho_0 = \rho , \qquad \delta_h = \frac{\rho_0}{2^{h+2}} , \qquad \rho_{h+1} = \rho_h - \delta_h , \qquad h \geq 0 .$$

Given that the error is quadratic, we can write for some $a, b > 0$ and a constant $C_E > 0$:

$$\|E(K_{h+1}, \underline{\mu}_{h+1})\|_{\rho_{h+1}} \leq C_E \, \nu^a \delta_h^b \, \|E(K_h, \underline{\mu}_h)\|_{\rho_h}^2 .$$

If the quantity $\varepsilon_0 \equiv \|E(K_0, \underline{\mu}_0)\|_{\rho_0}$ is small enough, then one can prove that

$$\|K_h - K_0\|_{\rho_h} \leq C_K \varepsilon_0 , \qquad |\underline{\mu}_h - \underline{\mu}_0| \leq C_\mu \varepsilon_0$$

for some constants $C_K, C_\mu > 0$. A finite number of conditions on parameters and norms will imply the indefinite iterability of the procedure and its convergence.

The a-posteriori approach for conformally symplectic systems has a number of consequences and further developments that we briefly summarize below, referring to the cited literature for full details:

- the method provides an efficient algorithm to determine the breakdown threshold, very suitable for computer implementations [15];

- the a-posteriori method allows to find rigorous lower estimates of the breakdown threshold [56, 97]. The rigorous lower estimates for symplectic maps in [56, 97] are very close to to the rigorous upper estimates in [67]. In [26] one can find very detailed estimates (they do not control completely the round off error, but they control everything else), that are comparable with the best numerical estimates computed by other methods;
- one gets that the local behavior near quasi–periodic solutions is given by a rotation in the angles and a shrink in the actions [22];
- the method allows to obtain a partial justification of Greene's criterion for the computation of the breakdown threshold of invariant attractors [27];
- one obtains a bootstrap of regularity, which allows to state that all smooth enough tori are analytic, whenever the map is analytic [21];
- one gets a characterization of the analyticity domains of the quasi–periodic attractors in the symplectic limit [23];
- one can prove the existence of whiskered tori for conformally symplectic systems [24].

Concerning the first item above, we stress that the proof given in [21] leads to a very efficient KAM algorithm, which can be implemented numerically and it is shown to work very close to the boundary of validity [26]. Indeed, all steps of the algorithm involve diagonal operations in the Fourier space and/or diagonal operations in the real space. Moreover, if we represent a function in discrete points or in Fourier space, then we can compute the other functions by applying the Fast Fourier Transform (FFT). Using N Fourier modes to discretize the function, then we need $O(N)$ storage and $O(N \log N)$ arithmetic operations. Note that all the steps in the algorithm can be implemented in a few lines in a high level language so that the resulting algorithm is not very hard to implement (about 200 lines in `Octave`, see [52], and about 2000 lines in C). Even if the above transcription of the algorithm works extremely well in near integrable systems, when approaching the breakdown, one needs to take some standard precautions (e.g. monitoring the size of the tails of Fourier series).

We also remark that the KAM proof requires a computer to make very long computations, which are needed to determine, for example, the initial approximate solution or to check the KAM algorithm. However, the computer introduces rounding-off and propagation errors, which can be controlled through interval arithmetic for which we refer to the specialized literature (see, e.g., [61, 74, 80, 89]).

5 Breakdown of Quasi–periodic Tori and Quasi–periodic Attractors

The analytical estimates which can be obtained through the implementation of the KAM theorem represent a rigorous lower bound of the breakdown threshold of invariant tori. In problems with a well-defined physical meaning, one can compare the KAM results with a measure of the parameter(s). For example in the restricted

3-body problem, one aims to prove the theorem for the true value of the mass ratio of the primaries. If we consider an asteroid under the gravitational attraction of Jupiter and the Sun, then the mass ratio amounts to $\varepsilon \simeq 10^{-3}$, which represents the benchmark that one wants to reach through rigorous KAM estimates.

Model problems like the standard maps do not have a physical reference value; therefore, one needs to apply numerical techniques that allow to determine the KAM breakdown threshold. Among the others, we mention Greene's technique [59], frequency analysis [81], Sobolev's method [15].

In the next Sections we review two methods for the numerical computation of the breakdown threshold that have been successfully applied to the standard map [15, 16, 59]: one is based on Sobolev's method (Sect. 5.1) and the other is based on Greene's method (Sect. 5.2). The problem of breakdown of KAM tori has been studied by many methods. The paper [16] contains a small survery and comparison of several different methods, some of which we will not mention here.

5.1 Sobolev Breakdown Criterion

To illustrate the method, we focus on the specific examples of the conservative and dissipative standard maps; hence we have a two-dimensional discrete system, which can be parametrized by a one-dimensional variable $\theta \in \mathbb{T}$. In particular, in the conservative case we write the invariance equation for K as

$$f \circ K(\theta) = K(\theta + \omega) \, ,$$

while in the dissipative case we write the invariance equation for (K, μ) as

$$f_\mu \circ K(\theta) = K(\theta + \omega) \, . \tag{28}$$

As shown rigorously in [16] for the conservative case and in [15] for the dissipative case, the continuation method based in the constructive Newton method can (if given enough computer resources) reach arbitrarily close to the breakdown. Furthermore, the breakdown of analytic tori happens if and only if some Sobolev norm of sufficiently high order blows up.

This rigorous result can, of course, be readily implemented. Today's computers, of course, do not have infinite resources, but they are fairly impressive for people who started to work with a PDP-11 with 16K of RAM. Since the algorithms we describe are based on computing Fourier series, one can get readily the Sobolev norms of the embedding K and monitor their blow up.

The blow up of the Sobolev norm gives a clear indication that the torus is breaking down. Note that, given the a-posteriori theorem, and the bootstrap of regularity results, if the norm of the computed solution is not blowing up, it is a very clear indication that the torus is there.

Table 1 Breakdown values ε_{crit} of the golden mean curve obtained implementing Sobolev's method for the conservative case (left column) and for the dissipative case (right column), the latter one for two different values of the dissipative parameter

Conservative case		Dissipative case
ε_{crit}	λ	ε_{crit}
0.9716	0.9	0.9721
	0.5	0.9792

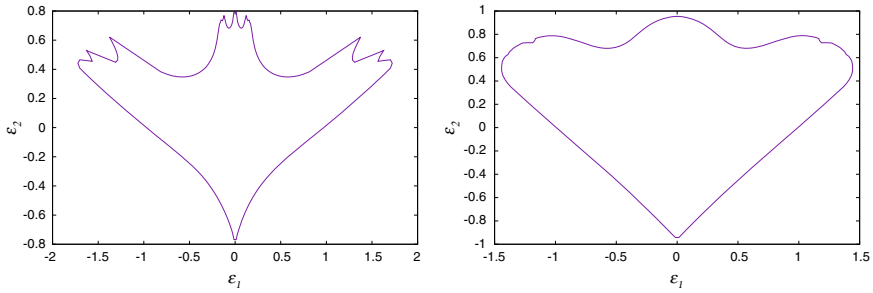


Fig. 7 Existence domain for invariant circles of the dissipative standard map with potential (29). Left: $\lambda = 0.9$. Right: $\lambda = 0.1$

Remark 2 Something that increases the possible effectiveness of this method is that it has been found empirically that the blow up of Sobolev norms is given by power laws whose exponents are universal. Even if this is mainly an empirical observation (that needs to be somehow tone down since [16] contains several warnings for some maps), it can improve dramatically the computation of breakdowns. Many of these empirical results are organized using *Renormalization Group methods* [88, 98–100]. Even if some aspects of renormalization group have been made rigorous [73, 75–77, 107, 108], much more mathematical work seems to remain.

We implement the method for the conservative and dissipative standard maps (1) and (3), computing in Table 1 the value of ε_{crit} for the frequency equal to the golden ratio: $\omega = 2\pi\frac{\sqrt{5}-1}{2}$. The result in the conservative case is in full agreement with the value which can be obtained by implementing Greene's method (see [59]). The values for the dissipative case given in Table 1 will be compared in Sect. 5.2 to those obtained implementing a version of Greene's method for the dissipative standard map.

In Fig. 7, we present the existence domain of the dissipative standard map (3) with a two harmonic potential given by

$$V(x) = \varepsilon_1 \sin(x) + \varepsilon_2 \sin(2x) \ . \tag{29}$$

We call attention to the fact that this region contains parts with smooth boundaries, but–specially in the conservative case–it contains some parts of the boundary that

are rather ragged. A tentative explanation ([85]) is that the smooth parts of the the boundary of the region of existence are the intersection of the family considered with the stable manifold of fixed point of renormalization. Even if this is not a completely rigorous picture, there has been significant mathematical progress in verifying it in an open set of families. We hope that, in the future, there could be more progress in this area.

One important advantage of the Sobolev method is that it can be programmed systematically and run unattended. The Greene's method relies on periodic orbits and one has to pay attention to making sure that the periodic orbits are continued correctly. We also note that the Sobolev method works for models of long range interaction in Statistical Mechanics without a dynamical interpretation.

5.2 *Greene's Method, Periodic Orbits and Arnold's Tongues*

The method by J. Greene, developed for the standard map in [59], is based on the conjecture that the breakdown of an invariant curve with frequency ω, say $\mathcal{C}(\omega)$, is related to a change from stability to instability of the periodic orbits $\mathcal{P}(\frac{p_j}{q_j})$ with frequencies $\frac{p_j}{q_j}$, $j = 1, 2, ...$, tending to ω. We observe that a standard procedure to obtain the rational approximants of ω is to compute the successive truncations of the continued fraction representation of ω.

Greene's method has been successfully developed for the conservative standard map for which a partial justification is given in [54, 85]. In the dissipative case, there appears an extra difficulty due to the fact that the periodic orbits with frequency $\frac{p_j}{q_j}$ occur in a whole interval of the drift parameter. This phenomenon gives rise to the appearance of the so-called *Arnold tongues*. Figure 8, left panel, gives a graphical representation of the Arnold tongues; having fixed a value of the dissipative parameter

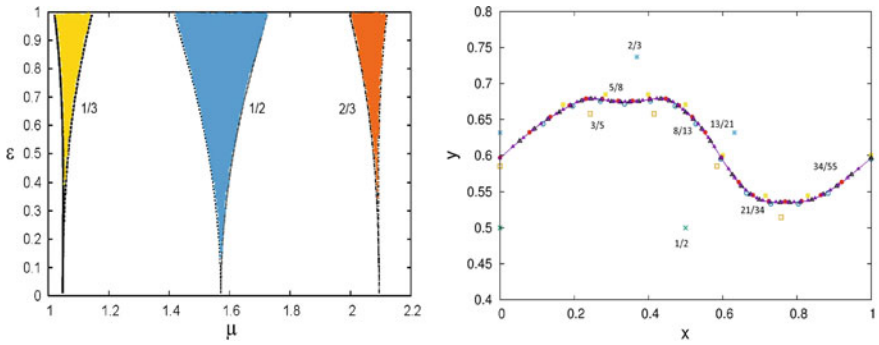


Fig. 8 Left: Arnold's tongues providing μ versus ε for three periodic orbits of the dissipative standard map with periods 1/3, 1/2, 2/3. Right: periodic orbits of the dissipative standard map approximating the golden mean curve

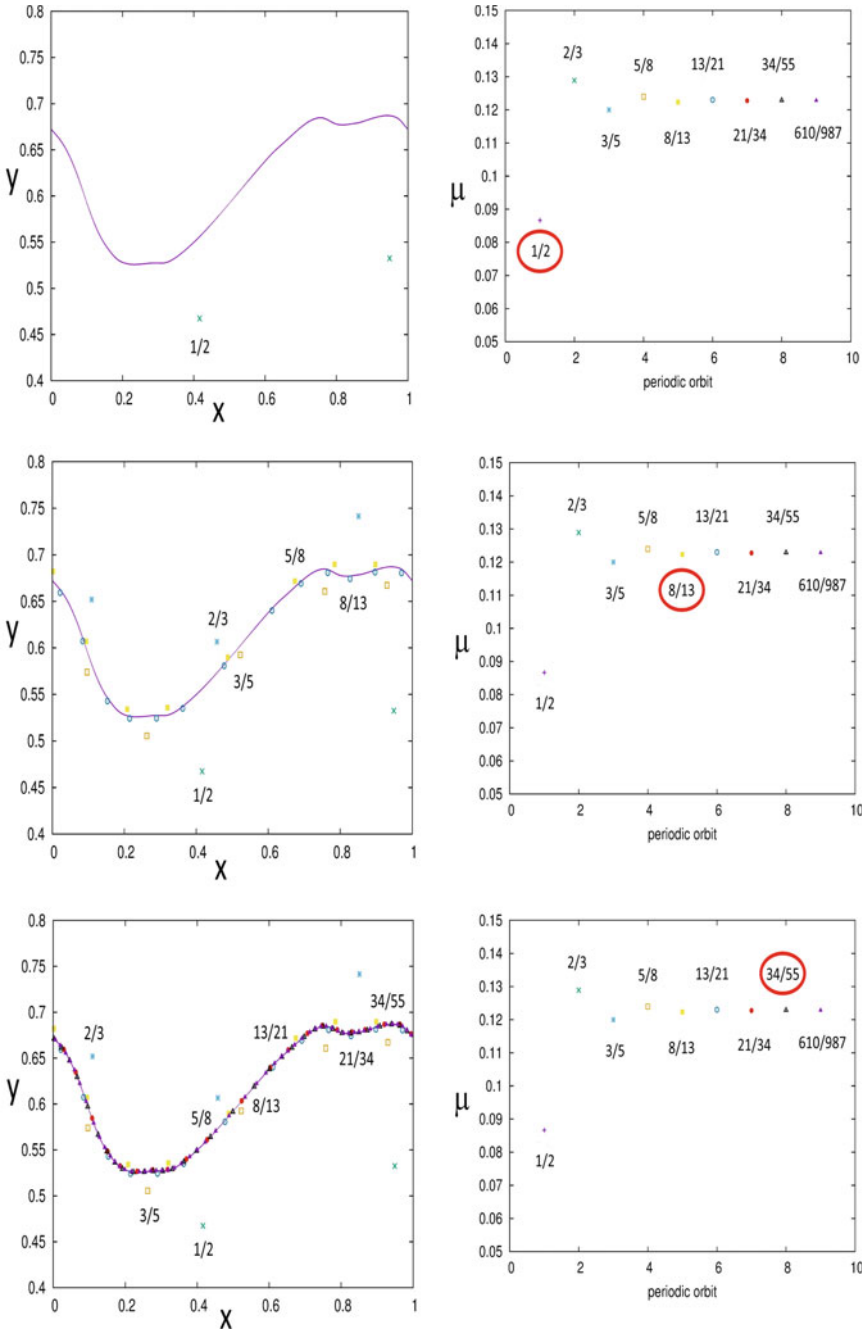


Fig. 9 Left: periodic orbits with increasing periods, approximating the golden mean curve. Right: the corresponding drift parameters with the successive periodic orbits labeled by integer numbers on the x-axis

ε, there is a whole interval of the drift parameter μ which admits a periodic orbit of the same period. The right panel of Fig. 8 shows several periodic orbits approaching the torus with frequency equal to the golden mean; such periodic orbits have frequency equal to the rational approximants which are given by the ratio of the Fibonacci numbers.

A partial justification of an extension of Greene's criterion in the conformally symplectic case is presented in [27], where it is proved that if there exists a smooth invariant attractor, one can predict the eigenvalues of the periodic orbits approximating the torus for parameters close to those of the attractor.

Figure 9 shows some approximating periodic orbits (left panels) and the corresponding behaviour of the drift parameter (right panels) that, in the limit, tends to the value of the drift that corresponds to the golden mean torus.

We also call attention to [37] which contains tentative results on the non-existence of invariant tori for the spin-orbit models. Even if the methods developed there are not rigorous, they may present a counterpoint to the methods to study the existence.

6 Collision of Invariant Bundles of Quasi-periodic Attractors

Quasi-periodic attractors of conformally symplectic maps are normally hyperbolic invariant manifolds (NHIM). As developed in [20] for the dissipative standard map, one can obtain the Lyapunov multipliers of the attractor from a simple computation. We start from the invariance Eq. (28) for a pair (K, μ). We then introduce a change of variables to reduce the cocycle. Let $\tilde{M}(\theta) = [DK(\theta) \mid E^s(\theta)]$, $\theta \in \mathbb{T}$, be the matrix whose columns are the tangent and stable bundles of $\mathcal{K} = K(\mathbb{T}^n)$:

$$Df_\mu \circ K(\theta)\tilde{M}(\theta) = \tilde{M}(\theta + \omega) \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} . \tag{30}$$

From Eq. (30) we can write the stable bundle as follows

$$E^s(\theta) = DK(\theta)B(\theta) + J^{-1}DK(\theta)N(\theta) ,$$

where $B(\theta)$ is the function that satisfies

$$B(\theta) - \lambda B(\theta + \omega) = -S(\theta) .$$

Indeed, after j iterates of the map we have that,

$$Df_\mu^j \circ K(\theta) = \tilde{M}(\theta + \omega) \begin{pmatrix} 1 & 0 \\ 0 & \lambda^j \end{pmatrix} \tilde{M}^{-1}(\theta) ,$$

which shows that the tangent space of $\mathcal{M}$ at $K(\theta)$ is

$$T_{K(\theta)}\mathcal{M} = \text{Range}(DK(\theta)) \oplus E^s_{K(\theta)} .$$

We can conclude that there exists a constant C such that

$$\begin{aligned} C^{-1}\lambda^j|v| \leq |Df_\mu^j \circ K(\underline{\theta})\, v| \leq C\lambda^j|v| , \qquad & v \in E^s_{K(\underline{\theta})} , \\ C^{-1}\,|v| \leq |Df_\mu^j \circ K(\underline{\theta})\, v| \leq C\,|v| , \qquad & v \in E^c_{K(\underline{\theta})} , \end{aligned}$$

showing that $\mathcal{K} = K(\mathbb{T}^n)$ is a NHIM. Equation (30) also tells us that the Lyapunov multipliers are constant along the family of quasi-periodic attractors for fixed Diophantine vectors.

In the case of maps of the cylinder $\mathcal{M} = \mathbb{R} \times \mathbb{T}$, we know that the curve $\mathcal{K}$ is C^r, one dimensional, and since ω satisfies the Diophantine condition, we know by the results of [66, 68, 69, 71] that the map conjugating the dynamics in $\mathcal{K}$ to a rigid rotation is in $C^{r-\tau-\delta}$ for a small $\delta > 0$. Therefore, by the bootstrap of regularity results[1], the conjugacy is analytic for analytic maps. Since the bundles depend on the conjugacy, then the regularity of the manifold implies the analyticity of K and the bundles up to the breakdown.

To investigate the breakdown of normal hyperbolicity, we note that, because of the pairing rule of Lyapunov exponents [51, 115], since one Lyapunov multiplier is 1—the one along the tangent directional (remember that the map on the torus is smoothly conjugate to the torus)—the other one is precisely λ.

We recall that hyperbolicity is equivalent to the existence of *transversal* invariant bundles with different rates. In our case, if the tori have to cease to be normally hyperbolic, because the exponents remain constant, the only thing that can happen is that the transversality of the bundles deteriorates.

What is found empirically is that the breakdown happens because at the same time the regularity of the conjugacy deteriorates quantitatively (even if the conjugacy remains analytic, some Sobolev norm blows up). See [20] for full details.

At the same parameter values, the breakdown of hyperbolicity happens via the stable and tangent bundle collision. Even if the Lyapunov exponents remain safely away, the transversality deteriorates and the tangent and stable bundles become close to tangent.

In the case at hand, we can make a very detailed study: the bundles are one dimensional and we compute a formula for the angle between the bundles for every θ. In fact, let $\alpha(\theta)$ be the angle between the stable and tangent bundles for every $\theta \in \mathbb{T}$, then we have

$$\alpha(\theta) = \arctan\left(\frac{1}{B(\theta)(DK(\theta)^T DK(\theta))}\right).$$

[1] i.e., all tori which are smooth enough are analytic if the map is analytic ([21]).

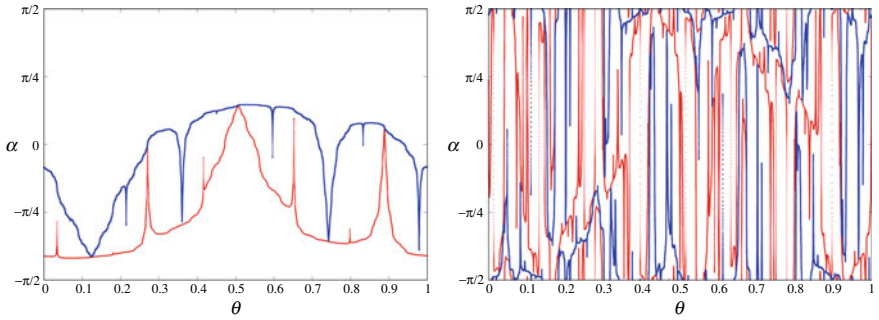


Fig. 10 Invariant bundles close to their collision. Left: dissipative standard map. Right: dissipative standard non-twist map. Reproduced from [17]

This formula says that the angle $\alpha(\theta)$ goes to zero at points where the functions in the denominator go to infinity.

We present figures (see Fig. 10) of the angle between the bundles close to the breakdown.

Rather remarkably these two phenomena (the blow up of Sobolev norms and the stable bundles and the tangent becoming parallel) happen at the same time and present very unexpected regularities. There are scaling relations that seem to be independent of the family considered and they happen in codimension 1 smooth submanifolds in the space of maps. We think that this is a very interesting mathematical phenomenon that deserves rigorous study. It seems quite unlikely that it would have been discovered except for the very careful numerics that can explore with confidence close to the breakdown. Such delicate numerics are only possible because of the rigorous mathematical development.

7 Applications

In this Section we want to briefly review some *constructive* applications of KAM theory for conservative and dissipative models. We will consider applications to the standard map and to the spin-orbit problem, both in the conservative and dissipative settings. Although we will not present other applications of KAM theory, it is worth mentioning also the constructive KAM results to the N-body and planetary problems in Celestial Mechanics ([94]); in this context, for results obtained in the conservative framework we refer the reader to [30–32, 58, 103] and to [38] for numerical investigations including dissipative effects.

7.1 Applications to the Standard Maps

The first applications of computer-assisted KAM proofs have been given for the conservative standard map; these results show that the golden mean torus persists for values of the perturbing parameter equal to 93% of the numerical breakdown value (see [45, 46]); we also mention [29] which, at the same epoch but using a different approach than [45, 46], reached 86% of the numerical breakdown value.

Rigorous estimates for the conservative standard map using the a-posteriori method have been proved in the remarkable paper [56], where for the twist and non-twist conservative standard maps the golden mean torus is proved to persist for values of the perturbing parameter as high as 99.9% of the numerical breakdown value.

For the dissipative standard map, the paper [26] analyzes the persistence of the invariant attractor with frequency equal to the golden mean and for a fixed value of the dissipative parameter (precisely $\lambda = 0.9$); such persistence is shown for values of the perturbing parameter equal to 99.9% of the breakdown value, where the numerical value has been obtained through the techniques presented in Sects. 5.1 and 5.2.

7.2 Applications to the Spin–Orbit Problems

The first application of KAM theory to the conservative spin-orbit problem is found in [33, 34]. In those articles some satellites in synchronous spin-orbit resonance have been considered; the synchronous or 1:1 spin-orbit resonance implies that the satellite always points the same face to the host planet. In particular, the following satellites have been considered: the Moon, and three satellites of Saturn, Rhea, Enceladus, Dione. Being the normalized frequency (namely, the ratio between the rotational and orbital frequency) equal to one, two Diophantine numbers bounding unity from above and below have been considered. Through a computer-assisted KAM theorem, the existence of invariant tori with frequency equal to the bounding numbers have been established for the true values of the parameters of the satellites, namely the eccentricity and the equatorial oblateness.

Such result guarantees the stability for infinite times in the sense of confinement in the phase space. In fact, the phase space associated to the Hamiltonian describing the conservative spin-orbit problem is 3-dimensional; since the KAM tori are 2-dimensional, one gets a confinement of the motion between the bounding invariant tori.

We remark that the confinement is no more valid for $n > 2$ degrees of freedom, since the motion can diffuse through invariant tori, reaching arbitrarily far regions; this phenomenon is known as Arnold's diffusion [4] for which we refer to the extensive literature on this topic (see, e.g., [49, 57] and references therein).

For the dissipative spin-orbit problem, we refer to [19, 36, 87, 105] for the development of KAM theory for a model of spin-orbit interaction with tidal torque as

in (11). Precisely, for $\lambda_0 \in \mathbb{R}_+$ and ω Diophantine, it is proven that there exists $0 < \varepsilon_0 < 1$, such that for any $\varepsilon \in [0, \varepsilon_0]$ and any $\lambda \in [-\lambda_0, \lambda_0]$ there exists a unique function $K = K(\theta, t)$ and a drift term μ which is the solution of the invariance equation for the dissipative spin-orbit model.

Explicit estimates for the dissipative spin-orbit problem, even in the more general case with a time-dependent tidal torque as in (9), are given in [19] (see also [105]). Here, the a-posteriori method is implemented to construct invariant attractors with Diophantine frequency; the results are valid for values of the perturbing parameter consistent with the astronomical quantities and very close to the numerical breakdown threshold, which has been computed in [28] through Sobolev and Greene's method (see also [106]).

Acknowledgements R.C. was partially supported by DGAPA-UNAM PASPA A.C. acknowledges the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C18000100006, EU-ITN Stardust-R, MIUR-PRIN 20178CJA2B "New Frontiers of Celestial Mechanics: theory and Applications". R.L was partially supported by NSF grant DMS 1800241.

References

1. Agrachëv, A.A.: Invariant Lagrangian submanifolds of dissipative systems. Uspekhi Mat. Nauk **65**(5(395)), 185–186 (2010)
2. Arnol'd, V.I.: Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations. Russian Math. Surv. **18**(5), 9–36 (1963)
3. Arnol'd, V.I.: Small denominators and problems of stability of motion in classical and celestial mechanics. Russian Math. Surv. **18**(6), 85–191 (1963)
4. Arnold, V.I.: Instability of dynamical systems with several degrees of freedom. Sov. Math. Doklady **5**, 581–585 (1964)
5. Banyaga, A.: Some properties of locally conformal symplectic structures. Comment. Math. Helv. **77**(2), 383–398 (2002)
6. Bensoussan, A.: Perturbation Methods in Optimal Control. Wiley/Gauthier-Villars Series in Modern Applied Mathematics. Wiley, Chichester (1988). Translated from the French by C. Tomson
7. Bohr, T.: A bound for the existence of invariant circles in a class of two-dimensional dissipative maps. Phys. Lett. A **104**(9), 441–443 (1984)
8. Bohr, T., Bak, P., Jensen, M.H.: Transition to chaos by interaction of resonances in dissipative systems. II. Josephson junctions, charge-density waves, and standard maps. Phys. Rev. A (3) **30**(4), 1970–1981 (1984)
9. Broer, H.W., Huitema, G.B., Sevryuk, M.B.: Families of quasi-periodic motions in dynamical systems depending on parameters. In: Nonlinear Dynamical Systems and Chaos (Groningen, 1995), pp. 171–211. Birkhäuser, Basel (1996)
10. Broer, H.W., Huitema, G.B., Sevryuk, M.B.: Quasi-Periodic Motions in Families of Dynamical Systems, Order Amidst Chaos. Springer, Berlin (1996)
11. Broer, H.W., Huitema, G.B., Takens, F., Braaksma, B.L.J.: Unfoldings and bifurcations of quasi-periodic tori. Mem. Am. Math. Soc. **83**(421), viii+175 (1990)
12. Bustamante, A.P., Calleja, R.C.: Computation of domains of analyticity for the dissipative standard map in the limit of small dissipation. Phys. D **395**, 15–23 (2019)

13. Bustamante, A.P., Calleja, R.C.: Corrigendum and addendum to "Computation of domains of analyticity for the dissipative standard map in the limit of small dissipation" [Physica D 395 (2019) 15–23]. Phys. D **417**, Paper No. 132837, 7 (2021)
14. Bustamante, A.P., de la Llave, R.: Gevrey Estimates for the Asymptotic Expansion of Tori of Weakly Dissipative Systems (2020). https://arxiv.org/abs/2010.06006
15. Calleja, R., Celletti, A.: Breakdown of invariant attractors for the dissipative standard map. Chaos **20**(1), 013121 (2010)
16. Calleja, R., de la Llave, R.: A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification. Nonlinearity **23**(9), 2029–2058 (2010)
17. Calleja, R., Canadell, M., Haro, A.: Non-twist invariant circles in conformally symplectic systems. Commun. Nonlinear Sci. Numer. Simul. **96**, Paper No. 105695, 15 (2021)
18. Calleja, R., Celletti, A., Gimeno, J., de la Llave, R.: Efficient and accurate KAM Tori construction for the dissipative spin-orbit problem using a map reduction. J. Nonlinear Sci. **32**(1), Paper No. 4 (2022)
19. Calleja, R., Celletti, A., Gimeno, J., de la Llave, R.: KAM quasi-periodic tori for the dissipative spin-orbit problem. Commun. Nonlinear Sci. Numer. Simul. **106**, Paper No. 106099 (2022)
20. Calleja, R., Figueras, J.-L.: Collision of invariant bundles of quasi-periodic attractors in the dissipative standard map. Chaos **22**(3), 033114, 10 (2012)
21. Calleja, R.C., Celletti, A., de la Llave, R.: A KAM theory for conformally symplectic systems: efficient algorithms and their validation. J. Diff. Equ. **255**(5), 978–1049 (2013)
22. Calleja, R.C., Celletti, A., de la Llave, R.: Local behavior near quasi-periodic solutions of conformally symplectic systems. J. Dynam. Diff. Equ. **25**(3), 821–841 (2013)
23. Calleja, R.C., Celletti, A., de la Llave, R.: Domains of analyticity and Lindstedt expansions of KAM tori in some dissipative perturbations of Hamiltonian systems. Nonlinearity **30**(8), 3151–3202 (2017)
24. Calleja, R.C., Celletti, A., de la Llave, R.: Existence of whiskered KAM tori of conformally symplectic systems. Nonlinearity **33**(1), 538–597 (2020)
25. Calleja, R.C., Celletti, A., de la Llave, R.: Whiskered KAM tori of conformally symplectic systems. Math. Res. Rep. **1**, 15–29 (2020)
26. Calleja, R.C., Celletti, A., de la Llave, R.: KAM quasi-periodic solutions for the dissipative standard map. Commun. Nonlinear Sci. Numer. Simul. **106**, Paper No. 106111 (2022)
27. Calleja, R.C., Celletti, A., Falcolini, C., de la Llave, R.: An extension of greene's criterion for conformally symplectic systems and a partial justification. SIAM J. Math. Anal. **46**(4), 2350–2384 (2014)
28. Calleja, R.C., Celletti, A., Gimeno, J., de la Llave, R.: Breakdown threshold of invariant attractors in the dissipative spin-orbit problem. Preprint (2020)
29. Celletti, A., Chierchia, L.: A constructive theory of Lagrangian tori and computer-assisted applications. In: Dynamics Reported, pp. 60–129. Springer, Berlin (1995)
30. Celletti, A., Chierchia, L.: On the stability of realistic three-body problems. Commun. Math. Phys. **186**(2), 413–449 (1997)
31. Celletti, A., Chierchia, L.: KAM tori for N-body problems: a brief history. Celest. Mech. Dynam. Astronom. **95**(1–4), 117–139 (2006)
32. Celletti, A., Chierchia, L.: KAM stability and celestial mechanics. Mem. Am. Math. Soc. **187**(878) (2007)
33. Celletti, A.: Analysis of resonances in the spin-orbit problem in celestial mechanics: higher order resonances and some numerical experiments. II. Z. Angew. Math. Phys. **41**(4), 453–479 (1990)
34. Celletti, A.: Analysis of resonances in the spin-orbit problem in celestial mechanics: the synchronous resonance. I. Z. Angew. Math. Phys. **41**(2), 174–204 (1990)
35. Celletti, A.: Stability and Chaos in Celestial Mechanics. Springer, Published in Association with Praxis Publishing, Berlin, Chichester (2010)
36. Celletti, A., Chierchia, L.: Quasi-periodic attractors in celestial mechanics. Arch. Ration. Mech. Anal. **191**(2), 311–345 (2009)

37. Celletti, A., MacKay, R.: Regions of nonexistence of invariant tori for spin-orbit models. Chaos **17**(4), 043119, 12 (2007)
38. Celletti, A., Stefanelli, L., Lega, E., Froeschlé, C.: Some results on the global dynamics of the regularized restricted three-body problem with dissipation. Celest. Mech. Dynam. Astronom. **109**(3), 265–284 (2011)
39. Chierchia, L., Gallavotti, G.: Smooth prime integrals for quasi-integrable Hamiltonian systems. Nuovo Cimento B (11) **67**(2), 277–295 (1982)
40. Chirikov, B.V.: A universal instability of many-dimensional oscillator systems. Phys. Rep. **52**(5), 264–379 (1979)
41. Ciocci, M.-C., Litvak-Hinenzon, A., Broer, H.: Survey on dissipative KAM theory including quasi-periodic bifurcation theory (H. Broer). London Mathematical Society Lecture Note Series, pp. 303–356. Cambridge University Press (2005)
42. Correia, A.C.M., Laskar, J.: Mercury's capture into the 3/2 spin-orbit resonance as a result of its chaotic dynamics. Nature **429**(6994), 848–850 (2004)
43. de la Llave, R.: A tutorial on KAM theory. Smooth ergodic theory and its applications (Seattle. WA, 1999). Am. Math. Soc. Providence, RI 175–292 (2001)
44. de la Llave, R., González, A., Jorba, À., Villanueva, J.: KAM theory without action-angle variables. Nonlinearity **18**(2), 855–895 (2005)
45. de la Llave, R., Rana, D.: Accurate strategies for small divisor problems. Bull. Am. Math. Soc. (N.S.) **22**(1), 85–90 (1990)
46. de la Llave, R., Rana, D.: Accurate strategies for K.A.M. bounds and their implementation. In: Computer Aided Proofs in Analysis (Cincinnati. OH, 1989), pp. 127–146. Springer, New York (1991)
47. del Castillo-Negrete, D., Greene, J.M., Morrison, P.J.: Area preserving nontwist maps: periodic orbits and transition to chaos. Phys. D **91**(1–2), 1–23 (1996)
48. del Castillo-Negrete, D., Greene, J.M., Morrison, P.J.: Renormalization and transition to chaos in area preserving nontwist maps. Phys. D **100**(3–4), 311–329 (1997)
49. Delshams, A., de la Llave, R., Seara, T.M. (2006) A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. Mem. Am. Math. Soc. **179**(844), viii+141
50. Dettmann, C.P., Morris, G.P.: Proof of Lyapunov exponent pairing for systems at constant kinetic energy. Phys. Rev. E **53**(6), R5545–R5548 (2006)
51. Dettmann, C.P., Morriss, G.P.: Hamiltonian formulation of the Gaussian isokinetic thermostat. Phys. Rev. E **54**, 2495–2500 (1996)
52. Eaton, J.W., Bateman, D., Hauberg, S., Wehbring, R.: GNU Octave Version 5.2.0 Manual: a High-Level Interactive Language for Numerical Computations (2020)
53. Eliasson, L.H., Fayad, B., Krikorian, R.: Around the stability of KAM tori. Duke Math. J. **164**(9), 1733–1775, 06 (2015)
54. Falcolini, C., de la Llave, R.: A rigorous partial justification of Greene's criterion. J. Statist. Phys. **67**(3–4), 609–643 (1992)
55. Feudel, U., Grebogi, C., Hunt, B.R., Yorke, J.A.: Map with more than 100 coexisting low-period periodic attractors. Phys. Rev. E (3) **54**(1), 71–81 (1996)
56. Figueras, J.-Ll., Haro, A., Luque, A.: Rigorous computer-assisted application of KAM theory: a modern approach. Found. Comput. Math. **17**(5), 1123–1193 (2017)
57. Gidea, M., de la Llave, R., M-Seara, T.: A general mechanism of diffusion in Hamiltonian systems: qualitative results. Commun. Pure Appl. Math. **73**(1), 150–209 (2020)
58. Giorgilli, A., Locatelli, U., Sansottera, M.: An extension of Lagrange's theory for secular motions. Rend. Cl. Sci. Mat. Nat. **143**, 223–239 (2009)
59. Greene, J.M.: A method for determining a stochastic transition. J. Math. Phys. **20**, 1183–1201 (1979)
60. Guckenheimer, J., Holmes, P.: Nonlinear oscillations, dynamical systems, and bifurcations of vector fields. Applied Mathematical Sciences, vol. 42. Springer, New York (1990). Revised and corrected reprint of the 1983 original

61. Haro, A.: Automatic differentiation tools in computational dynamical systems. Work in Progress (2011)
62. Haro, À., Canadell, M., Figueras, J.-L., Luque, A., Mondelo, J.-M.: The parameterization method for invariant manifolds. Applied Mathematical Sciences, vol. 195. Springer, [Cham] (2016). From rigorous results to effective computations
63. Haro, A., Luque, A.: A-posteriori KAM theory with optimal estimates for partially integrable systems. J. Diff. Equ. **266**(2–3), 1605–1674 (2019)
64. Haro, A., Luque, A.: A-posteriori KAM theory with optimal estimates for partially integrable systems. J. Diff. Equ. **266**(2), 1605–1674 (2019)
65. Herman, M., Sergeraert, F.: Sur un théorème d'Arnold et Kolmogorov. C. R. Acad. Sci. Paris Sér. A-B **273**, A409–A411 (1971)
66. Herman, M.-R.: Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. Inst. Hautes Études Sci. Publ. Math. **49**, 5–233 (1979)
67. Jungreis, I.: A method for proving that monotone twist maps have no invariant circles. Ergod. Theory Dyn. Syst. **11**(1), 79–84 (1991)
68. Katznelson, Y., Ornstein, D.: The absolute continuity of the conjugation of certain diffeomorphisms of the circle. Ergod. Theory Dyn. Syst. **9**(4), 681–690 (1989)
69. Katznelson, Y., Ornstein, D.: The differentiability of the conjugation of certain diffeomorphisms of the circle. Ergod. Theory Dyn. Syst. **9**(4), 643–680 (1989)
70. Ketoja, J.A., Satija, I.I.: Harper Equation, the Dissipative Standard Map and Strange Nonchaotic Attractors: Relationship Between an Eigenvalue Problem and Iterated Maps, vol. 109, pp. 70–80 (1997). Physics and Dynamics Between Chaos, Order, and Noise (Berlin, 1996)
71. Khanin, K.M., Sinai, Y.G.: A new proof of M. Herman's theorem. Commun. Math. Phys. **112**(1), 89–101 (1987)
72. Kim, S.-Y., Lee, D.-S.: Transition to chaos in a dissipative standardlike map. Phys. Rev. A (3) **45**(8), 5480–5487 (1992)
73. Koch, H.: A renormalization group for Hamiltonians, with applications to KAM tori. Ergod. Theory Dyn. Syst. **19**, 1–47 (1999)
74. Koch, H., Schenkel, A., Wittwer, P.: Computer-assisted proofs in analysis and programming in logic: a case study. SIAM Rev. **38**(4), 565–604 (1996)
75. Koch, H.: A renormalization group fixed point associated with the breakup of golden invariant tori. Disc. Contin. Dyn. Syst. **11**(4), 881–909 (2004)
76. Koch, H.: Existence of critical invariant tori. Ergod. Theory Dyn. Syst. **28**(6), 1879–1894 (2008)
77. Koch, H.: On hyperbolicity in the renormalization of near-critical area-preserving maps. Discrete Contin. Dyn. Syst. **36**(12), 7029–7056 (2016)
78. Kolmogorov, A.N.: On conservation of conditionally periodic motions for a small change in Hamilton's function. Dokl. Akad. Nauk SSSR (N.S.) **98**, 527–530 (1954). English translation in Stochastic Behavior in Classical and Quantum Hamiltonian Systems (Volta Memorial Conf., Como, 1977). Lecture Notes in Physics, vol. 93, pp. 51–56. Springer, Berlin (1979)
79. Kyner, W.T.: Rigorous and formal stability of orbits about an oblate planet. Mem. Am. Math. Soc. No. 81. Am. Math. Soc. Providence, R.I. (1968)
80. Lanford, O.E.: III. Computer-assisted proofs in analysis. In: Proceedings of the International Congress of Mathematicians, vol. 1, 2 (Berkeley, Calif., 1986), pp. 1385–1394, Providence, RI (1987). Amer. Math. Soc
81. Laskar, J., Froeschlé, C., Celletti, A.: The measure of chaos by the numerical analysis of the fundamental frequencies. Application to the standard mapping. Physica D **56**, 253–269 (1992)
82. Levi, M.: Qualitative analysis of the periodically forced relaxation oscillations. Mem. Am. Math. Soc. **32**(244), vi+147 (1981)
83. Lichtenberg, A.J., Lieberman, M.A.: Regular and chaotic dynamics. Applied Mathematical Sciences, 2nd edn., vol. 38. Springer, New York (1992)
84. Lin, K.K., Young, L.-S.: Shear-induced chaos. Nonlinearity **21**(5), 899–922 (2008)
85. MacKay, R.S.: Greene's residue criterion. Nonlinearity **5**(1), 161–187 (1992)

86. Massetti, J.E.: A normal form à la Moser for diffeomorphisms and a generalization of Rüssmann's translated curve theorem to higher dimensions. Anal. PDE **11**(1), 149–170 (2018)
87. Massetti, J.E.: Normal forms for perturbations of systems possessing a Diophantine invariant torus. Ergod. Theory Dyn. Syst. **39**(8), 2176–2222 (2019)
88. McKay, R.S.: Renormalisation in Area Preserving Maps. Ph.D. thesis, Princeton University (1982)
89. Meyer, K.R.: Lie transform tutorial. II. In: Computer Aided Proofs in Analysis (Cincinnati. OH, 1989), pp. 190–210. Springer, New York (1991)
90. Moser, J.: On invariant curves of area-preserving mappings of an annulus. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II **1962**, 1–20 (1962)
91. Moser, J.: A rapidly convergent iteration method and non-linear differential equations. II. Ann. Scuola Norm. Sup. Pisa **3**(20), 499–535 (1966)
92. Moser, J.: Convergent series expansions for quasi-periodic motions. Math. Ann. **169**, 136–176 (1967)
93. Peale, S.J.: The free precession and libration of Mercury. Icarus **178**(1), 4–18 (2005)
94. Poincaré, H.: New Methods of Celestial Mechanics, vol. 1–3. American Institute of Physics, New York (1993)
95. Pöschel, J.: Integrability of Hamiltonian systems on Cantor sets. Commun. Pure Appl. Math. **35**(5), 653–696 (1982)
96. Pöschel, J.: A lecture on the classical KAM theorem. In: Smooth Ergodic Theory and its Applications (Seattle, WA, 1999). Proceedings of Symposia in Pure Mathematics, vol. 69, pp. 707–732 (2001). . Am. Math. Soc., Providence, RI
97. Rana, D.: Proof of Accurate Upper and Lower Bounds to Stability Domains in Small Denominator Problems. Ph.D. thesis, Princeton University (1987)
98. Rand, D.A.: Existence, nonexistence and universal breakdown of dissipative golden invariant tori. I. Golden critical circle maps. Nonlinearity **5**(3), 639–662 (1992)
99. Rand, D.A.: Existence, nonexistence and universal breakdown of dissipative golden invariant tori. II. Convergence of renormalization for mappings of the annulus. Nonlinearity **5**(3), 663–680 (1992)
100. Rand, D.A.: Existence, nonexistence and universal breakdown of dissipative golden invariant tori. III. Invariant circles for mappings of the annulus. Nonlinearity **5**(3), 681–706 (1992)
101. Rüssmann, H.: Kleine Nenner. I. Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II **1970**, 67–105 (1970)
102. Rüssmann, H.: On optimal estimates for the solutions of linear difference equations on the circle. Celest. Mech. **14**(1), 33–37 (1976)
103. Sansottera, M., Locatelli, U., Giorgilli, A.: On the stability of the secular evolution of the planar Sun-Jupiter-Saturn-Uranus system. Math. Comput. Simul. **88**, 1–14 (2013)
104. Schmidt, G., Wang, B.W.: Dissipative standard map. Phys. Rev. A **32**(5), 2994–2999 (1985)
105. Stefanelli, L., Locatelli, U.: Kolmogorov's normal form for equations of motion with dissipative effects. Disc. Contin. Dyn. Syst. **17**(7), 2561–2593 (2012)
106. Stefanelli, L., Locatelli, U.: Quasi-periodic motions in a special class of dynamical equations with dissipative effects: A pair of detection methods. Disc. Contin. Dyn. Syst. Ser. B **20**(4), 1155–1187 (2015)
107. Stirnemann, A.: Renormalization for golden circles. Commun. Math. Phys. **152**(2), 369–431 (1993)
108. Stirnemann, A.: Towards an existence proof of MacKay's fixed point. Commun. Math. Phys. **188**(3), 723–735 (1997)
109. Vaisman, I.: Locally conformal symplectic manifolds. Int. J. Math. Math. Sci. **8**(3), 521–536 (1985)
110. Vlasova, O.F., Zaslavsky, G.M.: Nonergodic regions in the standard dissipative mapping. Phys. Lett. A **105**(1–2), 1–5 (1984)
111. Wang, Q., Young, L.-S.: Toward a theory of rank one attractors. Ann. Math. (2) **167**(2), 349–480 (2008)

112. Wayne, C.E.: An introduction to KAM theory. In: Dynamical Systems and Probabilistic Methods in Partial Differential Equations (Berkeley, CA, 1994), pp. 3–29. Am. Math. Soc, Providence, RI (1996)
113. Wenzel, W., Biham, O., Jayaprakash, C.: Periodic orbits in the dissipative standard map. Phys. Rev. A **43**(12), 6550–6557 (1991). cited By 34
114. Wisdom, J., Peale, S.J., Mignard, F.: The chaotic rotation of Hyperion. Icarus **58**(2), 137–152 (1984)
115. Wojtkowski, M.P., Liverani, C.: Conformally symplectic dynamics and symmetry of the Lyapunov spectrum. Commun. Math. Phys. **194**(1), 47–60 (1998)
116. Yamaguchi, Y.: Breakup of an invariant curve in a dissipative standard mapping. Phys. Lett. A **116**(7), 307–310 (1986). cited By 1

Tidal Effects and Rotation of Extended Bodies

Gwenaël Boué

Abstract These lecture notes provide basic tools to study the rotation of extended bodies which are either rigid or deformable by tides. The problem is written in a Lagrangian formalism using the "forme nouvelle des équations de la mécanique" developed by H. Poincaré in 1901. In the rigid body case, the corresponding Hamiltonian formalism is also presented. When the deformation of the extended body is taken into account, the mathematical description of the problem closely follows the approach presented by C. Ragazzo and L. Ruiz in their first two papers (2015, 2017). This choice is motivated by the compactness and the clarity of the equations stemmed from this formalism. These notes are illustrated with selected applications related to the rotation and libration of celestial bodies.

Keywords Extended body · Tide · Rotation

1 Introduction

The aim of these lecture notes is to provide basic tools to study the rotation of extended bodies which are either rigid or deformable by tides. The problem is written in a Lagrangian formalism using the "*forme nouvelle des équations de la mécanique*" developed by H. Poincaré in 1901 [30]. For the rigid body case, I also present the corresponding Hamiltonian formalism. When the deformation of the extended body is taken into account, I closely follow the approach presented by C. Ragazzo and L. Ruiz in their first two papers [32, 33]. This choice is motivated by the compactness and the clarity of the equations stemmed from this formalism.

These notes are illustrated with examples. Because it is not possible to be exhaustive on this subject, I selected a few problems for which the formalism is well suited and others presenting important dynamical features.

G. Boué (✉)
IMCCE, UMR8028 CNRS, Observatoire de Paris, PSL University and Sorbonne Université, IMCCE, 77 av. Denfert-Rochereau, 75014 Paris, France
e-mail: gwenael.boue@obspm.fr

G. Baù et al. (eds.), *New Frontiers of Celestial Mechanics: Theory and Applications*, Springer Proceedings in Mathematics & Statistics 399,
https://doi.org/10.1007/978-3-031-13115-8_4

The manuscript is organised as follows: the next three sections deal with the rigid motion in a broad sense. I present a few coordinate systems, their respective kinematic equations, and the derivation of the dynamical equations. The rotation of a rigid celestial body on a Keplerian orbit is studied in Sects. 5 and 6. In the last two sections, I describe tidal deformation and its consequence on the rotational motion of a deformable celestial body.

2 Coordinate System

The physical space is represented by an Euclidean affine space $\mathcal{E}$ of dimension 3 whose associated vector space is denoted by $\overrightarrow{\mathcal{E}}$. We attach to this space an inertial Cartesian reference frame $\mathcal{R}_0 = (\mathrm{O}, \overrightarrow{\mathrm{e}_1}, \overrightarrow{\mathrm{e}_2}, \overrightarrow{\mathrm{e}_3})$ in which Newton's laws hold.

Let $\mathcal{B}$ be an extended rigid body. For instance, this can be a planet or a satellite whose spatial extension is taken into account but not its ability of being deformed. We assign to this body $\mathcal{B}$, a so-called *body-fixed frame* in which the coordinates of all points of $\mathcal{B}$ remain constant. We denote by $\mathcal{R}_1 = (\mathrm{O}', \overrightarrow{\mathrm{f}_1}, \overrightarrow{\mathrm{f}_2}, \overrightarrow{\mathrm{f}_3})$ such a frame. Usually O' is the barycentre of $\mathcal{B}$ and $(\overrightarrow{\mathrm{f}_1}, \overrightarrow{\mathrm{f}_2}, \overrightarrow{\mathrm{f}_3})$ its principal axes of inertia.

Obtaining the motion of $\mathcal{B}$ amounts to get the motion of the frame $\mathcal{R}_1$ with respect to the frame $\mathcal{R}_0$. Therefore a rigid motion is, by definition, an Euclidean transformation of $\mathcal{E}$ that preserves both the Euclidean distance between every pair of points and the handedness. It consists of the identity, a translation, a rotation, or a combination of them.

The set of rigid motions forms the so-called special Euclidean group denoted by SE(3). This is a Lie group of dimension 6 which can be identified to the set $\mathbb{R}^3 \times \mathrm{SO}(3)$. The element of $\mathbb{R}^3$ is either $\mathbf{r}^0$ or $\mathbf{r}^1$, namely the coordinates of the radius vector $\overrightarrow{\mathrm{OO}'} \in \overrightarrow{\mathcal{E}}$ in the frame $\mathcal{R}_0$ and $\mathcal{R}_1$, respectively; and $\mathsf{R} \in \mathrm{SO}(3)$ is the rotation matrix whose entries are the coordinates of $(\overrightarrow{\mathrm{f}_1}, \overrightarrow{\mathrm{f}_2}, \overrightarrow{\mathrm{f}_3})$ in $(\overrightarrow{\mathrm{e}_1}, \overrightarrow{\mathrm{e}_2}, \overrightarrow{\mathrm{e}_3})$. The coordinates of $\overrightarrow{\mathrm{OO}'}$ are associated with the three degrees of freedom of translation of $\mathcal{B}$ and R with its three degrees of freedom of rotation.

Let $\mathrm{P} \in \mathcal{E}$ be a point belonging to $\mathcal{B}$ with coordinates $\mathbf{x}^0$ in $\mathcal{R}_0$ and $\mathbf{x}^1$ in $\mathcal{R}_1$. By construction $\mathbf{x}^1$ remains constant for any time, whereas $\mathbf{x}^0$ varies as $\mathcal{B}$ moves in $\mathcal{R}_0$. Hereafter, we highlight the dependency on time t of the coordinates of P in $\mathcal{R}_0$ by writing $\mathbf{x}^0(t)$. The mapping between these two sets of coordinates is given by

$$\mathbf{x}^0(t) = \mathbf{r}^0(t) + \mathsf{R}(t)\mathbf{x}^1 , \tag{1}$$

or equivalently, by

$$\mathbf{x}^0(t) = \mathsf{R}(t)\left(\mathbf{r}^1(t) + \mathbf{x}^1\right) . \tag{2}$$

The variables $(\mathbf{r}^0, \mathsf{R})$ and $(\mathbf{r}^1, \mathsf{R})$ are natural sets of coordinates to describe the motion of a rigid body. In the following, we will use them as generalised coordinates

in the Lagrangian formalism. However, although the set SO(3) is of dimension 3, rotation matrices have 9 entries. To avoid such a redundancy, we briefly present a few alternative parametrisations of SO(3).

2.1 The 3-1-3 Euler Angles

The 3-1-3 Euler angles (ϕ, θ, ψ), called precession angle, nutation angle and proper rotation, are such that the basis $(\overrightarrow{f_1}, \overrightarrow{f_2}, \overrightarrow{f_3})$ is deduced from $(\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3})$ by a rotation of angle ϕ around the third axis followed by a rotation of angle θ around the new first axis itself followed by a rotation of angle ψ around the final third axis. Mathematically, this reads $\mathsf{R} = \mathsf{R}_3(\phi)\mathsf{R}_1(\theta)\mathsf{R}_3(\psi)$ where $\mathsf{R}_k(\alpha)$, $k = 1, 2, 3$, are elemental rotation matrices of angle α defined as

$$\begin{aligned} \mathsf{R}_1(\alpha) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}, \qquad \mathsf{R}_2(\alpha) = \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix}, \\ \mathsf{R}_3(\alpha) &= \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{3}$$

A direct calculation gives

$$\mathsf{R} = \begin{bmatrix} \cos\phi\cos\psi - \sin\phi\cos\theta\sin\psi & -\cos\phi\sin\psi - \sin\phi\cos\theta\cos\psi & \sin\phi\sin\theta \\ \sin\phi\cos\psi + \cos\phi\cos\theta\sin\psi & \cos\phi\cos\theta\cos\psi - \sin\phi\sin\psi & -\cos\phi\sin\theta \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{bmatrix}. \tag{4}$$

2.2 Unitary Quaternion

A quaternion $q \in \mathbb{H}$ is a set of 4 coordinates (q_0, q_1, q_2, q_3) which can be decomposed into a scalar part $q_0 \in \mathbb{R}$ and a vector part $\mathbf{q} = (q_1, q_2, q_3) \in \mathbb{R}^3$ with the following formulas for the addition and the multiplication

$$(a_1, \mathbf{v}_1) + (a_2, \mathbf{v}_2) = (a_1 + a_2, \mathbf{v}_1 + \mathbf{v}_2), \tag{5}$$

$$(a_1, \mathbf{v}_1)(a_2, \mathbf{v}_2) = (a_1 a_2 - \mathbf{v}_1 \cdot \mathbf{v}_2,\ a_1\mathbf{v}_2 + a_2\mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2). \tag{6}$$

The set of quaternions $\mathbb{H}$ possesses a conjugation relation such that the conjugate of $q = (q_0, \mathbf{q})$ is $q^* = (q_0, -\mathbf{q})$. This allows to construct the norm of a quaternion defined as $\|q\|^2 = qq^* = q_0^2 + q_1^2 + q_2^2 + q_3^2$.

A rotation of angle α around an axis of unit vector $\mathbf{n}$ can be represented by the quaternion $q = (\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2}\mathbf{n})$ of length 1 (by construction the set of unitary quaternions is of dimension 3 like SO(3)). In terms of $q = (q_0, q_1, q_2, q_3)$ the rotation matrix reads

$$\mathsf{R} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_3q_0) & 2(q_1q_3 + q_2q_0) \\ 2(q_1q_2 + q_3q_0) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_1q_0) \\ 2(q_1q_3 - q_2q_0) & 2(q_2q_3 + q_1q_0) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix}. \quad (7)$$

The image of a vector $\mathbf{v}$ by a rotation of quaternion q is given by $\mathsf{R}\mathbf{v} = q\mathbf{v}q^*$. Here, we identify a quaternion $q = (0, \mathbf{q}) \in \mathbb{H}$ whose scalar part is null with its vector $\mathbf{q} \in \mathbb{R}^3$ and vice versa. The quaternion associated with the rotation $\mathsf{R} = \mathsf{R}(q_a)\mathsf{R}(q_b)$ is thus $q = q_a q_b$.

2.3 Special Case: Axisymmetric Body

When the body $\mathcal{B}$ is axisymmetric, any dynamical problem involving $\mathcal{B}$ is invariant by rotation of $\mathcal{B}$ around its axis of symmetry. Let $\overrightarrow{\mathrm{f}_3}$ be a unit vector along this axis. The kinetic energy and the potential energy shall not depend on the other vectors $(\overrightarrow{\mathrm{f}_1}, \overrightarrow{\mathrm{f}_2})$ of the base frame of $\mathcal{R}_1$. Therefore, the direction of $\overrightarrow{\mathrm{f}_3}$ is enough to describe the orientation of $\mathcal{B}$. In that case, only two degrees of freedom of rotation are left since the set $\{\overrightarrow{\mathrm{f}_3} \in \overrightarrow{\mathcal{E}}, \|\overrightarrow{\mathrm{f}_3}\| = 1\}$ is of dimension 2. Indeed, among all three possible rotations, we have discarded that around the vector $\overrightarrow{\mathrm{f}_3}$.

3 Generalised Velocity and Kinematic Equation

To describe the evolution of the configuration of $\mathcal{B}$ with respect to $\mathcal{R}_0$, it is natural to introduce the translation rate, i.e. the linear velocity $\overrightarrow{v}$ of the origin O', and the rotation rate, namely the angular velocity $\overrightarrow{\omega}$. More specifically, we use their coordinates $(\mathbf{v}^0, \boldsymbol{\omega}^0)$ in the frame $\mathcal{R}_0$ or $(\mathbf{v}^1, \boldsymbol{\omega}^1)$ in the frame $\mathcal{R}_1$. These velocities are related to the generalised coordinates described above through differential equations called *kinematic equations*. In the standard Lagrangian formalism, where coordinates are denoted by $\mathbf{q}$ and velocities $\dot{\mathbf{q}}$, kinematic equations are simply $\mathrm{d}\mathbf{q}/\mathrm{d}t = \dot{\mathbf{q}}$. In our problem, the same relation holds for the translation motion between $\mathbf{r}$ and $\mathbf{v}$ as we do have by definition $\mathrm{d}\mathbf{r}^0/\mathrm{d}t = \mathbf{v}^0$ and $\mathrm{d}\mathbf{r}^1/\mathrm{d}t = \mathbf{v}^1$. This is nevertheless not the case with the rotation motion and the associated kinematic equations have to be adapted to the chosen coordinate system.

3.1 Kinematic Equation Satisfied by the Rotation Matrix

To get the kinematic equation satisfied by the rotation matrix R, we consider a vector $\overrightarrow{u} \in \overrightarrow{\mathcal{E}}$ attached to $\mathcal{B}$ (practically, $\overrightarrow{u} = \overrightarrow{\mathrm{PP}'}$ where P and P′ are two points of $\mathcal{B}$). Its coordinates $\mathbf{u}^1$ in $\mathcal{R}_1$ are constant while its coordinates $\mathbf{u}^0$ in $\mathcal{R}_0$ vary in time according to

$$\mathbf{u}^0(t) = \mathsf{R}(t)\mathbf{u}^1 \,. \tag{8}$$

The time derivative of this equality gives

$$\dot{\mathbf{u}}^0(t) = \dot{\mathsf{R}}(t)\mathbf{u}^1 = \dot{\mathsf{R}}(t)\mathsf{R}(t)^\top \mathbf{u}^0(t) \,. \tag{9}$$

By construction $\dot{\mathsf{R}}(t)\mathsf{R}^\top(t)$ belongs to the set skew(3) of antisymmetric matrices. This set of degree 3 is in bijection with $\mathbb{R}^3$. Among the bijections from $\mathbb{R}^3$ to skew(3), we choose the one—which we refer to as the *hat operator* : $\mathbf{a} \in \mathbb{R}^3 \mapsto \hat{\mathbf{a}} \in \mathrm{skew}(3)$—such that for any two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, $\hat{\mathbf{a}}\,\mathbf{b} = \mathbf{a} \times \mathbf{b}$. The hat operator is defined for any vector $\mathbf{a} \in \mathbb{R}^3$ as

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mapsto \hat{\mathbf{a}} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} . \tag{10}$$

By definition, the angular velocity $\boldsymbol{\omega}^0$ is the vector of $\mathbb{R}^3$ such that $\dot{\mathbf{u}}^0(t) = \boldsymbol{\omega}^0(t) \times \mathbf{u}^0(t)$. Therefore, $\hat{\boldsymbol{\omega}}^0 = \dot{\mathsf{R}}\mathsf{R}^\top$ and the kinematic equation associated with the rotation motion reads in $\mathcal{R}_0$ as

$$\dot{\mathsf{R}} = \hat{\boldsymbol{\omega}}^0 \mathsf{R} \,. \tag{11}$$

In the frame $\mathcal{R}_1$ we have $\boldsymbol{\omega}^1 = \mathsf{R}^\top \boldsymbol{\omega}^0$, or equivalently, $\hat{\boldsymbol{\omega}}^1 = \mathsf{R}^\top \hat{\boldsymbol{\omega}}^0 \mathsf{R}$. By consequence, $\hat{\boldsymbol{\omega}}^1 = \mathsf{R}^\top \dot{\mathsf{R}}$ and the kinematic equation written in $\mathcal{R}_1$ is

$$\dot{\mathsf{R}} = \mathsf{R}\,\hat{\boldsymbol{\omega}}^1 \,. \tag{12}$$

3.2 Kinematic Equation Satisfied by the 3-1-3 Euler Angles

Let us represent the three rotations associated with the three Euler angles as in Fig. 1. The expression of the angular velocity $\overrightarrow{\omega}$ can then be read directly from the figure, namely,

$$\begin{aligned} \overrightarrow{\omega} &= \dot{\phi}\,\overrightarrow{\mathrm{e}_3} + \dot{\theta}\,\overrightarrow{\mathrm{e}_1}' + \dot{\psi}\,\overrightarrow{\mathrm{f}_3} \\ &= (\dot{\theta}\cos\phi + \dot{\psi}\sin\theta\sin\phi)\overrightarrow{\mathrm{e}_1} + (\dot{\theta}\sin\phi - \dot{\psi}\sin\theta\cos\phi)\overrightarrow{\mathrm{e}_2} + (\dot{\phi} + \dot{\psi}\cos\theta)\overrightarrow{\mathrm{e}_3} \\ &= (\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi)\overrightarrow{\mathrm{f}_1} + (\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)\overrightarrow{\mathrm{f}_2} + (\dot{\phi}\cos\theta + \dot{\psi})\overrightarrow{\mathrm{f}_3} \,. \end{aligned} \tag{13}$$

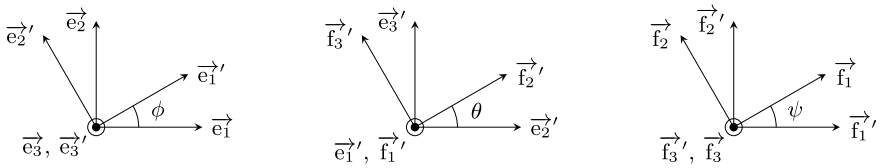


Fig. 1 Three Euler rotations: $(\vec{e_1}, \vec{e_2}, \vec{e_3}) \xrightarrow{\mathsf{R}_3(\phi)} (\vec{e_1}', \vec{e_2}', \vec{e_3}') \xrightarrow{\mathsf{R}_1(\theta)} (\vec{f_1}', \vec{f_2}', \vec{f_3}') \xrightarrow{\mathsf{R}_3(\psi)} (\vec{f_1}, \vec{f_2}, \vec{f_3})$

The components along $(\vec{e_1}, \vec{e_2}, \vec{e_3})$ of $\vec{\omega}$ are linear combinations of $(\dot{\phi}, \dot{\theta}, \dot{\psi})$. Inverting these relations, we get the desired kinematic equations in terms of the coordinates $\boldsymbol{\omega}^0$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\phi\\ \theta\\ \psi\end{pmatrix} = \begin{bmatrix}-\cot\theta\sin\phi & \cot\theta\cos\phi & 1\\ \cos\phi & \sin\phi & 0\\ \csc\theta\sin\phi & -\csc\theta\cos\phi & 0\end{bmatrix}\begin{pmatrix}\omega_1^0\\ \omega_2^0\\ \omega_3^0\end{pmatrix}. \tag{14}$$

Using the decomposition of $\vec{\omega}$ into $(\vec{f_1}, \vec{f_2}, \vec{f_3})$, one gets the equivalent kinematic equations in terms of the coordinates $\boldsymbol{\omega}^1$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\phi\\ \theta\\ \psi\end{pmatrix} = \begin{bmatrix}\csc\theta\sin\psi & \csc\theta\cos\psi & 0\\ \cos\psi & -\sin\psi & 0\\ -\cot\theta\sin\psi & -\cot\theta\cos\psi & 1\end{bmatrix}\begin{pmatrix}\omega_1^1\\ \omega_2^1\\ \omega_3^1\end{pmatrix}. \tag{15}$$

3.3 Kinematic Equation Satisfied by Unitary Quaternions

As said in Sect. 2.2, the rotation of a vector can be written in terms of quaternion products. For our purpose, if q_t denotes the quaternion associated with $\mathsf{R}(t)$, we have $\mathbf{u}^0(t) = q_t\,\mathbf{u}^1\,q_t^*$. Taking the time derivative of this expression, we get

$$\dot{\mathbf{u}}^0(t) = \dot{q}_t\,\mathbf{u}^1\,q_t^* + q_t\,\mathbf{u}^1\,\dot{q}_t^* = \dot{q}_t q_t^*\,\mathbf{u}^0(t) + \mathbf{u}^0(t)\,q_t\dot{q}_t^*. \tag{16}$$

We shall now remember that q_t is unitary, therefore $q_t q_t^* = 1$ which leads to $q_t\dot{q}_t^* = -\dot{q}_t q_t^*$. Hence,

$$\dot{\mathbf{u}}^0(t) = [\dot{q}_t q_t^*, \mathbf{u}^0(t)] = (\dot{q}_t q_t^*)\mathbf{u}^0(t) - \mathbf{u}^0(t)(\dot{q}_t q_t^*). \tag{17}$$

A direct calculation shows that the scalar part of $\dot{q}_t q_t^*$ is null, therefore $\dot{q}_t q_t^*$ can be considered as a vector of $\mathbb{R}^3$. Moreover the commutator of two vectors assimilated to their quaternion counterparts is equal to twice their vector product: $[\mathbf{a}, \mathbf{b}] = 2\,(\mathbf{a} \times \mathbf{b})$ which means that $\dot{q}_t q_t^*$ is equal to $\frac{1}{2}\boldsymbol{\omega}^0$. Therefore, the kinematic equation in these variables simply is

$$\dot{q}_t = \frac{1}{2}\boldsymbol{\omega}^0 q_t \,. \tag{18}$$

Moreover, substituting $\boldsymbol{\omega}^0$ for $q_t\, \boldsymbol{\omega}^1\, q_t^*$ in the previous equation, one get the kinematic equation written in the body frame $(\overrightarrow{\mathrm{f}_1}, \overrightarrow{\mathrm{f}_2}, \overrightarrow{\mathrm{f}_3})$, namely,

$$\dot{q}_t = \frac{1}{2} q_t\, \boldsymbol{\omega}^1 \,. \tag{19}$$

3.4 *Kinematic Equation Satisfied by a Unit Vector of the Figure Axis*

In the case of an axisymmetric body, the rotation is parameterised by the unit vector $\overrightarrow{\mathrm{f}_3}$ along the figure axis. Given that $\overrightarrow{\mathrm{f}_3}$ belongs to $\mathcal{B}$, its evolution with respect to $(\overrightarrow{\mathrm{e}_1}, \overrightarrow{\mathrm{e}_2}, \overrightarrow{\mathrm{e}_3})$ is, by definition of the angular velocity,

$$\dot{\mathbf{f}}_3^0 = \boldsymbol{\omega}^0 \times \mathbf{f}_3^0 \,, \tag{20}$$

where $\mathbf{f}_3^0$ are the coordinates of $\overrightarrow{\mathrm{f}_3}$ in $(\overrightarrow{\mathrm{e}_1}, \overrightarrow{\mathrm{e}_2}, \overrightarrow{\mathrm{e}_3})$.

4 Least Action Principle and Dynamical Equations

Let us consider the Lagrangian $\mathcal{L}(\mathbf{r}^\alpha, \mathsf{R}, \mathbf{v}^\alpha, \boldsymbol{\omega}^\alpha)$, with $\alpha \in \{0, 1\}$, describing the evolution of a rigid body. Here the matrix R is written as a variable of the Lagrangian to recall that the problem depends on the rotation of $\mathcal{B}$. But as explained above, the rotation can be parameterised by other variables and the Lagrangian itself does not necessarily depend explicitly on R. Our purpose in this section is to derive the *dynamical equations* of the system—because calculations are very similar in both frames $\mathcal{R}_0$ and $\mathcal{R}_1$, we only explicit the method in $\mathcal{R}_0$. In the standard Lagrangian formalism $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$ these are the well-known Euler-Lagrange equations given by $\mathrm{d}_t \partial_{\dot{\mathbf{q}}} \mathcal{L} = \partial_{\mathbf{q}} \mathcal{L}$. Rotations modify the structure of these equations which are then called Poincaré-Lagrange equations after [30]. As we will see, one does not need to know in which coordinates the Lagrangian is written to derive the dynamical equations of our problem.

4.1 Parametrisation of the Tangent Space

We introduce a path $\gamma\colon \mathbb{R} \to \mathbb{R}^3 \times \mathrm{SO}(3)$ which assigns to a time t an element $(\mathbf{r}^0(t), \mathsf{R}(t))$ of the Euclidean group. As previously seen, there exist two vectors $\mathbf{v}^0(t)$ and $\boldsymbol{\omega}^0(t)$ of $\mathbb{R}^3$ such that the time derivative of $\gamma(t)$ is given by

$$\dot{\mathbf{r}}^0(t) = \mathbf{v}^0(t)\,, \qquad \dot{\mathsf{R}}(t) = \hat{\boldsymbol{\omega}}^0(t)\mathsf{R}(t)\,. \tag{21}$$

It is thus natural to parameterise an infinitesimal displacement $\delta\gamma(t)$ on the tangent space of the Euclidean group by two vectors $\delta\mathbf{r}^0(t)$ and $\delta\boldsymbol{\theta}^0(t)$ of $\mathbb{R}^3$ such that

$$\delta\gamma(t) = (\ \delta\mathbf{r}^0(t)\,,\ \delta\hat{\boldsymbol{\theta}}^0(t)\mathsf{R}(t)\ )\,. \tag{22}$$

We now impose that the infinitesimal variation of $\dot{\gamma}$ (Eq. 21) is equal to the time derivative of the infinitesimal displacement $\delta\gamma$ (Eq. 22). This step allows to relate the variation of the velocities $(\delta\mathbf{v}^0(t), \delta\boldsymbol{\omega}^0(t))$ to the evolution rate of the displacement vectors $(\delta\dot{\mathbf{r}}^0(t), \delta\dot{\boldsymbol{\theta}}^0(t))$. We get

$$\delta\mathbf{v}^0(t) = \delta\dot{\mathbf{r}}^0(t)\,, \qquad \delta\boldsymbol{\omega}^0(t) = \delta\dot{\boldsymbol{\theta}}^0(t) + \delta\boldsymbol{\theta}^0(t) \times \boldsymbol{\omega}^0(t)\,. \tag{23}$$

4.2 Variation of the Action

We now are in position to differentiate the action $\mathcal{S}(\gamma)$ associated with the Lagrangian $\mathcal{L}$. The action reads

$$\mathcal{S}(\gamma) = \int_{t_0}^{t_1} \mathcal{L}(\gamma, \dot{\gamma})\,\mathrm{d}t\,. \tag{24}$$

Its variation $\delta\mathcal{S}(\gamma) = \mathcal{S}(\gamma + \delta\gamma) - \mathcal{S}(\gamma)$ in the vicinity of γ is

$$\begin{aligned}\delta\mathcal{S}(\gamma) &= \int_{t_0}^{t_1} \left(\frac{\partial\mathcal{L}}{\partial\mathbf{v}^0}\cdot\delta\mathbf{v}^0 + \frac{\partial\mathcal{L}}{\partial\mathbf{r}^0}\cdot\delta\mathbf{r}^0 + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^0}\cdot\delta\boldsymbol{\omega}^0 + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}^0}\cdot\delta\boldsymbol{\theta}^0 \right) \mathrm{d}t \\ &= \int_{t_0}^{t_1} \left(\frac{\partial\mathcal{L}}{\partial\mathbf{v}^0}\cdot\delta\dot{\mathbf{r}}^0 + \frac{\partial\mathcal{L}}{\partial\mathbf{r}^0}\cdot\delta\mathbf{r}^0 + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^0}\cdot(\delta\dot{\boldsymbol{\theta}}^0 + \delta\boldsymbol{\theta}^0\times\boldsymbol{\omega}^0) + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}^0}\cdot\delta\boldsymbol{\theta}^0 \right) \mathrm{d}t\,.\end{aligned} \tag{25}$$

After an integration by part to remove $\delta\dot{\mathbf{r}}^0$ and $\delta\dot{\boldsymbol{\theta}}^0$ and a circular permutation in the triple product $\partial_{\boldsymbol{\omega}^0}\mathcal{L}\cdot(\delta\boldsymbol{\theta}^0\times\boldsymbol{\omega}^0)$, we obtain

$$\delta\mathcal{S}(\gamma) = \left[\frac{\partial\mathcal{L}}{\partial\mathbf{v}^0}\cdot\delta\mathbf{r}^0 + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^0}\cdot\delta\boldsymbol{\theta}^0\right]_{t_0}^{t_1} - \int_{t_0}^{t_1}\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\mathbf{v}^0} - \frac{\partial\mathcal{L}}{\partial\mathbf{r}^0}\right)\cdot\delta\mathbf{r}^0 + \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^0} - \boldsymbol{\omega}\times\frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^0} - \frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}^0}\right)\cdot\delta\boldsymbol{\theta}^0\right]\mathrm{d}t\,. \tag{26}$$

4.3 Dynamical Equations

The principle of least action tells us that the path $t \mapsto \gamma(t)$ effectively followed by the system is the one satisfying $\delta\mathcal{S}(\gamma) = 0$ for any infinitesimal displacement $\delta\gamma$ with $\delta\gamma(t_0) = \delta\gamma(t_1) = 0$. Therefore, the equations of motion deduced from (26) are

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\mathbf{v}^0} = \frac{\partial\mathcal{L}}{\partial\mathbf{r}^0}\,, \tag{27}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^0} = \boldsymbol{\omega}^0\times\frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^0} + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}^0}\,. \tag{28}$$

The first of these two equations is the common Euler-Lagrange equation. It relates the evolution rate of the linear momentum $\mathbf{p}^0 = \partial_{\mathbf{v}^0}\mathcal{L}$ to the force $\mathbf{F}^0$ equal to the variation of the Lagrangian under an infinitesimal translation $\partial_{\mathbf{r}^0}\mathcal{L}$. The second is called Poincaré-Lagrange equation after Poincaré's 1901 article "*Sur une forme nouvelle des équations de la Mécanique*" [30] (see also Sect. 2 of [5]). It links the evolution rate of the angular momentum $\boldsymbol{\pi}^0 = \partial_{\boldsymbol{\omega}^0}\mathcal{L}$ to the torque $\mathbf{T}^0$ which is the variation of the Lagrangian under an infinitesimal rotation $\partial_{\boldsymbol{\theta}^0}\mathcal{L}$. The Poincaré-Lagrange equation also contains an inertial torque of the form $\boldsymbol{\omega}^0\times\partial_{\boldsymbol{\omega}^0}\mathcal{L} = \boldsymbol{\omega}^0\times\boldsymbol{\pi}^0$.

Were the dynamical equations derived in the body-fixed frame $\mathcal{R}_1$, we would have reached

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\mathbf{v}^1} = \frac{\partial\mathcal{L}}{\partial\mathbf{r}^1}\,, \tag{29}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^1} = \frac{\partial\mathcal{L}}{\partial\boldsymbol{\omega}^1}\times\boldsymbol{\omega}^1 + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}^1}\,. \tag{30}$$

4.4 Rayleigh Dissipation Function

Dissipative systems can be studied within the Lagrangian formalism. Let $\mathcal{L}(\mathbf{q},\dot{\mathbf{q}})$ be a Lagrangian written in terms of standard coordinates and velocities. Dissipation is modelled by a function $\mathcal{D}$ called *Rayleigh dissipation function* of the form $\mathcal{D} = \frac{1}{2}\eta\|\dot{\mathbf{q}}\|^2$. The associated equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \frac{\partial \mathcal{L}}{\partial \mathbf{q}} - \frac{\partial \mathcal{D}}{\partial \dot{\mathbf{q}}}\,. \tag{31}$$

4.5 Spin Operator

The operators $\partial_{\boldsymbol{\theta}^0}$ and $\partial_{\boldsymbol{\theta}^1}$ are well-known in quantum theory of angular momentum. These are the *spin operator*, usually denoted by $\hat{S}$ or $\hat{J}$, with $\partial_{\boldsymbol{\theta}^0}$ being expressed in the inertial frame $\mathcal{R}_0$ and $\partial_{\boldsymbol{\theta}^1}$ in the rotated frame $\mathcal{R}_1$. In particular, its action on Wigner's D-matrices (used to rotate spherical harmonics) is described in many textbooks such as [36]. This can be particularly useful for studying the evolution of binary asteroids with highly non-spherical shapes (e.g., [4]). In this course we limit ourselves to simple problems for which $\partial_{\boldsymbol{\theta}}\mathcal{L}$ can easily be computed. All calculations are based on the proposition that under an infinitesimal rotation of angle $\delta\boldsymbol{\theta}$ a vector $\mathbf{v}$ is transformed into a vector $\mathbf{v} + \delta\mathbf{v}$ with $\delta\mathbf{v} = \delta\boldsymbol{\theta} \times \mathbf{v}$.

4.6 Hamiltonian Formalism

For conservative systems, such as in the rigid body case, it can be convenient to switch to the Hamiltonian formalism. In that case, we define the linear momentum $\mathbf{p}^0$ and the angular momentum $\boldsymbol{\pi}^0$ in $\mathcal{R}_0$ as follows

$$\mathbf{p}^0 = \frac{\partial \mathcal{L}}{\partial \mathbf{v}^0}\,, \qquad \boldsymbol{\pi}^0 = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\omega}^0}\,. \tag{32}$$

The Hamiltonian $\mathcal{H}$ is then obtained by a *Legendre transformation* of the Lagrangian $\mathcal{H} = \mathbf{p}^0 \cdot \mathbf{v}^0 + \boldsymbol{\pi}^0 \cdot \boldsymbol{\omega}^0 - \mathcal{L}$. Taking into account the definition of $\mathbf{p}^0$ and $\boldsymbol{\pi}^0$, the infinitesimal variation of $\mathcal{H}$ reads

$$\delta\mathcal{H} = \mathbf{v}^0 \cdot \delta\mathbf{p}^0 + \boldsymbol{\omega}^0 \cdot \delta\boldsymbol{\pi}^0 - \frac{\partial \mathcal{L}}{\partial \mathbf{r}^0} \cdot \delta\mathbf{r}^0 - \frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}^0} \cdot \delta\boldsymbol{\theta}^0\,. \tag{33}$$

The Hamiltonian is thus a function of $(\mathbf{p}^0, \boldsymbol{\pi}^0, \mathbf{r}^0, \mathsf{R})$ whose variation is

$$\delta\mathcal{H} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}^0} \cdot \delta\mathbf{p}^0 + \frac{\partial \mathcal{H}}{\partial \boldsymbol{\pi}^0} \cdot \delta\boldsymbol{\pi}^0 + \frac{\partial \mathcal{H}}{\partial \mathbf{r}^0} \cdot \delta\mathbf{r}^0 + \frac{\partial \mathcal{H}}{\partial \boldsymbol{\theta}^0} \cdot \delta\boldsymbol{\theta}^0\,. \tag{34}$$

Identifying the two expressions, we get

$$\mathbf{v}^0 = \frac{\partial \mathcal{H}}{\partial \mathbf{p}^0}\,, \qquad \boldsymbol{\omega}^0 = \frac{\partial \mathcal{H}}{\partial \boldsymbol{\pi}^0}\,, \qquad \frac{\partial \mathcal{L}}{\partial \mathbf{r}^0} = -\frac{\partial \mathcal{H}}{\partial \mathbf{r}^0}\,, \qquad \frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}^0} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{\theta}^0}\,. \tag{35}$$

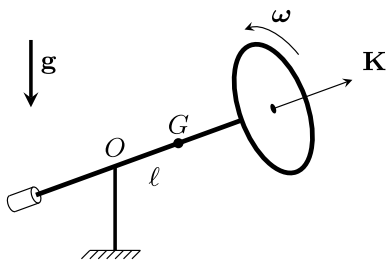


Fig. 2 Gyroscope of mass m rotating at the angular speed ω around its figure axis $\mathbf{K}$ of maximum inertia J. Its barycentre is at G. The gyroscope, subject to the uniform gravity field $\mathbf{g}$, is free to rotate about the attachment point O. The distance OG is denoted by ℓ

Substituting these formulas in the kinematic and dynamical equations, we obtain the Poincaré-Hamilton equations

$$\dot{\mathbf{r}}^0 = \frac{\partial \mathcal{H}}{\partial \mathbf{p}^0}\,, \qquad \dot{\mathbf{p}}^0 = -\frac{\partial \mathcal{H}}{\partial \mathbf{r}^0}\,, \qquad \dot{\mathsf{R}} = \left(\frac{\partial \hat{\mathcal{H}}}{\partial \boldsymbol{\pi}^0}\right)\mathsf{R}\,, \qquad \dot{\boldsymbol{\pi}}^0 = \frac{\partial \mathcal{H}}{\partial \boldsymbol{\pi}^0} \times \boldsymbol{\pi}^0 - \frac{\partial \mathcal{H}}{\partial \boldsymbol{\theta}^0}\,. \tag{36}$$

For completeness, we provide the same equations written in the body-fixed frame $\mathcal{R}_1$

$$\dot{\mathbf{r}}^1 = \frac{\partial \mathcal{H}}{\partial \mathbf{p}^1}\,, \qquad \dot{\mathbf{p}}^1 = -\frac{\partial \mathcal{H}}{\partial \mathbf{r}^1}\,, \qquad \dot{\mathsf{R}} = \mathsf{R}\left(\frac{\partial \hat{\mathcal{H}}}{\partial \boldsymbol{\pi}^1}\right), \qquad \dot{\boldsymbol{\pi}}^1 = \boldsymbol{\pi}^1 \times \frac{\partial \mathcal{H}}{\partial \boldsymbol{\pi}^1} - \frac{\partial \mathcal{H}}{\partial \boldsymbol{\theta}^1}\,. \tag{37}$$

4.7 Example: The Gyroscope

We consider the gyroscope drawn in Fig. 2. It is assumed to rotate rapidly around its axis of maximum inertia $\mathbf{K}$ (here all vectors are written in the inertial frame $\mathcal{R}_0$). We denote by $\boldsymbol{\omega} = \omega\mathbf{K}$ its angular velocity and by J its moment of inertia along $\mathbf{K}$. Let m be its mass centred on the barycentre G shifted from the attachment point O by a distance ℓ along the figure axis. We set $\boldsymbol{\ell} = \ell\mathbf{K}$. The gyroscope evolves in the gravity field $\mathbf{g}$ of the Earth. For this system, the Lagrangian reads

$$\mathcal{L} = \frac{1}{2} J\omega^2 + m\,\mathbf{g}\cdot\boldsymbol{\ell}\,.$$

The angular momentum is $\boldsymbol{\pi} = \partial_\omega \mathcal{L} = J\boldsymbol{\omega}$. To get the action of the spin operator $\partial_{\boldsymbol{\theta}}$ on the Lagrangian $\mathcal{L}$ we compute its variation under an infinitesimal rotation of vector $\delta\boldsymbol{\theta}$. Under this transformation only $\boldsymbol{\ell}$ evolves and becomes $\boldsymbol{\ell} + \delta\boldsymbol{\theta} \times \boldsymbol{\ell}$. Therefore

$$\delta\mathcal{L} = m\,\mathbf{g}\cdot\delta\boldsymbol{\ell} = m\,\mathbf{g}\cdot(\delta\boldsymbol{\theta}\times\boldsymbol{\ell}) = m\,(\boldsymbol{\ell}\times\mathbf{g})\cdot\delta\boldsymbol{\theta} \qquad \Rightarrow \qquad \frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}} = m\,(\boldsymbol{\ell}\times\mathbf{g})\,.$$

Moreover, $\boldsymbol{\omega}$ and $\boldsymbol{\pi}$ being colinear, their product $\boldsymbol{\omega} \times \boldsymbol{\pi}$ is null. Then the Poincaré-Lagrange equations (28) give

$$\dot{\boldsymbol{\pi}} = m\,(\boldsymbol{\ell} \times \mathbf{g}) = -\boldsymbol{\omega}_p \times \boldsymbol{\pi}\,,$$

where $\boldsymbol{\omega}_p = m\ell\mathbf{g}/(J\omega)$. This is the well-known equation of motion for a gyroscope with $\boldsymbol{\omega}_p$ being the precession frequency.

5 Lagrangian of a Rigid Body Interacting with a Point Mass

Here we consider a rigid body $\mathcal{B}$ of mass m interacting with a mass m_0. We wish to follow the evolution of the system with respect to the inertial frame $\mathcal{R}_0$ of origin O and basis $(\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3})$. As above, we attach to $\mathcal{B}$ a frame $\mathcal{R}_1$ centred on O′ with unit vectors $(\overrightarrow{f_1}, \overrightarrow{f_2}, \overrightarrow{f_3})$. We denote by M the position of the mass m_0. Here again, we choose to study the problem in $\mathcal{R}_0$, therefore all vectors are written in this frame, but for clarity we drop the index 0.

5.1 Kinematic Energy

By definition, the kinematic energy T of the whole system is the sum of the individual kinematic energies of each mass element. Therefore,

$$T = \frac{1}{2} m_0 v_{\mathrm{M}}^2 + \int_{\mathcal{B}} \frac{1}{2} v_{\mathrm{P}}^2 \,\mathrm{d}m(\mathrm{P})\,. \tag{38}$$

Moreover, if we set $\mathbf{v}_{\mathrm{O}'} = \mathrm{dOO}'/\mathrm{d}t$ and denote by $\boldsymbol{\omega}$ the angular velocity of $\mathcal{B}$, the velocity of a point $\mathrm{P} \in \mathcal{B}$ with respect to $\mathcal{R}_0$ is

$$\mathbf{v}_{\mathrm{P}} = \mathbf{v}_{\mathrm{O}'} + \boldsymbol{\omega} \times \mathrm{O}'\mathrm{P}\,. \tag{39}$$

Substituting the expression of $\mathbf{v}_{\mathrm{P}}$ in that of the kinematic energy T, we get

$$T = \frac{1}{2} m_0 v_{\mathrm{M}}^2 + \frac{1}{2} m v_{\mathrm{O}'}^2 + \mathbf{v}_{\mathrm{O}'} \cdot \left(\boldsymbol{\omega} \times \int_{\mathcal{B}} \mathrm{O}'\mathrm{P}\,\mathrm{d}m(\mathrm{P}) \right) + \frac{1}{2} \int_{\mathcal{B}} \|\boldsymbol{\omega} \times \mathrm{O}'\mathrm{P}\|^2 \,\mathrm{d}m(\mathrm{P})\,. \tag{40}$$

The integral of $\mathrm{O}'\mathrm{P}\,\mathrm{d}m(\mathrm{P})$ vanishes when O′ is chosen to be the barycentre of $\mathcal{B}$. This indeed is a common practice for this kind of problem because it is a natural choice and it simplifies computations. But one could choose the geometric centre instead and keep this term. In our case, we hereafter assume that O′ is the barycentre of $\mathcal{B}$.

For the last integral of (40), we use $\|\boldsymbol{\omega} \times \mathrm{O'P}\|^2 = \omega^2 \|\mathrm{O'P}\|^2 - (\boldsymbol{\omega} \cdot \mathrm{O'P})^2$. Then

$$T = \frac{1}{2} m_0 v_{\mathrm{M}}^2 + \frac{1}{2} m v_{\mathrm{O'}}^2 + \frac{1}{2} \boldsymbol{\omega} \cdot \mathsf{I}' \boldsymbol{\omega} \,, \tag{41}$$

where I' is the *inertia matrix* of $\mathcal{B}$ written in the frame $(\vec{e_1}, \vec{e_2}, \vec{e_3})$ and whose expression is

$$\begin{aligned} \mathsf{I}' &= \int_{\mathcal{B}} \left(\|\mathrm{O'P}\|^2 \, \mathbb{I} - \mathrm{O'P} \otimes \mathrm{O'P} \right) \mathrm{d}m(\mathrm{P}) \\ &= \begin{bmatrix} \int (y^2 + z^2) \, \mathrm{d}m & -\int xy \, \mathrm{d}m & -\int xz \, \mathrm{d}m \\ -\int xy \, \mathrm{d}m & \int (x^2 + z^2) \, \mathrm{d}m & -\int yz \, \mathrm{d}m \\ -\int xz \, \mathrm{d}m & -\int yz \, \mathrm{d}m & \int (x^2 + y^2) \, \mathrm{d}m \end{bmatrix} , \end{aligned} \tag{42}$$

with (x, y, z) the coordinates of $\mathrm{O'P}$ in $\mathcal{R}_0$. This matrix is symmetric. Usually, the basis vectors $(\vec{f_1}, \vec{f_2}, \vec{f_3})$ of the frame $\mathcal{R}_1$ are chosen so as to diagonalise the inertia matrix. In that case, $\vec{f_1}$, $\vec{f_2}$ and $\vec{f_3}$ are called the *principal axes of inertia* of $\mathcal{B}$ and the diagonal terms its *principal moments of inertia*. Let us denote by I the inertia matrix written in $(\vec{f_1}, \vec{f_2}, \vec{f_3})$. We have $\mathsf{I}' = \mathsf{R}\,\mathsf{I}\,\mathsf{R}^{\top}$.

Regarding the translational motion represented by the vectors $\mathbf{v}_{\mathrm{M}}$ and $\mathbf{v}_{\mathrm{O'}}$, it is often more convenient to express it in terms of the velocity $\mathbf{v}_{\mathrm{G}} = (m_0 \mathbf{v}_{\mathrm{M}} + m \mathbf{v}_{\mathrm{O'}})/(m + m_0)$ of the system barycentre G and of the relative velocity $\mathbf{v} = \mathbf{v}_{\mathrm{M}} - \mathbf{v}_{\mathrm{O'}}$. We get

$$T = \frac{1}{2}(m_0 + m) v_{\mathrm{G}}^2 + \frac{1}{2} \beta v^2 + \frac{1}{2} \boldsymbol{\omega} \cdot \mathsf{I}' \boldsymbol{\omega} \,, \tag{43}$$

where $\beta = m_0 m/(m_0 + m)$ is the *reduced mass*.

5.2 Potential Energy

We only consider the Newtonian interaction between $\mathcal{B}$ and the mass m_0. Therefore, the potential energy U reads

$$U = -\mathcal{G} m_0 \int_{\mathcal{B}} \frac{\mathrm{d}m(\mathrm{P})}{\|\mathrm{MP}\|} \,, \tag{44}$$

where $\mathcal{G} = 6.6743 \times 10^{-11}\ \mathrm{m^3 \cdot kg^{-1} \cdot s^{-2}}$ is the gravitational constant. Let $\mathbf{r}$ be the vector $\mathrm{O'M}$ expressed in $(\vec{e_1}, \vec{e_2}, \vec{e_3})$ and whose time derivative is $\dot{\mathbf{r}} = \mathbf{v}$. We still use the vector $\mathbf{y} = \mathrm{O'P}$ written in the same basis. In terms of these two vectors $\mathrm{MP} = \mathbf{y} - \mathbf{r}$. We now assume that the distance r between the two objects is much larger than the size $\max(y)$ of $\mathcal{B}$ and perform a Taylor expansion of U in the small parameter y/r. Let us introduce Legendre polynomials $P_n(X)$ such that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ with

$y < x$ and $\mathbf{x} \cdot \mathbf{y} = xy \cos S$,

$$\frac{1}{\|\mathbf{x} - \mathbf{y}\|} = \sum_{n=0}^{\infty} \frac{y^n}{x^{n+1}} P_n(\cos S) \,. \tag{45}$$

The first polynomials are

$$P_0(X) = 1 \,, \quad P_1(X) = X \,, \quad P_2(X) = \frac{1}{2}(3X^2 - 1) \,, \quad P_3(X) = \frac{1}{2}(5X^3 - 3X) \,. \tag{46}$$

Here, we limit the expansion to degrees $n \leq 2$ which is referred to as the *quadrupole* approximation. Under this simplification, the potential energy becomes

$$U = -\frac{\mathcal{G}mm_0}{r} - \frac{\mathcal{G}m_0}{r^3} \left(\mathbf{r} \cdot \int_{\mathcal{B}} \mathbf{y} \, \mathrm{d}m(\mathrm{P}) \right) - \frac{\mathcal{G}m_0}{2r^5} \left(\mathbf{r} \cdot \int_{\mathcal{B}} \left(3\, \mathbf{y} \otimes \mathbf{y} - y^2 \, \mathbb{I} \right) \mathrm{d}m(\mathrm{P}) \, \mathbf{r} \right) . \tag{47}$$

As in the previous section, O′ is supposed to be the barycentre of $\mathcal{B}$, therefore the first integral vanishes. The second integral can be expressed in terms of the inertia matrix I'. We get

$$U = -\frac{\mathcal{G}mm_0}{r} + \frac{3}{2}\frac{\mathcal{G}m_0}{r^5} \, \mathbf{r} \cdot \left(\mathsf{I}' - \frac{\mathrm{Tr}(\mathsf{I}')}{3} \, \mathbb{I} \right) \mathbf{r} \,. \tag{48}$$

In this expression, the first term is the usual gravitational interaction between two point masses. The second term represents the effect of the triaxiality of $\mathcal{B}$ and couples the orbit $\mathbf{r}$ to the spin through the rotation matrix in $\mathsf{I}' = \mathsf{R}\,\mathsf{I}\,\mathsf{R}^{\top}$.

5.3 Lagrangian and Hamiltonian of the System

The Lagrangian $\mathcal{L} = T - U$ of the system reads

$$\mathcal{L} = \frac{1}{2}(m_0 + m)v_{\mathrm{G}}^2 + \frac{1}{2}\beta v^2 + \frac{1}{2}\boldsymbol{\omega} \cdot \mathsf{I}'\boldsymbol{\omega} + \frac{\mathcal{G}mm_0}{r} - \frac{3}{2}\frac{\mathcal{G}m_0}{r^5} \, \mathbf{r} \cdot \left(\mathsf{I}' - \frac{\mathrm{Tr}(\mathsf{I}')}{3}\mathbb{I} \right) \mathbf{r} \,. \tag{49}$$

As expected, the system being isolated, the Lagrangian is independent of the coordinates $\mathbf{r}_{\mathrm{G}}$ of the barycentre G. The variable $\mathbf{r}_{\mathrm{G}}$ is said to be *cyclic*. The associated momentum $\partial_{\mathbf{v}_{\mathrm{G}}}\mathcal{L} = (m_0 + m)\mathbf{v}_{\mathrm{G}}$ is conserved. Hereafter we discard this degree of freedom or, equivalently, we assume that $\mathbf{v}_{\mathrm{G}} = \mathbf{0}$ and consider $\mathcal{L}$ as being only a function of $(\mathbf{r}, \mathbf{v}, \mathsf{R}, \boldsymbol{\omega})$. We thus write

$$\mathcal{L}(\mathbf{r}, \mathbf{v}, \mathsf{R}, \boldsymbol{\omega}) = \frac{1}{2}\beta v^2 + \frac{1}{2}\boldsymbol{\omega} \cdot \mathsf{I}'\boldsymbol{\omega} + \frac{\mathcal{G}mm_0}{r} - \frac{3}{2}\frac{\mathcal{G}m_0}{r^5} \, \mathbf{r} \cdot \left(\mathsf{I}' - \frac{\mathrm{Tr}(\mathsf{I}')}{3}\mathbb{I} \right) \mathbf{r} \,. \tag{50}$$

In this course, we shall eventually neglect the effect of the quadrupole on the orbit to focus on the rotation motion of $\mathcal{B}$. The orbital evolution is studied in, e.g., [18] (see also [6]). Neglecting the perturbation on the orbit remains valid whenever the angular momentum of rotation $\pi = \|\mathsf{I}'\boldsymbol{\omega}\|$ is much smaller than the orbital angular momentum $\beta\|\mathbf{r} \times \mathbf{v}\|$. This is particularly true when $\mathcal{B}$ is a moon orbiting a planet or a planet orbiting a star. In that case the position vector $\mathbf{r}$ can be seen as a function of time $\mathbf{r}(t)$. Furthermore, the orbit might not even be Keplerian due to interactions with other bodies not accounted for in (50), e.g., other moons or other planets. In this situation, the Lagrangian becomes non-autonomous and reads

$$\mathcal{L}(\mathsf{R}, \boldsymbol{\omega}, t) = \frac{1}{2}\boldsymbol{\omega} \cdot \mathsf{I}'\boldsymbol{\omega} - \frac{3}{2}\frac{\mathcal{G}m_0}{r^5}\,\mathbf{r} \cdot \mathsf{I}'\mathbf{r}\,. \tag{51}$$

For the sake of conciseness, we dropped the time t and wrote $\mathbf{r}$ instead of $\mathbf{r}(t)$. The corresponding Hamiltonian written in terms of $\boldsymbol{\pi} = \mathsf{I}'\boldsymbol{\omega}$ is

$$\mathcal{H}(\mathsf{R}, \boldsymbol{\pi}, t) = \frac{1}{2}\boldsymbol{\pi} \cdot \mathsf{I}'^{-1}\boldsymbol{\pi} + \frac{3}{2}\frac{\mathcal{G}m_0}{r^5}\,\mathbf{r} \cdot \mathsf{I}'\mathbf{r}\,. \tag{52}$$

5.4 Equations of Motion

To get the Poincaré-Lagrange equations for the Lagrangian (51) (or the Poincaré-Hamilton equations for the Hamiltonian (52)), we shall first notice that the variation of $\mathsf{I}' = \mathsf{R}\,\mathsf{I}\,\mathsf{R}^\top$ under an infinitesimal rotation of vector $\delta\boldsymbol{\theta}$ is[1]

$$\delta\mathsf{I}' = [\delta\hat{\boldsymbol{\theta}}, \mathsf{I}'] := \delta\hat{\boldsymbol{\theta}}\,\mathsf{I}' - \mathsf{I}'\,\delta\hat{\boldsymbol{\theta}}\,, \tag{53}$$

with $\delta\hat{\boldsymbol{\theta}}$ the image of $\delta\boldsymbol{\theta}$ by the hat operator (10). Furthermore, for any vector $\mathbf{x} \in \mathbb{R}^3$, we have

$$\mathbf{x} \cdot [\delta\hat{\boldsymbol{\theta}}, \mathsf{I}']\,\mathbf{x} = 2\delta\boldsymbol{\theta} \cdot (\mathsf{I}'\mathbf{x} \times \mathbf{x})\,, \tag{54}$$

implying $\partial_{\boldsymbol{\theta}}(\mathbf{x} \cdot \mathsf{I}'\mathbf{x}) = 2(\mathsf{I}'\mathbf{x} \times \mathbf{x})$. Therefore, the equations of motion are

[1] Let us detail the calculation. Under any rotation of $\mathcal{B}$, the matrix of inertia I expressed in the body frame $\mathcal{R}_1$ remains constant. Therefore,

$$\begin{aligned}\delta\mathsf{I}' &= \delta\mathsf{R}\,\mathsf{I}\,\mathsf{R}^\top + \mathsf{R}\,\mathsf{I}\,\delta\mathsf{R}^\top \\ &= \delta\mathsf{R}\,\mathsf{R}^\top\,\mathsf{R}\,\mathsf{I}\,\mathsf{R}^\top + \mathsf{R}\,\mathsf{I}\,\mathsf{R}^\top\,\mathsf{R}\,\delta\mathsf{R}^\top \\ &= (\delta\mathsf{R}\,\mathsf{R}^\top)\,\mathsf{I}' + \mathsf{I}'(\delta\mathsf{R}\,\mathsf{R}^\top)^\top\,.\end{aligned}$$

By definition, $\delta\hat{\boldsymbol{\theta}} := \delta\mathsf{R}\,\mathsf{R}^\top$ is skew symmetric, hence the result

$$\delta\mathsf{I}' = \delta\hat{\boldsymbol{\theta}}\,\mathsf{I}' - \mathsf{I}'\,\delta\hat{\boldsymbol{\theta}}\,.$$

.

$$\dot{\boldsymbol{\pi}} = -3\frac{\mathcal{G}m_0}{r^5}(\mathsf{I}'\mathbf{r}) \times \mathbf{r}\,, \qquad \dot{\mathsf{R}} = \hat{\boldsymbol{\omega}}\mathsf{R} \tag{55}$$

with $\boldsymbol{\pi} = \partial_{\boldsymbol{\omega}}\mathcal{L} = \mathsf{I}'\boldsymbol{\omega}$. In practice, in numerical integrations, the state vector is composed of the rotation matrix R (or any equivalent representation of SO(3)) and the angular momentum $\boldsymbol{\pi}$. At each time step the inertia matrix I' is computed from $\mathsf{R}\,\mathsf{I}\,\mathsf{R}^\top$ and $\boldsymbol{\omega}$ is given by $\mathsf{I}'^{-1}\boldsymbol{\pi}$.

6 Libration in the Vicinity of the Synchronous State

In this section, we assume that $\mathcal{B}$ is close to the 1:1 spin-orbit resonance state such that $\omega \approx n$ where n is the orbital mean motion. The orbit is represented by its elliptical elements $a, e, i, \lambda, \varpi, \Omega$ which are respectively the semimajor axis, the eccentricity, the inclination with respect to $\mathcal{R}_0$, the mean longitude, the longitude of the periapsis, and the longitude of the ascending node. We allow the angles ϖ and Ω to vary linearly with time (due to the flattening of the central body m_0 or to the presence of additional bodies in the system). The mean longitude λ is also assumed to vary linearly with time. Here we define the orbital mean motion as $n = \dot{\lambda}$ but we also use Kepler's third law $n^2a^3 = \mathcal{G}(m + m_0)$ to approximate its value (the two definitions are not equivalent). Moreover, we further assume that $m_0 \gg m$ such that $n^2a^3 \approx \mathcal{G}m_0$. These two simplifications are made with the sole purpose to avoid cumbersome expressions.

6.1 Rotating Frame

To study this specific problem, we place ourselves in a new reference frame $\mathcal{R}'_0$ of basis $(\overrightarrow{e_1}', \overrightarrow{e_2}', \overrightarrow{e_3}')$ rotating at the angular velocity $\boldsymbol{\omega}_{\mathcal{R}'_0} = n\overrightarrow{e_3}$ with respect to the inertial frame $(\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3})$. For conciseness, we keep the same notation as above although the meaning of the variables has changed. The matrix R now represents the rotation of $\mathcal{R}_1$ relative to $\mathcal{R}'_0$ and $\boldsymbol{\omega}$ is the angular velocity of $\mathcal{B}$ with respect to $\mathcal{R}'_0$, i.e., $\boldsymbol{\omega}$ is still defined by the equality $\dot{\mathsf{R}}\mathsf{R}^\top = \hat{\boldsymbol{\omega}}$. The inertia matrix $\mathsf{I}' = \mathsf{R}\,\mathsf{I}\,\mathsf{R}^\top$ and the position vector $\mathbf{r}$ are now written in $\mathcal{R}'_0$. The rotation speed of $\mathcal{B}$ relative to $\mathcal{R}_0$ is then $\boldsymbol{\omega}_{\mathcal{R}'_0} + \boldsymbol{\omega}$. In this new reference frame, the Lagrangian is thus

$$\mathcal{L} = \frac{1}{2}(\boldsymbol{\omega}_{\mathcal{R}'_0} + \boldsymbol{\omega}) \cdot \mathsf{I}'(\boldsymbol{\omega}_{\mathcal{R}'_0} + \boldsymbol{\omega}) - \frac{3}{2}\frac{\mathcal{G}m_0}{r^5}\,\mathbf{r} \cdot \mathsf{I}'\mathbf{r}\,. \tag{56}$$

According to (54) we have

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}} = \mathsf{I}'(\boldsymbol{\omega}_{\mathcal{R}'_0} + \boldsymbol{\omega}) \times (\boldsymbol{\omega}_{\mathcal{R}'_0} + \boldsymbol{\omega}) - 3\frac{\mathcal{G}m_0}{r^5}(\mathsf{I}'\mathbf{r}) \times \mathbf{r}\,. \tag{57}$$

This equation will be used to infer the Poincaré-Lagrange equation (28), which here reads as $\dot{\boldsymbol{\pi}} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}} + \boldsymbol{\omega} \times \frac{\partial \mathcal{L}}{\partial \boldsymbol{\omega}}$. As for the kinematic equation, let us consider I' as our generalised coordinate instead of R. We get

$$\dot{\boldsymbol{\pi}} = -\boldsymbol{\omega}_{\mathcal{R}'_0} \times \boldsymbol{\pi} - 3\frac{\mathcal{G}m_0}{r^5}(\mathsf{I}'\mathbf{r}) \times \mathbf{r}\,, \qquad \dot{\mathsf{I}}' = [\hat{\boldsymbol{\omega}}, \mathsf{I}']\,, \tag{58}$$

with

$$\boldsymbol{\pi} := \frac{\partial \mathcal{L}}{\partial \boldsymbol{\omega}} = \mathsf{I}'(\boldsymbol{\omega}_{\mathcal{R}'_0} + \boldsymbol{\omega}) \qquad \Rightarrow \qquad \boldsymbol{\omega} = -\boldsymbol{\omega}_{\mathcal{R}'_0} + \mathsf{I}'^{-1}\boldsymbol{\pi}\,. \tag{59}$$

In the following, we assume the orbit to be almost circular with low inclination with respect to the reference plane defined by $\overrightarrow{\mathrm{e}_1}$ and $\overrightarrow{\mathrm{e}_2}$. In the rotated frame, we thus have $\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{r}_1(t)$ with $\mathbf{r}_0 = a\,\mathbf{e}'_1$ and $\|\mathbf{r}_1(t)\| \ll \|\mathbf{r}_0\|$.

6.2 Equilibrium State in the Case of a Circular Orbit with No Inclination

In this subsection, we neglect $\mathbf{r}_1(t)$ and look for a fixed point $(\boldsymbol{\pi}_0, \mathsf{I}'_0)$ of the problem described by the Lagrangian (56). The evolution rate of $\boldsymbol{\pi}$ becomes

$$\dot{\boldsymbol{\pi}} = -\boldsymbol{\omega}_{\mathcal{R}'_0} \times \boldsymbol{\pi} - 3\,n^2\,(\mathsf{I}'\mathbf{e}'_1) \times \mathbf{e}'_1\,. \tag{60}$$

For any point $\mathrm{P} \in \mathcal{B}$ to be at rest in $\mathcal{R}'_0$, $\boldsymbol{\omega}$ must be equal to $\boldsymbol{\omega}_0 = \mathbf{0}$. Hence, Eq. (59) leads to $\boldsymbol{\pi}_0 = \mathsf{I}'_0\,\boldsymbol{\omega}_{\mathcal{R}'_0}$. Substituting $\boldsymbol{\omega}_{\mathcal{R}'_0}$ by its expression $n\,\overrightarrow{\mathrm{e}_3}$ and $\boldsymbol{\pi}$ by its equilibrium value $\boldsymbol{\pi}_0 = \mathsf{I}'_0\,\boldsymbol{\omega}_{\mathcal{R}'_0}$ in (60), and exploiting the equality $\mathbf{e}_3 = \mathbf{e}'_3$, we find that the following condition must be fulfilled at equilibrium

$$\mathbf{e}'_3 \times (\mathsf{I}'_0\mathbf{e}'_3) + 3\,(\mathsf{I}'_0\mathbf{e}'_1) \times \mathbf{e}'_1 = \mathbf{0}\,. \tag{61}$$

For the relation (61) to be satisfied, the symmetric matrix I'_0 has to be diagonal in the frame $\mathcal{R}'_0$. Therefore, $\mathcal{B}$ is at rest in the rotating frame $\mathcal{R}'_0$ if and only if its principal axes of inertia coincide with the base frame vectors $(\overrightarrow{\mathrm{e}_1}', \overrightarrow{\mathrm{e}_2}', \overrightarrow{\mathrm{e}_3}')$. We denote by A, B, C its principal moments of inertia such that $\mathsf{I}'_0 = \mathrm{diag}(A, B, C)$ in this frame.

6.3 Eigenfrequencies

We still discard the perturbation $\mathbf{r}_1(t)$ and look at the motion in the vicinity of the fixed point $(\boldsymbol{\pi}_0, \mathsf{I}'_0)$ found above. For this purpose, we expand $\boldsymbol{\pi}$ and I' to first order of a small parameter $\varepsilon \ll 1$ as $\boldsymbol{\pi} = \boldsymbol{\pi}_0 + \varepsilon\boldsymbol{\pi}_1$ and $\mathsf{I}' = \mathsf{I}'_0 + \varepsilon\mathsf{I}'_1$. In order to remain in the tangent space of the phase space of the problem, we parameterise I'_1 by the

vector $\boldsymbol{\theta}_1 \in \mathbb{R}^3$ such that

$$\mathsf{I}'_1 = [\hat{\boldsymbol{\theta}}_1, \mathsf{I}'_0]\,. \tag{62}$$

In the tangent space, the variables used to describe the system are thus $(\boldsymbol{\pi}_1, \boldsymbol{\theta}_1) \in \mathbb{R}^3 \times \mathbb{R}^3$. We also expand the angular velocity as $\boldsymbol{\omega} = \boldsymbol{\omega}_0 + \varepsilon\boldsymbol{\omega}_1$ with $\boldsymbol{\omega}_0 = \mathbf{0}$. From Eq. (59), we get

$$\boldsymbol{\omega}_1 = \mathsf{I}'^{-1}_0 (\boldsymbol{\pi}_1 - \boldsymbol{\theta}_1 \times \boldsymbol{\pi}_0) + \boldsymbol{\theta}_1 \times \boldsymbol{\omega}_{\mathcal{R}'_0}\,. \tag{63}$$

The identification of the time derivative of (62) with the kinematic equation $\dot{\mathsf{I}}'_1 = [\hat{\boldsymbol{\omega}}_1, \mathsf{I}'_0]$ shows that $\dot{\boldsymbol{\theta}}_1 = \boldsymbol{\omega}_1$. Therefore at first order in ε, the kinematic equation reads

$$\dot{\boldsymbol{\theta}}_1 = \mathsf{I}'^{-1}_0 \boldsymbol{\pi}_1 + \left(\mathsf{I}'^{-1}_0 \hat{\boldsymbol{\pi}}_0 - \hat{\boldsymbol{\omega}}_{\mathcal{R}'_0}\right) \boldsymbol{\theta}_1\,. \tag{64}$$

As for the dynamical equation, the expansion of (60) gives

$$\dot{\boldsymbol{\pi}}_1 = -\hat{\boldsymbol{\omega}}_{\mathcal{R}'_0} \boldsymbol{\pi}_1 + 3n^2 \left[(\mathbf{e}'_1 \cdot \mathsf{I}'_0 \mathbf{e}'_1)\mathbb{I} - \mathsf{I}'_0 \mathbf{e}'_1 \otimes \mathbf{e}'_1 + \hat{\mathbf{e}}'_1 \mathsf{I}'_0 \hat{\mathbf{e}}'_1\right] \boldsymbol{\theta}_1\,. \tag{65}$$

Hence, the linearised system of differential equations is

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} \boldsymbol{\pi}_1 \\ \boldsymbol{\theta}_1 \end{pmatrix} = \mathsf{A} \begin{pmatrix} \boldsymbol{\pi}_1 \\ \boldsymbol{\theta}_1 \end{pmatrix}\,, \tag{66}$$

with

$$\mathsf{A} = \begin{bmatrix} 0 & n & 0 & 0 & 0 & 0 \\ -n & 0 & 0 & 0 & 3n^2(A-C) & 0 \\ 0 & 0 & 0 & 0 & 0 & 3n^2(A-B) \\ 1/A & 0 & 0 & 0 & n(A-C)/A & 0 \\ 0 & 1/B & 0 & n(C-B)/B & 0 & 0 \\ 0 & 0 & 1/C & 0 & 0 & 0 \end{bmatrix}\,. \tag{67}$$

Because of the three degrees of freedom of rotation of $\mathcal{B}$, we shall expect three eigenmodes (also called *proper modes*). A quick inspection of the matrix A shows that the motion involving the third components of $\boldsymbol{\pi}_1$ and $\boldsymbol{\theta}_1$, called *libration in longitude*, is decoupled from those involving the first two components of the same vectors, called *libration in latitude* and *wobble*. When $A \leq B \leq C$ all the eigenvalues of A are imaginary and equal to $\pm\mathrm{i}\nu_1$, $\pm\mathrm{i}\nu_2$ and $\pm\mathrm{i}\nu_3$ where ν_1, ν_2 and ν_3 are the eigenfrequencies given by

$$\begin{aligned}
&\text{(libration in longitude)} && \nu_1 = n\sqrt{3\gamma}\,,\\
&\text{(libration in latitude)} && \nu_2 = n\sqrt{\frac{(1+\alpha\beta+3\beta)+\sqrt{\Delta}}{2}}\,,\\
&\text{(wobble)} && \nu_3 = n\sqrt{\frac{(1+\alpha\beta+3\beta)-\sqrt{\Delta}}{2}}\,,
\end{aligned} \tag{68}$$

with $\Delta = (1+\alpha\beta+3\beta)^2 - 16\alpha\beta$. Here we have introduced the relative moments of inertia

$$\alpha = \frac{C-B}{A}\,, \qquad \beta = \frac{C-A}{B}\,, \qquad \gamma = \frac{B-A}{C}\,. \tag{69}$$

The three proper motions are represented in Fig. 3.

6.4 Driven Solution

The perturbation $\mathbf{r}_1(t)$ imposes a driven rotation motion of $\mathcal{B}$. At first order in inclination and eccentricity the coordinates (x_1, y_1, z_1) of $\mathbf{r}_1(t)$ in the frame $(\vec{e_1}', \vec{e_2}', \vec{e_3}')$ are

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \mathsf{R}_3(\Omega-\lambda)\mathsf{R}_1(i)\mathsf{R}_3(\varpi-\Omega)\begin{pmatrix} r\cos v \\ r\sin v \\ 0 \end{pmatrix} - \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} \approx a\begin{pmatrix} -e\cos(\lambda-\varpi) \\ 2e\sin(\lambda-\varpi) \\ \sin i\,\sin(\lambda-\Omega) \end{pmatrix}. \tag{70}$$

In the following, we consider a fixed eccentricity e and a fixed inclination i with respect to the $(\vec{e_1}, \vec{e_2})$ plane that we call *invariant plane*. This departure from a circular and coplanar orbit introduces an additional torque $\mathbf{T}_1$ in the Poincaré-Lagrange equation (60). Its expression is given by a first order expansion in $\mathbf{r}_1$ of the torque $\mathbf{T} = -3(\mathcal{G}m_0/r^5)\,(\mathsf{I}'\mathbf{r}) \times \mathbf{r}$ taken from (58). We have

$$\begin{aligned}
\mathbf{T}_1 &= -3\frac{n^2}{a}\left[(\mathsf{I}_0'\mathbf{r}_1)\times\mathbf{e}_1' + (\mathsf{I}_0'\mathbf{e}_1')\times\mathbf{r}_1\right] = -3n^2\begin{pmatrix} 0 \\ (C-A)z_1 \\ (A-B)y_1 \end{pmatrix}\\
&= \begin{pmatrix} 0 \\ -3n^2(C-A)\sin i\,\sin(\lambda-\Omega) \\ 6n^2(B-A)e\sin(\lambda-\varpi) \end{pmatrix}.
\end{aligned} \tag{71}$$

The new linearised equations of motion become

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \boldsymbol{\pi}_1 \\ \boldsymbol{\theta}_1 \end{pmatrix} = \mathsf{A}\begin{pmatrix} \boldsymbol{\pi}_1 \\ \boldsymbol{\theta}_1 \end{pmatrix} + \begin{pmatrix} \mathbf{T}_1 \\ \mathbf{0} \end{pmatrix}. \tag{72}$$

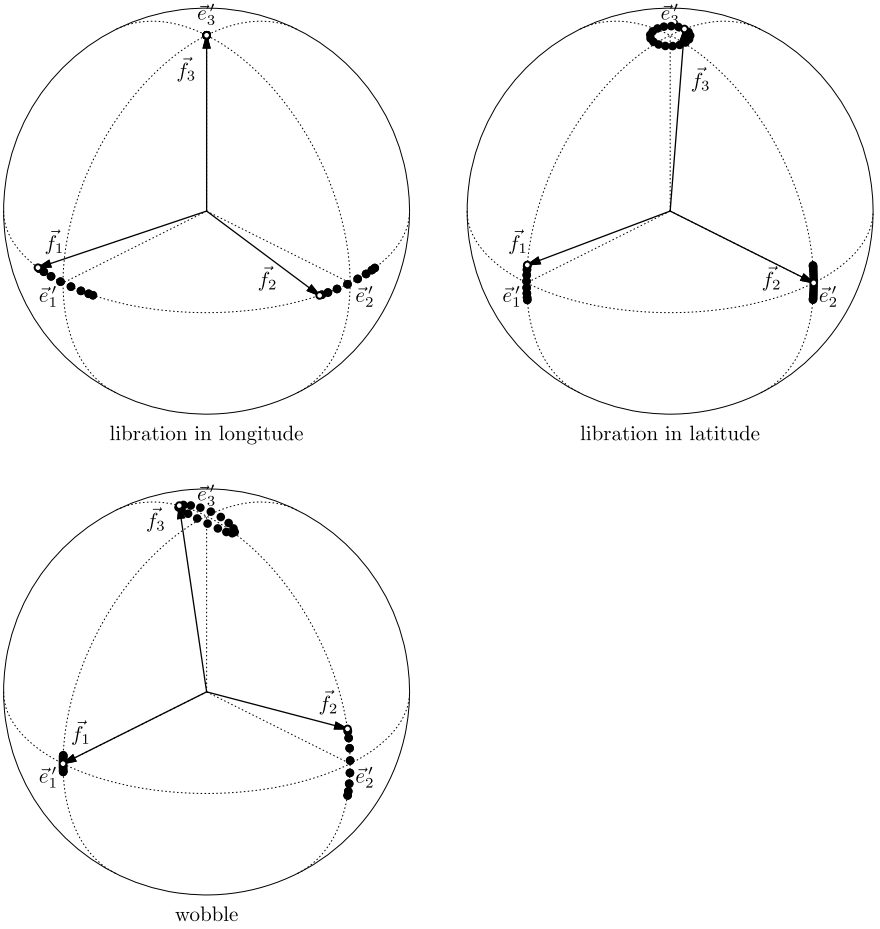


Fig. 3 Proper motions of the rigid body. Positions at constant time intervals over a full libration period of the principal axes $(\vec{f}_1, \vec{f}_2, \vec{f}_3)$ are depicted by black dots. Empty circles indicate the initial condition. Intersections of the dotted great circles of the unit sphere represent the base frame vectors $(\vec{e}_1', \vec{e}_2', \vec{e}_3')$

6.4.1 Libration in Longitude

The libration in longitude, involving the third components (π_{1z}, θ_{1z}) of $\boldsymbol{\pi}_1$ and $\boldsymbol{\theta}_1$ respectively, is driven by the eccentricity e of the orbit present in T_{1z}. From the expression (67) of the matrix A, we get the differential equation satisfied by θ_{1z}, namely,

$$\ddot{\theta}_{1z} + 3n^2\gamma\,\theta_{1z} = 6n^2\gamma e \sin(\lambda - \varpi). \tag{73}$$

For $|\dot{\varpi}| \ll n$ and $\gamma \ll 1$, the forced solution of this equation is

$$\theta_{1z} = -6\gamma e \sin(\lambda - \varpi)\,, \tag{74}$$

which means that $\mathcal{B}$ librates in longitude with an amplitude equal to $6\gamma e$. Hence, measuring this amplitude allows to estimate the equatorial flattening $\gamma = (B - A)/C$ of $\mathcal{B}$.

6.4.2 Libration in Latitude

The libration in latitude is a precession of the angular momentum $\boldsymbol{\pi}$ around the invariant vector $\mathbf{e}_3 = \mathbf{e}'_3$. This motion is excited by the inclination i of the orbit which appears in the expression of T_{1y}. Assuming $\dot{\Omega} \ll n$ and $\alpha, \beta \ll 1$, the driven solution of (72) when T_{1z} is set to zero reads

$$\begin{cases} \pi_{1x} = -\vartheta\pi_0 \sin(\lambda - \Omega)\,, \\ \pi_{1y} = -\vartheta\pi_0 \cos(\lambda - \Omega)\,, \\ \pi_{1z} = 0\,, \end{cases} \qquad \begin{cases} \theta_{1x} = \vartheta \cos(\lambda - \Omega)\,, \\ \theta_{1y} = -\vartheta \sin(\lambda - \Omega)\,, \\ \theta_{1z} = 0\,, \end{cases} \tag{75}$$

with

$$\vartheta = \frac{3\beta n \sin i}{2\dot{\Omega} + 3\beta n}\,. \tag{76}$$

The solution (75) is expressed in the rotating frame $(\vec{e_1}', \vec{e_2}', \vec{e_3}')$. In this frame, the coordinates of the orbit normal $\vec{N}$ are

$$\mathbf{N} = \mathsf{R}_3(\Omega - \lambda)\mathsf{R}_1(i) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin i \sin(\Omega - \lambda) \\ -\sin i \cos(\Omega - \lambda) \\ \cos i \end{pmatrix}. \tag{77}$$

The orbit normal $\vec{N}$, the angular momentum $\vec{\pi} = \vec{\pi_0} + \vec{\pi_1}$, the rotation vector $\vec{\theta} = \vec{\theta_1}$ and the spin axis $\vec{f_3} = \vec{e_3}' + \vec{\theta_1} \times \vec{e_3}'$ are displayed in Fig. 4. In this driven solution, the pole of the invariant plane $\vec{e_3}$, the orbit normal $\vec{N}$ and the angular momentum $\vec{\pi}$ remain coplanar. Such a configuration is called a *Cassini state* [7]. This motion is characterised by, e.g., the tilt angle ϵ between $\vec{\pi}$ and the orbit pole $\vec{N}$. This angle, called *obliquity*, is equal to[2]

$$\epsilon = i - \vartheta \approx \frac{\dot{\Omega}}{\dot{\Omega} + \frac{3}{2}\beta n} \sin i\,. \tag{78}$$

[2] By definition, the obliquity is a positive angle between 0 and 180 degrees. But in the particular case of a Cassini state where $\vec{N}$, $\vec{e_3}'$ and $\vec{f_3}'$ are coplanar, an orientation is defined by $\vec{N}$ and $\vec{e_3}'$. Here, we follow the convention defined in [28] (this convention yields a positive obliquity for the Moon).

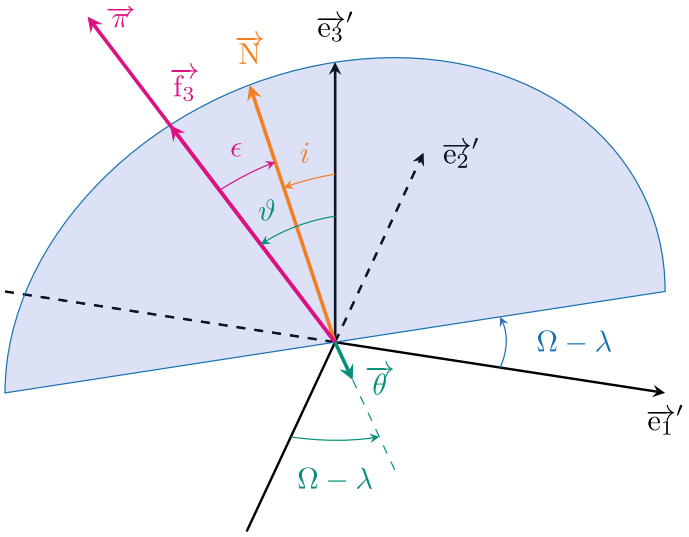


Fig. 4 Forced libration in latitude seen in the $(\vec{e_1}', \vec{e_2}', \vec{e_3}')$ rotating plane. The orbit normal $\vec{N}$, the normal to the Laplace plane $\vec{e_3}'$, the body spin axis $\vec{f_3}$ and its angular momentum $\vec{\pi}$ are all coplanar. $\vec{f_3}$ is deduced from $\vec{e_3}'$ by the rotation of vector $\vec{\theta}$. The same relation holds between $\vec{\pi_0}$ and $\vec{\pi}$

It should be stressed that $\dot{\Omega}$ is usually negative and can eventually be equal to $-\frac{3}{2}\beta n$. The equality between these two terms corresponds to a resonance between the precession of the orbit and the free libration in latitude. In that case Eq. (78) is no more valid and non-linearities have to be taken into account (see, e.g., [34]).

7 Tidal Deformation

Although at first approximation celestial bodies can be considered perfectly rigid, important behaviours referred to as *tidal effects* require to model them as dissipative deformable extended bodies. The shape of a body $\mathcal{B}$ is subject to deformation as long as it evolves in a varying non-uniform gravity field with respect to its body-fixed frame. This is particularly the case when $\mathcal{B}$ orbits a point mass if its orbit is not circular or if its rotation is not synchronous. Sufficiently massive bodies are expected to be spherical when they are at rest (i.e., isolated without proper rotation) because their figure is controlled by gravity and for a given volume the sphere minimises the potential energy. But due to their proper rotation, celestial bodies are flatten at the poles and the presence of a massive companion produces two additional bulges approximately aligned with the neighbour (one at the front and the other at the far side). The redistribution of mass associated with the bulges induces friction and

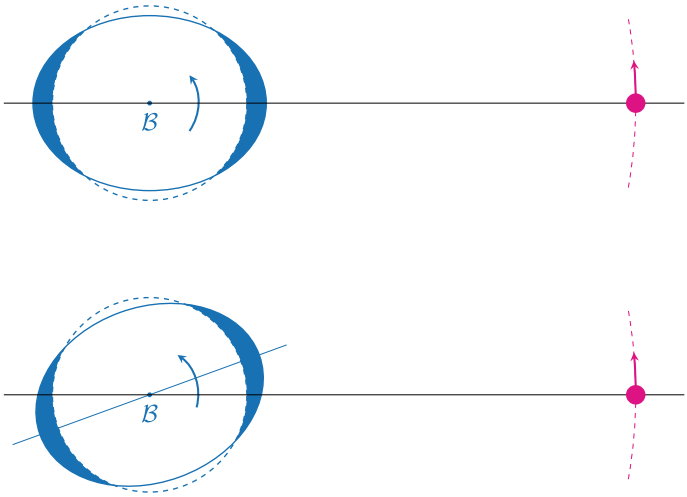


Fig. 5 The upper diagram represents tidal deformation with no friction: the bulges are aligned with the companion. When friction is added (lower diagram), there is a delay between the high tides and the direction of the companion. In this figure, the spin rate is assumed to be faster than the revolution frequency. The resulting orientation of the tidal bulges leads to a deceleration of the spin velocity and an acceleration of the orbital motion [23]

dissipation within the body and the slight misalignment with the direction of the companion introduces a torque affecting the rotation (see Fig. 5).

7.1 Inertia Matrix and Stokes Coefficients

Because of its deformation, the inertia matrix I of $\mathcal{B}$ changes with time. At each instant we could choose to write I in a frame where it is diagonal but even in that case, the diagonal elements would vary. As discussed in the next subsection, it is more convenient to consider another frame attached to $\mathcal{B}$ in which all entries of I can eventually oscillate. Here we only consider cases where $\mathcal{B}$'s shape is almost spherical and I almost proportional to the identity matrix. Let $\mathrm{I}_\circ$ be the *mean moment of inertia* defined as $\mathrm{I}_\circ = \frac{1}{3}\,\mathrm{Tr}\,\mathsf{I}$. We then follow Ragazzo & Ruiz [32, 33] approach and define the traceless symmetric matrix B by $\mathsf{I} = \mathrm{I}_\circ(\mathbb{I} - \mathsf{B})$. This matrix is related to the second order *Stokes coefficients* C_{lm} and S_{lm} (which are commonly used to parameterise the gravity field of an extended body [20]) through

$$\mathsf{B} = \frac{2}{3}\frac{mR^2}{\mathrm{I}_\circ}\begin{bmatrix} 3C_{22} - \frac{1}{2}C_{20} & 3S_{22} & \frac{3}{2}C_{21} \\ 3S_{22} & -3C_{22} - \frac{1}{2}C_{20} & \frac{3}{2}S_{21} \\ \frac{3}{2}C_{21} & \frac{3}{2}S_{21} & C_{20} \end{bmatrix}. \tag{79}$$

For a homogeneous body, the matrix B is also related to the displacement field (strain deformation) of the material within the body. Indeed, if $\mathrm{P} \in \mathcal{B}$ is a mass element initially at the position $\mathbf{y} = \mathrm{O}'\mathrm{P}$ when $\mathcal{B}$ is spherical, under deformation it moves to the position $\mathbf{y} + \mathsf{B}\mathbf{y}$.

7.2 Tisserand Frame

Considering $\mathcal{B}$ as a deformable body raises the question of the definition of a reference frame $\mathcal{R}_1$ attached to the body. Observationally, the rotation of a solid body is deduced from the motion of surface features. Using Jupiter's red spot and characteristic patterns at the surface of Mars, Jean-Dominique Cassini managed to visually estimate the rotation rate of these planets since the years 1664–1666 [11]. Now, more precise measurements are obtained using radar signals [29]. But in both cases, observations give access to the rotation of the surface only.

In dynamical studies it is more convenient to choose $\mathcal{R}_1$ as the *Tisserand frame* defined as the frame in which $\mathcal{B}$ has no angular momentum. Let $\mathbf{y}^0$ be the coordinates of $\overrightarrow{\mathrm{O}'\mathrm{P}}$ in $\mathcal{R}_0$ and $\mathbf{y}^1$ the coordinates of the same vector in $\mathcal{R}_1$. The angular momentum $\boldsymbol{\ell}$ of $\mathcal{B}$ with respect to its barycentre O' and relative to the inertial frame $\mathcal{R}_0$ is, by definition,

$$\boldsymbol{\ell} = \int_{\mathcal{B}} \mathbf{y}^0 \times \dot{\mathbf{y}}^0 \, \mathrm{d}m(\mathrm{P}) \tag{80}$$

with

$$\mathbf{y}^0 = \mathsf{R}\mathbf{y}^1 \qquad \Rightarrow \qquad \dot{\mathbf{y}}^0 = \dot{\mathsf{R}}\mathbf{y}^1 + \mathsf{R}\dot{\mathbf{y}}^1 \, . \tag{81}$$

We recall that $\mathcal{B}$ being deformable, $\mathbf{y}^1$ is not constant anymore. Let us introduce the angular velocity $\boldsymbol{\omega}^1$ written in the frame $\mathcal{R}_1$, i.e., $\boldsymbol{\omega}^1 = \mathsf{R}^{\top}\boldsymbol{\omega}^0$ such that $\dot{\mathbf{y}}^0 = \mathsf{R}(\boldsymbol{\omega}^1 \times \mathbf{y}^1 + \dot{\mathbf{y}}^1)$. We have then

$$\begin{aligned} \boldsymbol{\ell} &= \mathsf{R} \int_{\mathcal{B}} \mathbf{y}^1 \times (\boldsymbol{\omega}^1 \times \mathbf{y}^1 + \dot{\mathbf{y}}^1) \, \mathrm{d}m(\mathrm{P}) \\ &= \mathsf{R} \left(\int_{\mathcal{B}} ((y^1)^2 \, \mathbb{I} - \mathbf{y}^1 \otimes \mathbf{y}^1) \, \mathrm{d}m(\mathrm{P}) \right) \boldsymbol{\omega}^1 + \mathsf{R} \int_{\mathcal{B}} \mathbf{y}^1 \times \dot{\mathbf{y}}^1 \, \mathrm{d}m(\mathrm{P}) \\ &= \mathsf{R} \left(\mathsf{I} \, \boldsymbol{\omega}^1 + \int_{\mathcal{B}} \mathbf{y}^1 \times \dot{\mathbf{y}}^1 \, \mathrm{d}m(\mathrm{P}) \right) . \end{aligned} \tag{82}$$

By construction, the Tisserand frame is chosen so that the last integral vanishes. Therefore, the angular momentum is simply given by $\boldsymbol{\ell} = \mathsf{R}\mathsf{I} \, \boldsymbol{\omega}^1 = \mathsf{I}'\boldsymbol{\omega}^0$.

7.3 Lagrangian of the Problem

Given that the elements of the matrix B are allowed to vary, they can be considered as new generalised coordinates of the problem [32, 33]. For that purpose, one shall add the contribution of $\dot{\mathsf{B}}$ to the kinetic energy of the system, include the potential energy associated with the auto-gravity of $\mathcal{B}$ and the eventual elasticity of its material, and complete the Lagrangian with a dissipation function. Nevertheless, the inertia of deformation is usually negligible [8] and we will discard the increment of kinetic energy induced by $\dot{\mathsf{B}}$.

From now on, since we work with matrices, we introduce a new scalar product and its related norm. For any two matrices A and B we define $\mathsf{A}\cdot\mathsf{B} = \frac{1}{2}\,\mathrm{Tr}(\mathsf{A}^\top\mathsf{B})$ and $\|\mathsf{A}\|^2 = \mathsf{A}\cdot\mathsf{A} = \frac{1}{2}\,\mathrm{Tr}(\mathsf{A}^\top\mathsf{A})$.

At the lowest order in $\|\mathsf{B}\|$, the gravitational potential energy U_g and the elastic potential energy U_el of an isotropic body[3] are of the form

$$U_\mathrm{g} = \frac{1}{2} g \mathrm{I}_\circ \|\mathsf{B}\|^2 , \qquad U_\mathrm{el} = \frac{1}{2} \mu \mathrm{I}_\circ \|\mathsf{B}\|^2 , \tag{83}$$

where g and μ are two coefficients to be determined. To model the viscosity of the material, we add the Rayleigh dissipation function

$$\mathcal{D} = \frac{1}{2} \eta \mathrm{I}_\circ \|\dot{\mathsf{B}}\|^2 , \tag{84}$$

where η is another physical parameter of $\mathcal{B}$. It should be stressed that for an isotropic body, the potential energy and the dissipation function have to be invariant by rotation as are $\|\mathsf{B}\|^2$ and $\|\dot{\mathsf{B}}\|^2$. More specifically, for an isotropic homogeneous body of radius R at rest,

$$g = \frac{4}{5}\frac{\mathcal{G}m}{R^3} , \qquad \mu = \frac{40\pi}{3}\frac{R}{m}\mu_0 , \qquad \eta = \frac{40\pi}{3}\frac{R}{m}\eta_0 , \tag{85}$$

where μ_0 and η_0 are effective microscopic shear modulus and shear viscosity of the material, respectively. The equations presented here are those for a visco-elastic body with a Kelvin-Voigt rheology. They can easily be generalised to any rheology [33].

Assuming as above that $\mathcal{B}$ is orbiting a point mass m_0, the Lagrangian of the system reads

$$\mathcal{L} = \frac{\mathrm{I}_\circ}{2}\left(\|\boldsymbol{\omega}\|^2 - \boldsymbol{\omega}\cdot\mathsf{B}'\boldsymbol{\omega}\right) - \frac{\mathrm{I}_\circ}{2} g \|\mathsf{B}\|^2 - \frac{\mathrm{I}_\circ}{2}\mu\|\mathsf{B}\|^2 + \frac{3}{2}\frac{\mathcal{G}m_0\mathrm{I}_\circ}{r^5}\mathbf{r}\cdot\mathsf{B}'\mathbf{r} , \tag{86}$$

where $\mathsf{B}' = \mathsf{R}\mathsf{B}\mathsf{R}^\top$ is the matrix B expressed in the inertial frame $\mathcal{R}_0$. The equations of motion (31) become

[3] An isotropic body is a body which tends to be spherical when tidal and rotation perturbations are artificially set to zero. In this asymptotic form, all physical variables (density, elasticity, viscosity, etc.) are radially distributed but not necessarily uniform throughout the body.

$$\eta\dot{\mathsf{B}} + (g+\mu)\mathsf{B} = \mathsf{F}\,, \tag{87}$$

$$\dot{\boldsymbol{\pi}} = -3\frac{\mathcal{G}m_0\mathrm{I}_\circ}{r^5}(\mathsf{B}'\mathbf{r})\times\mathbf{r}\,, \tag{88}$$

$$\dot{\mathsf{R}} = \hat{\boldsymbol{\omega}}\mathsf{R}\,, \tag{89}$$

where $\boldsymbol{\pi} = \partial_{\boldsymbol{\omega}}\mathcal{L} = \mathrm{I}_\circ(\boldsymbol{\omega} - \mathsf{B}'\boldsymbol{\omega})$ and $\mathsf{F} = \mathsf{R}\mathsf{F}'\mathsf{R}^\top$ with[4]

$$\mathsf{F}' = -\boldsymbol{\omega}\otimes\boldsymbol{\omega} + \frac{1}{3}\omega^2\mathbb{I} + 3\frac{\mathcal{G}m_0}{r^5}\left(\mathbf{r}\otimes\mathbf{r} - \frac{1}{3}r^2\mathbb{I}\right). \tag{90}$$

7.4 Love Number

In good approximation, tidal deformation is the linear response of an extended body to a perturbing potential (B is proportional to F). This behaviour can be interpreted as a transfer function between the tidal potential W and the gravitational potential V' induced by the mass redistribution within $\mathcal{B}$. This problem has been studied by Augustus Edward Hough Love [22] after whom *Love numbers* have been named. Let $\mathrm{P}\in\partial\mathcal{B}$ be a point at the surface of $\mathcal{B}$ of coordinate $\mathbf{y}^1$ in $\mathcal{R}_1$ (i.e., such that $\|\mathbf{y}^1\| = R$). In the Fourier domain, we define the *complex Love number* $\underline{k}_2(\nu)$ at the frequency ν as

$$\hat{V}'(\mathbf{y}^1,\nu) = \underline{k}_2(\nu)\hat{W}(\mathbf{y}^1,\nu)\,. \tag{91}$$

From the expression of V' and W, namely,

$$V'(\mathbf{y}^1,t) = -\frac{3}{2}\frac{\mathcal{G}\mathrm{I}_\circ}{R^5}\,\mathbf{y}^1\cdot\mathsf{B}(t)\mathbf{y}^1 \qquad\text{and}\qquad W(\mathbf{y}^1,t) = -\frac{1}{2}\mathbf{y}^1\cdot\mathsf{F}(t)\mathbf{y}^1\,, \tag{92}$$

we deduce that

$$\hat{\mathsf{B}}(\nu) = \underline{K}_2(\nu)\hat{\mathsf{F}}(\nu)\,, \qquad \underline{k}_2(\nu) = \frac{3\mathcal{G}\mathrm{I}_\circ}{R^5}\,\underline{K}_2(\nu)\,. \tag{93}$$

By identification with (87), we see that in the above model

$$\underline{K}_2(\nu) = \frac{K_{\mathrm{el}}}{1+\mathrm{i}\tau\nu}\,, \qquad K_{\mathrm{el}} = \frac{1}{g+\mu}\,, \qquad \tau = \frac{\eta}{g+\mu}\,, \tag{94}$$

where the constant τ represents a *viscous timescale*. From the formulae (93) and (94), one can derive the *elastic Love number* $k_{\mathrm{el}} = (3\mathcal{G}\mathrm{I}_\circ/R^5)\,K_{\mathrm{el}}$. In the case of a fluid material, $\mu = 0$, we obtain the so-called *fluid Love number* $k_{\mathrm{f}} = (3\mathcal{G}\mathrm{I}_\circ/R^5)\,K_{\mathrm{f}}$

[4] In the calculation of $\mathsf{F} - (g+\mu)\mathsf{B} = 1/\mathrm{I}_\circ\,\partial_{\mathsf{B}}\mathcal{L}$, we recall that B is traceless symmetric. Therefore, $\partial_{\mathsf{B}}\mathcal{L}$ is, by definition, the only traceless symmetric matrix such that $\mathrm{d}\mathcal{L} = \partial_{\mathsf{B}}\mathcal{L}\cdot\mathrm{d}\mathsf{B}$. Let K be the matrix with entries $K_{ij} = \partial\mathcal{L}/\partial B_{ij}$, then $\partial_{\mathsf{B}}\mathcal{L} = \mathsf{K} - \frac{1}{3}\,\mathrm{Tr}\,(\mathsf{K})\mathbb{I}$.

with $K_{\rm f} = 1/g$. In particular, for a homogeneous body, using (85), one gets[5]

$$k_{\rm f} = \frac{3}{2} \qquad k_{\rm el} = \frac{k_{\rm f}}{1 + \frac{25}{2}\frac{\mu_0}{\rho g R}} \qquad k_{\rm visc} = \frac{k_{\rm f}}{1 + \frac{25}{2}\frac{{\rm i}\eta\nu}{\rho g R}}\,, \tag{95}$$

where $\rho = 3m/(4\pi R^3)$ is the mean density and $g = Gm/R^2$ the surface gravity. The *viscous Love number* $k_{\rm visc}$ corresponds to the case $\mu = 0$ and $\eta \neq 0$. The (macroscopic) rheology described by (94) is that of a Kelvin-Voigt material with rigidity $g + \mu$ and viscosity η. Nevertheless, one can plug any arbitrary rheology in (93) and use it as a substitute of Eq. (87).

7.5 Constant Deformation

As we have seen, the deformation matrix B of the body $\mathcal{B}$ is proportional to the force F given in Eq. (90). Here we focus on the constant part $\bar{F}$ of the force which governs the permanent deformation $\bar{\mathsf{B}}$ of $\mathcal{B}$. The other harmonics introduce small oscillating deviations that we discard for the moment. In addition, we assume that $\mathcal{B}$ is in a $p:1$ spin-orbit resonance, i.e., $\omega = pn$ where p is a half-integer, and that its spin axis is orthogonal to the orbit plane (planar case). The spatial case is treated in [10]. To simplify the expression, we also assume that $m_0 \gg m$ such that $n^2 a^3 \approx \mathcal{G} m_0$. Within these hypotheses,

$$\begin{aligned}\mathsf{F} = & p^2 n^2 \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & -2/3 \end{bmatrix} \\ & + 3\,n^2 \left(\frac{a}{r}\right)^3 \begin{bmatrix} \cos^2(v - pnt) - 1/3 & \cos(v - pnt)\sin(v - pnt) & 0 \\ \cos(v - pnt)\sin(v - pnt) & \sin^2(v - pnt) - 1/3 & 0 \\ 0 & 0 & -1/3 \end{bmatrix},\end{aligned} \tag{96}$$

where v is the true anomaly. Therefore,

[5] Let us stress that the formulae (95) show a coefficient 25/2 in their denominator whereas "exact" expressions have a coefficient 19/2 (as obtained in 1863 by [35], see also Chap. V of [22]). This discrepancy comes from the homogeneous-deformation assumption in [32, 33]. In ibid., the viscoelastic deformation is computed in absence of gravity, which naturally leads to a homothetic transformation of the whole body. The effect of gravity field is only added afterwards. Viscoelasticity and gravity have to be treated jointly to retrieve the correct factor 19/2. Nevertheless, bodies are never strictly homogeneous. Therefore the small error 25/2 instead of 19/2 pertaining to the homogeneous-deformation approximation remains negligible. The advantage of the formalism used here is to provide a pedagogical interpretation of the effect of the gravity field.

$$\bar{\mathsf{F}} = p^2 n^2 \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & -2/3 \end{bmatrix} + 3n^2 \begin{bmatrix} \frac{1}{6}X_0^{-3,0}(e) + \frac{1}{2}X_{2p}^{-3,2}(e) & 0 & 0 \\ 0 & \frac{1}{6}X_0^{-3,0}(e) - \frac{1}{2}X_{2p}^{-3,2}(e) & 0 \\ 0 & 0 & -\frac{1}{3}X_0^{-3,0}(e) \end{bmatrix}, \tag{97}$$

where we have introduced *Hansen coefficients* $X_k^{n,m}(e)$ [15]. These are functions of the orbital eccentricity e defined as

$$\left(\frac{r}{a}\right)^n \exp(\mathrm{i}mv) = \sum_{k=-\infty}^{\infty} X_k^{n,m}(e)\, \exp(\mathrm{i}kM)\,, \tag{98}$$

where $M = nt$ is the mean anomaly, or alternatively as

$$X_k^{n,m}(e) = \frac{1}{2\pi}\int_0^{2\pi} (\cos E - e + \mathrm{i}\sqrt{1-e^2}\sin E)^m (1 - e\cos E)^{n-m+1} \times \exp[-\mathrm{i}k(E - e\sin E)]\,\mathrm{d}E\,. \tag{99}$$

Solving Eq. (93) and using the expression of B in terms of Stokes coefficients (Eq. 79), we see that $\bar{J}_2 = -\bar{C}_{20}$ and $\bar{C}_{22}$ are the only non-zero coefficients. We have

$$\bar{J}_2 = \frac{\mathrm{I}_\circ}{mR^2} n^2 K_{\mathrm{el}} \left(p^2 + \frac{3}{2}X_0^{-3,0}(e)\right), \qquad \bar{C}_{22} = \frac{\mathrm{I}_\circ}{mR^2} n^2 K_{\mathrm{el}}\, \frac{3}{4} X_{2p}^{-3,2}(e)\,. \tag{100}$$

Quite remarkably, their ratio is independent of the physical properties of $\mathcal{B}$. Indeed,

$$\frac{\bar{C}_{22}}{\bar{J}_2} = \frac{3X_{2p}^{-3,2}(e)}{4p^2 + 6X_0^{-3,0}(e)}\,. \tag{101}$$

For $2p = 1, .., 5$, the relevant Hansen coefficients are

$$X_0^{-3,0}(e) = (1-e^2)^{-3/2}\,, \tag{102}$$

$$X_1^{-3,2}(e) = -\frac{1}{2}e + \frac{1}{16}e^3 - \frac{5}{384}e^5 - \frac{143}{18432}e^7 - \frac{9097}{1474560}e^9 + O(e^{11})\,, \tag{103}$$

$$X_2^{-3,2}(e) = 1 - \frac{5}{2}e^2 + \frac{13}{16}e^4 - \frac{35}{288}e^6 - \frac{5}{576}e^8 - \frac{49}{3600}e^{10} + O(e^{12})\,, \tag{104}$$

$$X_3^{-3,2}(e) = \frac{7}{2}e - \frac{123}{16}e^3 + \frac{489}{128}e^5 - \frac{1763}{2048}e^7 + \frac{13527}{163840}e^9 + O(e^{11})\,, \tag{105}$$

$$X_4^{-3,2}(e) = \frac{17}{2}e^2 - \frac{115}{6}e^4 + \frac{601}{48}e^6 - \frac{1423}{360}e^8 + \frac{48619}{69120}e^{10} + O(e^{12})\,, \tag{106}$$

$$X_5^{-3,2}(e) = \frac{845}{48}e^3 - \frac{32525}{768}e^5 + \frac{208225}{6144}e^7 - \frac{6122725}{442368}e^9 + O(e^{11})\,. \tag{107}$$

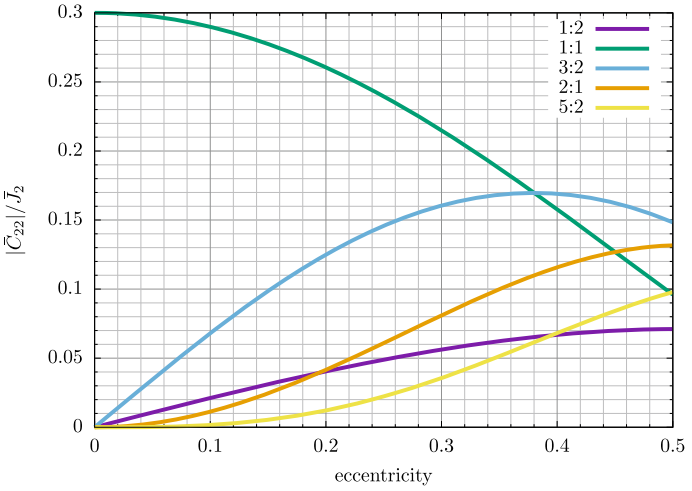


Fig. 6 Ratio of the equilibrium C_{22} over the equilibrium J_2 of a deformable body in different spin-orbit resonances as a function of the orbital eccentricity

In the case of the 1:2 spin-orbit resonance, $\bar{C}_{22}$ given by Eq. (100) is negative. This means that at pericentre the shortest equatorial axis points towards the central body. But for all the other spin-orbit resonances considered here, $\bar{C}_{22} > 0$, meaning that this is the long axis that lines up with the direction of the central body at pericentre.

The equilibrium shapes for different spin-orbit resonances are displayed as a function of the eccentricity in Fig. 6. In particular, for a synchronous object in circular orbit, we expect the ratio of its Stokes coefficients to be $\bar{C}_{22}/\bar{J}_2 = 0.3$.

These coefficients can be estimated from the analysis of the trajectory of a particle in the gravity field of the extended body. They are known for all the planets and for several satellites of the solar system, a few of which are displayed in Table 1. According to the principle of parsimony, without additional information, synchronous satellites satisfying the condition $\bar{C}_{22}/\bar{J}_2 = 0.3$ are commonly believed to be in hydrostatic equilibrium. In Table 1, all satellites but the Moon[6] are likely in hydrostatic equilibrium. Under this assumption, one can infer their mean moment of inertia $\mathrm{I}_\circ$ thanks to Clairaut's equation [31]. In turn, the mean moment of inertia constrains their internal mass distribution. Alternatively, the mean moment of inertia can be inferred from libration amplitudes (see Sect. 6.4).[7]

[6] The origin of the Moon's excess deformation is attributed to a fossil figure acquired when it was closer to the Earth and to the subsequent formation of the South Pole-Aitken Basin [19].

[7] In Sect. 6.4, it is shown that the amplitude of libration in longitude is proportional to γ while the amplitude of libration in latitude is a function of β. Both quantities are related to the mean moment of inertia as follows:

$$\gamma = \frac{4C_{22}}{\mathrm{I}_\circ}\,, \qquad \beta = \frac{J_2 + 2C_{22}}{\mathrm{I}_\circ}\,.$$

.

Table 1 Gravity field coefficients J_2 and C_{22} of synchronous satellites

Planet	Satellite	$J_2 \times 10^6$	$C_{22} \times 10^6$	C_{22}/J_2	References
Earth	Moon[a]	203.22	22.38	0.1101	[14, 21]
Jupiter	Io	1845.9 ± 4.2	553.7 ± 1.2	0.300 ± 0.001	[3]
	Europa	437.59 ± 77.47	138.62 ± 2.44	0.31 ± 0.06	[13]
	Ganymede[b]	127.27 ± 2.90	38.18 ± 0.87	0.3 (imposed)	[2]
	Callisto	32.7 ± 0.8	10.2 ± 0.3	0.31 ± 0.01	[1]
Saturn	Titan	33.599 ± 0.332	10.121 ± 0.029	0.301 ± 0.003	[16]

[a] The uncertainties on the gravity field coefficients of the Moon are not given in this table because they are much smaller than the significant digits written in this table. J_2 is precise at 10^{-11} and C_{22} at 10^{-12}.

[b] The value of Ganymede's coefficient J_2 has not been fitted to the observational data in [2]. Under the hydrostatic hypothesis, it has been set equal to $C_{22}/0.3$

7.6 Maximal Triaxiality

Some celestial bodies of the solar system show clear departure from hydrostatic equilibrium. For instance, Mercury's triaxiality is more than one order of magnitude greater than expected from the simple model of Sect. 7.5 (a discussion of the origin of Mercury's strong departure from hydrostaticity is given in [26]). These bodies behave like a spring whose "relaxed" shape is not the sphere. In that case, the elastic potential energy $U_{\rm el}$ given in Eq. (83) has to be upgraded to incorporate a relaxed shape B_0, viz.,

$$U_{\rm el} = \frac{1}{2}\mu \mathrm{I}_\circ \|\mathsf{B} - \mathsf{B}_0\|^2 . \tag{108}$$

In absence of viscosity, the expression of the equilibrium shape $\mathsf{B}_{\rm eq}$ (given by $\partial_{\mathsf{B}}\mathcal{L} = 0$) becomes (see Eq. (87))

$$\mathsf{B}_{\rm eq} = \frac{1}{g+\mu}(\mathsf{F} + \mu\mathsf{B}_0) = K_{\rm el}(\mathsf{F} + \mu\mathsf{B}_0) . \tag{109}$$

The term $\mu\mathsf{B}_0$ can be interpreted as a prestress modifying the equilibrium figure of $\mathcal{B}$. It should be noticed that a solid material cannot be stretched indefinitely. Above a given threshold, defined by the *yield criterion* [17], the material begins to plastically deform or fracture. The typical critical strain $u_{\rm crit}$ of the Earth's lithosphere is in the range 10^{-5}–10^{-3}. Therefore, the eigenvalues of $\Delta\mathsf{B} = \mathsf{B} - \mathsf{B}_0$ cannot exceed $u_{\rm crit}$. Let $\mathsf{B}_{\rm el} = K_{\rm el}\mathsf{F}$ be the elastic deformation in absence of prestress. At equilibrium, when $\mathsf{B} = \mathsf{B}_{\rm eq}$ and $\Delta B = \mathsf{B}_{\rm eq} - \mathsf{B}_0$, we have

$$\mathsf{B}_{\rm eq} - \mathsf{B}_{\rm el} = \frac{\mu}{g}(\mathsf{B}_{\rm el} - \Delta\mathsf{B}) = \left(\frac{k_{\rm f}}{k_{\rm el}} - 1\right)(\mathsf{B}_{\rm el} - \Delta\mathsf{B}) . \tag{110}$$

The permanent shape of a body which is not in spin-orbit resonance, like the Earth, departs from the sphere only through a flattening of the poles due to the proper rotation. Such a body is indeed expected to be axisymmetric because in that case both F and $\mathsf{B}_{\rm el}$ are invariant by rotation around the spin axis. Therefore, an eventual triaxiality $\gamma_{\rm eq} = (B_{\rm eq} - A_{\rm eq})/C_{\rm eq}$ of the equilibrium shape can only be the outcome of a non-axisymmetric prestress $\mu\mathsf{B}_0$. We can thus infer from Eq. (110) an upper bound on the permanent triaxiality given by

$$\frac{B_{\rm eq} - A_{\rm eq}}{C_{\rm eq}} < \left(\frac{k_{\rm f}}{k_{\rm el}} - 1\right) u_{\rm crit} \,. \tag{111}$$

We notice that the maximal triaxiality $\gamma_{\rm max} = (\frac{k_{\rm f}}{k_{\rm el}} - 1)u_{\rm crit}$ is a function of the ratio between the fluid Love number $k_{\rm f}$ and the elastic Love number $k_{\rm el}$. With the simple approach of Sect. 7.3, for a homogeneous body we get

$$\gamma_{\rm max} = \frac{25}{2}\frac{\mu_0}{\rho \bar{g} R} u_{\rm crit} \,, \tag{112}$$

where $\rho = \frac{3}{4\pi}\frac{m}{R^3}$ is the mean density and $\bar{g} = \frac{\mathcal{G}m}{R^2}$ the surface gravity. We stress that a similar study [37] performed with a different mathematical approach leads to a similar expression where the numerical factor $25/2$ is replaced by 7.9. Nevertheless, this difference is not significant as the main goal of Eq. (112) is to estimate the order of magnitude of the admissible triaxiality of a given celestial body for which we have very few information, such as an exoplanet. In the case of the Earth, we obtain $\gamma_{\rm max} \approx 3.5 \times 10^{-5}$ while the observed value is $\gamma = 1.5 \times 10^{-5}$.

8 Tidal Evolution

This section is dedicated to the rotation evolution of $\mathcal{B}$ due to tides. We assume the orbit to be Keplerian and unperturbed by the rotation. Moreover, to avoid tedious calculation, we here only consider the planar case where $\mathcal{B}$ rotates around its axis of maximum inertia which itself is orthogonal to the orbit plane. For this simplified problem, we slightly modify the notation used so far. The rotation matrix R is parameterised by a single angle which we denote θ such that $\mathsf{R} = \mathsf{R}_3(\theta)$. The origin of θ is the direction of the periapsis that is supposed to be fixed. We set $\Omega = \dot{\theta}$ as the angular speed[8] and we denote by ω any arbitrary frequency. Eventually, we approximate Kepler's third law by $n^2 a^3 = \mathcal{G}m_0$.

[8] The symbol Ω enters in conflict with that of the longitude of the ascending node defined in Sect. 6. Nevertheless, Ω is a standard notation for spin rates and, because the problem remains planar, the longitude of the ascending node never appears in this whole section. Therefore, we do not expect any confusion.

8.1 Equations of Motion

In the planar case, and in terms of Stokes coefficients (see Sect. 7.1), the set of Eqs. (88)–(89) reads

$$\frac{\mathrm{d}}{\mathrm{d}t}(C\dot{\theta}) = -6\frac{\mathcal{G}mm_0R^2}{r^3}\left[C_{22}\sin 2(\theta - v) + S_{22}\cos 2(\theta - v)\right], \tag{113}$$

where C itself is a function of the Stokes coefficient J_2, viz.,

$$C = \frac{2}{3}mR^2J_2 + \mathrm{I}_\circ\,. \tag{114}$$

We decompose each Stokes coefficient $Z \in \{J_2, C_{22}, S_{22}\}$ into a constant part Z_0 induced by an eventual prestress $\mu\mathsf{B}_0$ (see Sect. 7.6) and a tidal component Z_{TID} satisfying the tidal equation (87). In the Fourier domain, we have

$$\hat{Z}_{\mathrm{TID}}(\omega) = \underline{k}_2(\omega)\hat{Z}_{\mathrm{EXT}}(\omega)\,, \tag{115}$$

with Z_{EXT} the external perturbation deduced from the decomposition of the matrix F (90) in spherical harmonics [9]:

$$J_{2,\mathrm{EXT}}(t) = \frac{\Omega^2R^3}{3\mathcal{G}m} + \frac{1}{2}\frac{m_0}{m}\left(\frac{R}{r}\right)^3, \tag{116}$$

$$C_{22,\mathrm{EXT}}(t) = \frac{1}{4}\frac{m_0}{m}\left(\frac{R}{r}\right)^3\cos 2(\theta - v)\,, \tag{117}$$

$$S_{22,\mathrm{EXT}}(t) = -\frac{1}{4}\frac{m_0}{m}\left(\frac{R}{r}\right)^3\sin 2(\theta - v)\,. \tag{118}$$

Hereafter, the reference meridian of $\mathcal{B}$ is chosen such that $S_{22,0} = 0$.

8.2 Secular Tidal Torque Out of Spin-Orbit Resonances

To solve (115), we make the hypothesis that the rotation is almost uniform such that Ω can be approximated by a constant. The expansion of the Z_{EXT} in Fourier series gives

$$J_{2,\mathrm{EXT}}(t) = \frac{\Omega^2 R^3}{3\mathcal{G}m} + \frac{1}{2}\frac{m_0}{m}\left(\frac{R}{a}\right)^3 \sum_{k=-\infty}^{\infty} X_k^{-3,0}(e)\cos(kM)\,, \tag{119}$$

$$C_{22,\mathrm{EXT}}(t) = \frac{1}{4}\frac{m_0}{m}\left(\frac{R}{a}\right)^3 \sum_{k=-\infty}^{\infty} X_k^{-3,2}(e)\cos(2\theta - kM)\,, \tag{120}$$

$$S_{22,\mathrm{EXT}}(t) = -\frac{1}{4}\frac{m_0}{m}\left(\frac{R}{a}\right)^3 \sum_{k=-\infty}^{\infty} X_k^{-3,2}(e)\sin(2\theta - kM)\,. \tag{121}$$

It is customary to decompose the complex Love number as $\underline{k}_2(\omega) = k_2(\omega)[\cos\epsilon(\omega) - \mathrm{i}\sin\epsilon(\omega)]$ where the modulus $k_2(\omega)$ is an even function ω and the *phase lag* $\epsilon(\omega)$ an odd function of ω. With this notation, the solution of (115) is

$$J_{2,\mathrm{TID}}(t) = k_2(0)\frac{\Omega^2 R^3}{3\mathcal{G}m} + \frac{1}{2}\frac{m_0}{m}\left(\frac{R}{a}\right)^3 \sum_{k=-\infty}^{\infty} k_2(kn)\, X_k^{-3,0}(e)\cos[kM - \epsilon(kn)]\,, \tag{122}$$

$$C_{22,\mathrm{TID}}(t) = \frac{1}{4}\frac{m_0}{m}\left(\frac{R}{a}\right)^3 \sum_{k=-\infty}^{\infty} k_2(2\Omega - kn)\, X_k^{-3,2}(e)\cos[2\theta - kM - \epsilon(2\Omega - kn)]\,, \tag{123}$$

$$S_{22,\mathrm{TID}}(t) = -\frac{1}{4}\frac{m_0}{m}\left(\frac{R}{a}\right)^3 \sum_{k=-\infty}^{\infty} k_2(2\Omega - kn) X_k^{-3,2}(e)\sin[2\theta - kM - \epsilon(2\Omega - kn)]\,. \tag{124}$$

Substituting these formulae in the right-hand side of (113), we get the tidal torque

$$\begin{aligned} T_{\mathrm{TID}} &:= -6\frac{\mathcal{G}mm_0R^2}{r^3}\left[C_{22,\mathrm{TID}}\cos 2(\theta - \upsilon) + S_{22,\mathrm{TID}}\sin 2(\theta - \upsilon)\right] \\ &= -\frac{3}{2}\frac{\mathcal{G}m_0^2}{R}\left(\frac{R}{a}\right)^6 \\ &\times \sum_{k=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} k_2(2\Omega - kn) X_k^{-3,2}(e) X_j^{-3,2}(e)\sin[(k-j)M + \epsilon(2\Omega - kn)]\,. \end{aligned} \tag{125}$$

Often, the average of this torque over the mean anomaly M is sufficient to describe the evolution of the system. Doing so, we impose $j = k$ and obtain the *secular tidal torque* provided in [9], namely,

$$\langle T_{\mathrm{TID}}\rangle_M = -\frac{3}{2}\frac{\mathcal{G}m_0^2}{R}\left(\frac{R}{a}\right)^6 \sum_{k=-\infty}^{\infty}\left(X_k^{-3,2}(e)\right)^2 k_2(2\Omega - kn)\sin\epsilon(2\Omega - kn)\,. \tag{126}$$

This torque, seen as a function of the rotation frequency Ω, is, up to the constant normalisation factor $\frac{3}{2}\mathcal{G}m_0^2R^5/a^6$, a sum of elementary torques $T_1(x) =$

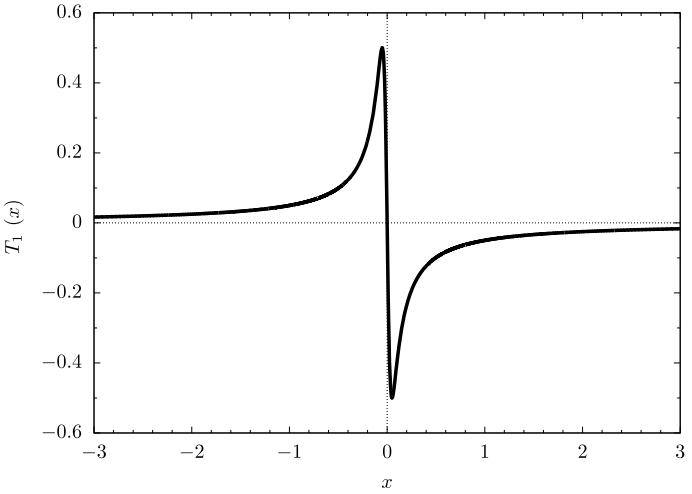


Fig. 7 Normalised typical elementary tidal torque $T_1(x)$ associated with one single harmonic of the Fourier decomposition of the orbit (cf. main text)

$-k_2(2x)\sin\epsilon(2x)$, called *kink*, weighted by the square of the Hansen coefficient $X_k^{-3,2}(e)$ and translated by $kn/2$ along the x-axis. For standard rheologies, $T_1(x)$ is an odd function of x, negative for $x > 0$, whose absolute value decreases to zero as $x \to \infty$. The typical shape of an individual kink T_1 is displayed in Fig. 7 and the overall tidal torque in Fig. 8. Depending on the rate at which $T_1(x)$ goes to zero as $x \to \infty$, $T_{\rm TID}$ can have one or several zeros located in the vicinity of $k:2$ spin-orbit resonances,[9] i.e., where $2\Omega \approx kn$.

8.3 Secular Evolution Out of Spin-Orbit Resonances

To get the secular evolution of the system, we take the average of both sides of (113). In the right-hand side, to the tidal torque, we shall add the torque $T_{\rm TRI}$, sometimes referred to as the *triaxial torque*, induced by the permanent deformation $C_{22,0}$.[10] Namely,

$$\frac{\mathrm{d}}{\mathrm{d}t}(C\Omega) = T_{\rm TRI} + T_{\rm TID}\,, \tag{127}$$

with $T_{\rm TID}$ given in (125) and

[9] It must be stressed that due to the vertical displacement of each kink imposed by its neighbours the total tidal torque does not cancel exactly at $k:2$ spin-orbit resonant states [25]. Exact resonances would only occur if the slope between the two extrema of individual kinks were infinite.

[10] We recall that $C_{22} = C_{22,0} + C_{22,\rm TID}$ and $S_{22} = S_{22,0} + S_{22,\rm TID}$ with $S_{22,0} = 0$ because the body-fixed frame of $\mathcal{B}$ is chosen such that B_0 is diagonal.

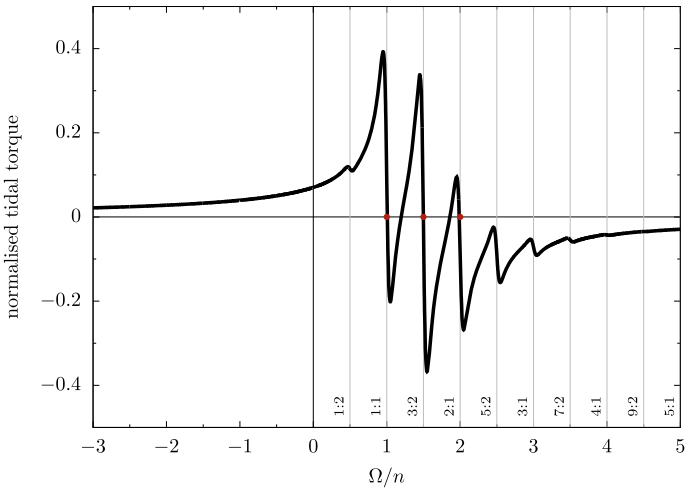


Fig. 8 Normalised typical tidal torque $\langle T_{\rm TID}\rangle_M$ as a function of the deformable body's spin rate Ω divided by the orbital mean motion n. A few spin-orbit resonances are represented by vertical lines. The three stable equilibria, defined by $\langle T_{\rm TID}\rangle_M = 0$ and $\partial\langle T_{\rm TID}\rangle_M/\partial\Omega < 0$, are highlighted by red points

$$T_{\rm TRI} = -6\frac{\mathcal{G}mm_0R^2}{r^3}C_{22,0}\cos 2(\theta - v). \tag{128}$$

The calculation of the torque $T_{\rm TRI}$ is presented in [12]. Its expansion in Fourier series gives

$$T_{\rm TRI} = -6\frac{\mathcal{G}mm_0}{R}\left(\frac{R}{a}\right)^3 C_{22,0}\sum_{k=-\infty}^{\infty} X_k^{-3,2}(e)\,\sin(2\theta - kM)\,. \tag{129}$$

We see that out of spin-orbit resonances this torque does not contain any constant term, therefore its average is nil: $\langle T_{\rm TRI}\rangle_M = 0$. Let us now develop the left-hand side of (127). We have

$$\frac{\rm d}{{\rm d}t}(C\Omega) = C\dot{\Omega} + \dot{C}\Omega\,. \tag{130}$$

The value of the moment of inertia C is dominated by $\mathrm{I}_\circ$ and can safely be replaced by its constant part in $C\dot{\Omega}$. However, because of the time derivative of C, the term $\dot{C}\Omega$ is of the same order of magnitude as $T_{\rm TID}$ and shall not be excluded a priori. Nevertheless, using (114) and (122), we get

$$\dot{C}\Omega = -\frac{1}{3}m_0R^2n\Omega\left(\frac{R}{a}\right)^3\sum_{k=-\infty}^{\infty} k_2(kn)\,k\,X_k^{-3,0}(e)\sin[kM - \epsilon(kn)] \tag{131}$$

goes to zero when averaged over the mean anomaly. Finally, the secular evolution of $\mathcal{B}$'s rotation out of spin-orbit resonance is only governed by the tidal torque $\langle T_{\rm TID}\rangle_M$ through

$$\dot{\Omega} = \frac{\langle T_{\rm TID}\rangle_M}{C} = \frac{3}{2}\frac{\mathcal{G}m_0^2}{CR}\left(\frac{R}{a}\right)^6 \sum_{k=-\infty}^{\infty} \left(X_k^{-3,2}(e)\right)^2 T_1(\Omega - kn/2)\,. \tag{132}$$

If the rotation speed Ω is initially greater than the synchronous frequency n, then, according to Fig. 8 it decreases until it reaches a $k:2$ spin-orbit resonance. At this stage, $\mathcal{B}$ can either remain captured in the resonant state or escape toward the next $(k-1):2$ spin-orbit resonance. Whether the capture takes place actually depends on the orientation of the triaxial body at the pericentre [24].

8.4 Libration in the Vicinity of the Synchronous Rotation

In the vicinity of the 1 : 1 synchronous rotation $\Omega \approx n$, it is instructive to compute the constant part of the Stokes coefficients induced by tides $J_{2,\rm TID}$, $C_{22,\rm TID}$ and $S_{22,\rm TID}$ [27]. From (122), (123) and (124), we have

$$(\bar{J}_{2,\rm TID})^{1:1} := \left\langle J_{2,\rm TID}\right\rangle_{\Omega=n} = k_2(0)\left(\frac{\Omega^2R^3}{3\mathcal{G}m} + \frac{1}{2}\frac{m_0}{m}\left(\frac{R}{a}\right)^3\right) X_0^{-3,0}(e)\,, \tag{133}$$

$$(\bar{C}_{22,\rm TID})^{1:1} := \left\langle C_{22,\rm TID}\right\rangle_{\Omega=n} = k_2(0)\frac{1}{4}\frac{m_0}{m}\left(\frac{R}{a}\right)^3 X_2^{-3,2}(e)\,, \tag{134}$$

$$(\bar{S}_{22,\rm TID})^{1:1} := \left\langle S_{22,\rm TID}\right\rangle_{\Omega=n} = 0\,. \tag{135}$$

Let us now determine the dominant terms in the expression of $T_{\rm TID}$ (125) in the vicinity of the synchronous rotation. Hansen coefficients have the property $X_k^{n,m}(e) = O(e^{|m-k|})$. Therefore, one can expect the dominant term to be that with $k = j = 2$. But the phase lag ϵ would be evaluated at the frequency $2\Omega - 2n = 0$ and $\epsilon(0) = 0$. Hence the dominant terms are those with $(k, j) \in \{(2, 1), (2, 3), (1, 2), (3, 2)\}$. Taking into account the parity of $k_2(x)$ and $\epsilon(x)$, we get

$$\begin{aligned}(T_{\rm TID})^{1:1} \approx & \frac{3}{2}\frac{\mathcal{G}m_0^2}{R}\left(\frac{R}{a}\right)^6 k_2(0)\left(X_3^{-3,2}(e) - X_1^{-3,2}(e)\right) X_2^{-3,2}(e) \\ & \times \left(\sin M - \frac{k_2(n)}{k_2(0)}\sin\left(M - \epsilon(n)\right)\right).\end{aligned} \tag{136}$$

Using Kepler's third law, the expansion of the Hansen coefficients (103)–(105), the expression of $(\bar{C}_{22,\rm TID})^{1:1}$ (134) and the definition $C_{22} = (B - A)/(4mR^2)$, we obtain

$$(T_{\rm TID})^{1:1} \approx 6\,n^2\,e\,(\bar{B}_{\rm TID} - \bar{A}_{\rm TID})\left(\sin M - \frac{k_2(n)}{k_2(0)}\sin\big(M - \epsilon(n)\big)\right). \tag{137}$$

In this equation, as for the Stockes coefficients J_2, C_{22} and S_{22}, we split the moments of inertia A, B and C into their permanent component A_0, B_0 and C_0 and their tidal deformation $A_{\rm TID}$, $B_{\rm TID}$, $C_{\rm TID}$ using

$$\begin{aligned} &\frac{A_0}{mR^2} = \frac{\mathrm{I}_\circ}{mR^2} - \frac{1}{3}J_{2,0} - 2C_{22,0}\,, \quad \frac{B_0}{mR^2} = \frac{\mathrm{I}_\circ}{mR^2} - \frac{1}{2}J_{2,0} + 2C_{22,0}\,, \\ &\frac{C_0}{mR^2} = \frac{\mathrm{I}_\circ}{mR^2} + \frac{2}{3}J_{2,0}\,, \end{aligned} \tag{138}$$

and

$$\begin{aligned} &\frac{A_{\rm TID}}{mR^2} = -\frac{1}{3}J_{2,\rm TID} - 2C_{22,\rm TID}\,, \quad \frac{B_{\rm TID}}{mR^2} = -\frac{1}{3}J_{2,\rm TID} + 2C_{22,\rm TID}\,, \\ &\frac{C_{\rm TID}}{mR^2} = \frac{2}{3}J_{2,\rm TID}\,. \end{aligned} \tag{139}$$

Moreover, to simplify the notation, instead of writting $(\bar{A}_{\rm TID})^{1:1}$, we simply write $\bar{A}_{\rm TID}$. Idem for $\bar{B}_{\rm TID}$ and $\bar{C}_{\rm TID}$.

Regarding the triaxial torque $T_{\rm TRI}$ (129), when applying the same method, we get at first order in eccentricity

$$(T_{\rm TRI})^{1:1} \approx 6\,n^2\,e\,(B_0 - A_0)\,\sin M\,. \tag{140}$$

For the expansion of $\dot{C}\Omega$ (131), we use $X_1^{-3,0}(e) \sim \frac{3}{2}e$ and get

$$\dot{C}\Omega = -\,n^2\,e\,(\bar{B}_{\rm TID} - \bar{A}_{\rm TID})\,\frac{k_2(n)}{k_2(0)}\sin(M - \epsilon(n))\,. \tag{141}$$

Finally, if we call $\bar{A} = A_0 + \bar{A}_{\rm TID}$, $\bar{B} = B_0 + \bar{B}_{\rm TID}$ and $\bar{C} = C_0 + \bar{C}_{\rm TID}$, the equation of motion (113) becomes

$$\ddot{\theta} = 6\,n^2\,e\left(\frac{\bar{B} - \bar{A}}{\bar{C}}\sin M - \frac{5}{6}\frac{\bar{B}_{\rm TID} - \bar{A}_{\rm TID}}{\bar{C}}\frac{k_2(n)}{k_2(0)}\sin\big(M - \epsilon(n)\big)\right). \tag{142}$$

Therefore, $\mathcal{B}$ librates with an amplitude $\Theta_{\rm elast}$ given by (see also [27])

$$\Theta_{\rm elast} = 6\,e\,\frac{\bar{B} - \bar{A}}{\bar{C}}\sqrt{1 - \frac{5}{3}\frac{\bar{B}_{\rm TID} - \bar{A}_{\rm TID}}{\bar{B} - \bar{A}}\frac{k_2(n)}{k_2(0)}\cos\epsilon(n) + \left(\frac{5}{6}\frac{\bar{B}_{\rm TID} - \bar{A}_{\rm TID}}{\bar{B} - \bar{A}}\frac{k_2(n)}{k_2(0)}\right)^2}\,. \tag{143}$$

Alternative expressions exist in the literature, see [27] for a comparison. It should be stressed that even in absence of permanent triaxiality $(B_0 - A_0)/C_0$, the thenceforth viscous body still librates with an amplitude

$$\Theta_{\text{visc}} = 6\,e\,\frac{\bar{B}_{\text{TID}} - \bar{A}_{\text{TID}}}{\bar{C}_{\text{TID}}}\sqrt{1 - \frac{5}{3}\frac{k_2(n)}{k_2(0)}\cos\epsilon(n) + \left(\frac{5}{6}\frac{k_2(n)}{k_2(0)}\right)^2}\,, \tag{144}$$

which means that $\mathcal{B}$ behaves like a rigid body even though it has no elasticity. Moreover, for a perfect fluid body, $k_2(n) = k_2(0) = k_f$ and $\epsilon(n) = 0$, i.e., there is no lag between the tidal bulge and the perturbing body and the amplitude is frequency independent. In that case also $\mathcal{B}$ is librating although with an amplitude Θ_{fluid} reduced by a factor 6 with respect to the rigid case (74), viz.,

$$\Theta_{\text{fluid}} = e\,\frac{\bar{B}_{\text{TID}} - \bar{A}_{\text{TID}}}{\bar{C}_{\text{TID}}}\,. \tag{145}$$

In the fluid case, the tidal torque is nil because the bulge is aligned with the direction of the external body. The libration of θ is actually due to the variation of the moment of inertia $C(t)$ which has to be compensated to keep a constant angular momentum $C\dot{\theta}$.

9 Conclusion

These lecture notes present a Lagrangian formalism of the rotation and tidal deformation of an extended body. After a general introduction to the topic of rotational motion describing different sets of coordinates and their associated equations of motion, the rotation and the deformation of a celestial body on a Keplerian orbit are studied. In particular, we determine the free libration modes (path and frequency) of a rigid body in the synchronous rotation state. Given that this problem possesses three degrees of freedom of rotation, it also has three eigenmodes respectively called libration in longitude, libration in latitude and wobble. For each of them, we calculate the proper frequency and the path in the configuration space. When the problem is driven by a periodic perturbation of the Keplerian orbit, we also compute the amplitude of the so-called "forced solution". In the case of the libration in latitude, we thereby retrieve the linear approximation of the Cassini state dynamics.

In a subsequent section, we allow for tidal deformation. This improvement of the model is made at the cost of additional generalised coordinates. Nevertheless, thanks to the matrix formulation of the problem, expressions remain relatively compact. In this part, we recall the definition of a Tisserand frame generalising the body-fixed frame concept to deformable bodies. We compute the equilibrium dynamical triaxiality of a fluid extended body in spin-orbit resonance, i.e., the mean of its moments of inertia (or of a combination of them) over time. These averaged moments of inertia allow in turn to compute the mean gravity field generated by the deformable body. Moreover, the comparison of a measured triaxiality with the equilibrium one tells us if a body can be in hydrostatic equilibrium or not. In the case where the body is in hydrostatic equilibrium, one can then infer its internal mass distribution.

In this section, we also discuss permanent triaxiality, representing departure from hydrostaticity, and the maximal amplitude it can reached before the body's material starts to fracture under the high pressure exerted by the deformation. In a last section, we analyse the long term evolution of a body's rotation speed under tidal torque. In particular, we revisit the libration in longitude problem, generalising the rigid body case treated beforehand. This full expression of the libration in longitude, compared with observations, provides insights on celestial body's Love numbers.

As a concluding remark, I would like to stress that these lecture notes were not intended to present new results nor to be exhaustive on the literature existing on this subject. The purpose was to present the Lagrangian formalism and its applications to classical problems encountered in the field. Moreover, often the geometry of the problems has been chosen as simple as possible to avoid cumbersome calculations although more general problems have already been solved in the literature. Nevertheless, I believe that these notes will allow the reader to face more general yet unsolved situations in a simple manner using the present formalism.

Acknowledgements These lecture notes have been made possible thanks to the numerous enlightening discussions I had on this subject. This is why I warmly thank all the colleagues with whom I had the opportunity to work. I am also very grateful towards the I-Celmech team who invited me to give this course.

References

1. Anderson, J.D., Jacobson, R.A., McElrath, T.P., Moore, W.B., Schubert, G., Thomas, P.C.: Shape, mean radius, gravity field, and interior structure of callisto. Icarus **153**(1), 157–161 (2001)
2. Anderson, J.D., Lau, E.L., Sjogren, W.L., Schubert, G., Moore, W.B.: Gravitational constraints on the internal structure of Ganymede. Nature **384**(6609), 541–543 (1996)
3. Anderson, J.D., Jacobson, R.A., Lau, E.L., Moore, W.B., Schubert, G.: Io's gravity field and interior structure. J. Geophys. Res. **106**(E12), 32963–32970 (2001)
4. Boué, G.: The two rigid body interaction using angular momentum theory formulae. Celest. Mech. Dyn. Astron. **128**(2–3), 261–273 (2017)
5. Boué, G., Rambaux, N., Richard, A.: Rotation of a rigid satellite with a fluid component: a new light onto Titan's obliquity. Celest. Mech. Dyn. Astron. **129**(4), 449–485 (2017)
6. Boué, G., Efroimsky, M.: Tidal evolution of the Keplerian elements. Celest. Mech. Dyn. Astron. **131**(7), 30 (2019)
7. Colombo, G.: Cassini's second and third laws. Astron. J. **71**, 891 (1966)
8. Correia, A.C.M., Ragazzo, C., Ruiz, L.S.: The effects of deformation inertia (kinetic energy) in the orbital and spin evolution of close-in bodies. Celest. Mech. Dyn. Astron. **130**(8), 51 (2018)
9. Alexandre C. M. Correia, Gwenaël Boué, Jacques Laskar, and Adrián Rodríguez. Deformation and tidal evolution of close-in planets and satellites using a Maxwell viscoelastic rheology. *Astronomy & Astrophysics*, 571:A50, November 2014
10. Correia, A.C.M., Rodríguez, A.: On the equilibrium figure of close-in planets and satellites. Astrophys. J. **767**(2), 128 (2013)
11. Danjon, A., Cassini, J.-D.: L'Astronomie **77**, 4 (1963)
12. Frouard, J., Efroimsky, M.: Tides in a body librating about a spin-orbit resonance: generalisation of the Darwin-Kaula theory. Celest. Mech. Dyn. Astron. **129**(1–2), 177–214 (2017)

13. Casajus, L.G., Zannoni, M., Modenini, D., Tortora, P., Nimmo, F., Van Hoolst, T., Buccino, D., Oudrhiri, K.: Updated Europa gravity field and interior structure from a reanalysis of Galileo tracking data. Icarus **358**, 114187 (2021)
14. Goossens, S., Lemoine, F.G., Sabaka, T.J., Nicholas, J.B., Mazarico, E., Rowlands, D.D., B.D. Loomis, Chinn, D.S., Neumann, G.A., Smith, D.E., Zuber, M.T.: A global degree and order 1200 model of the lunar gravity field using GRAIL mission data. In: Lunar and Planetary Science Conference, Lunar and Planetary Science Conference, page 1484, Mar. 2016
15. Hansen, P.A.: Entwickelung der products einer potenz des radius vectors mit dem sinus oder cosinus eines vielfachen der wahren anomalie in reihen. Abhandl. d. K. S. Ges. d. Wissensch : IV. S. Hirzel (1855)
16. Iess, L., Jacobson, R.A., Ducci, M., Stevenson, D.J., Lunine, J.I., Armstrong, J.W., Asmar, S.W., Racioppa, P., Rappaport, N.J., Tortora, P.: The tides of titan. Science **337**(6093), 457 (2012)
17. Jones, R.M.: Deformation Theory of Plasticity. Bull Ridge Publishing (2009)
18. Kaula, W.M.: Tidal dissipation by solid friction and the resulting orbital evolution. Rev. Geophys. Space Phys. **2**, 661–685 (1964)
19. Keane, J.T., Matsuyama, I.: Evidence for lunar true polar wander and a past low-eccentricity, synchronous lunar orbit. Geophys. Res. Lett. **41**(19), 6610–6619 (2014)
20. Lambeck, K.: Geophysical Geodesy: the Slow Deformations of the Earth Lambeck. Clarendon Press (1988)
21. Lemoine, F.G., Goossens, S., Sabaka, T.J., Nicholas, J.B., Mazarico, E., Rowlands, D.D., Loomis, B.D., Chinn, D.S., Neumann, G.A., Smith, D.E., Zuber, M.T.: GRGM900C: a degree 900 lunar gravity model from GRAIL primary and extended mission data. Geophys. Res. Lett. **41**(10), 3382–3389 (2014)
22. Love, A.E.H.: Some Problems of Geodynamics. Cambridge University Press (1911)
23. MacDonald, G.J.F.: Tidal friction. Rev. Geophys. Space Phys. **2**, 467–541 (1964)
24. Makarov, V.V.: Conditions of passage and entrapment of terrestrial planets in spin-orbit resonances. Astrophys. J. **752**(1), 73 (2012)
25. Makarov, V.V., Berghea, C., Efroimsky, M.: Dynamical evolution and spin-orbit resonances of potentially habitable exoplanets: the case of GJ 581d. Astrophys. J. **761**(2), 83 (2012)
26. Matsuyama, I., Nimmo, F.: Gravity and tectonic patterns of Mercury: effect of tidal deformation, spin-orbit resonance, nonzero eccentricity, despinning, and reorientation. J. Geophys. Res. (Planets) **114**(E1), E01010 (2009)
27. Noyelles, B.: Interpreting the librations of a synchronous satellite—How their phase assesses Mimas' global ocean. Icarus **282**, 276–289 (2017)
28. Peale, S.J.: Possible histories of the obliquity of Mercury. Astron. J. **79**, 722 (1974)
29. Pettengill, G.H., Shapiro, I.I.: Radar astronomy. Ann. Rev. Astron. Astrophys. **3**, 377 (1965)
30. Poincaré, H.: Sur une forme nouvelle des équations de la mécanique. Comptes rendus de l'Académie des Sciences **132**, 369–371 (1901)
31. Ragazzo, C., Paulo, S.: The theory of figures of Clairaut with focus on the gravitational modulus: inequalities and an improvement in the Darwin-Radau equation. J. Math. Sci. **14**, 1–14 (2020)
32. Ragazzo, C., Ruiz, L.S.: Dynamics of an isolated, viscoelastic, self-gravitating body. Celest. Mech. Dyn. Astron. **122**(4), 303–332 (2015)
33. Ragazzo, C., Ruiz, L.S.: Viscoelastic tides: models for use in Celestial Mechanics. Celest. Mech. Dyn. Astron. **128**(1), 19–59 (2017)
34. Stanton, J.: Generalized Cassini's Laws. Astron. J. **74**, 483 (1969)
35. Thomson, W.: On the rigidity of the earth. Philos. Trans. R. Soc. London Ser. **I**(153), 573–582 (1863)
36. Varshalovich, D., Moskalev, A., Khersonskii, V.: Quantum Theory of Angular Momentum. World Scientific (1988)
37. Zanazzi, J.J., Lai, D.: Triaxial deformation and asynchronous rotation of rocky planets in the habitable zone of low-mass stars. Mon. Not. R. Astron. Soc. **469**(3), 2879–2885 (2017)

Arnold Diffusion and Nekhoroshev Theory

Christos Efthymiopoulos and Rocío Isabel Paez

Abstract Starting with Arnold's pioneering work [2], the term "Arnold diffusion" has been used to describe the slow diffusion taking place in the space of the actions in Hamiltonian nonlinear dynamical systems with three or more degrees of freedom. The present text is an elaborated transcript of the introductory course given in the Milano I-CELMECH school on the topic of Arnold diffusion and its relation to Nekhoroshev theory. The course introduces basic concepts related to our current understanding of the mechanisms leading to Arnold diffusion. Emphasis is placed upon the identification of those invariant objects in phase space which drive chaotic diffusion, such as the stable and unstable manifolds emanating from (partially) hyperbolic invariant objects. Besides a qualitative understanding of the diffusion mechanisms, a precise quantification of the speed of Arnold diffusion can be achieved by methods based on canonical perturbation theory, i.e. by the construction of a suitable normal form at optimal order. As an example of such methods, we discuss the (quasi-)stationary-phase approximation for the selection of remainder terms acting as driving terms for the diffusion. Finally, we discuss the efficiency of such methods through numerical examples in which the optimal normal form is determined by a computer-algebraic implementation of a normalization algorithm.

Keywords Nekhoroshev theory · Arnold diffusion · Hamiltonian systems

C. Efthymiopoulos (✉)
Dipartimento di Matematica Tullio Levi-Civita, Università degli Studi di Padova, Via Trieste 63, 35121 Padova, Italy
e-mail: cefthym@math.unipd.it

R. I. Paez
School of Computer Science and Information Technology, University College Cork UCC, Cork, Ireland
e-mail: rpaez@ucc.ie

G. Baù et al. (eds.), *New Frontiers of Celestial Mechanics: Theory and Applications*, Springer Proceedings in Mathematics & Statistics 399,
https://doi.org/10.1007/978-3-031-13115-8_5

1 Introduction

In some introductory texts (see, for example, [17, 46, 58]), the topic of Arnold diffusion is introduced by a simplified topological argument, related to a difference between the cases of invariant tori in Hamiltonian systems with $n = 2$ and with $n \geq 3$ degrees of freedom. Consider a n-degrees of freedom Hamiltonian system $H(q, p)$, $q \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, whose phase space contains a large measure of n-dimensional Kolmogorov–Arnold–Moser (KAM) tori [1, 43, 51]. Any orbit $(q(t), p(t))$ with initial conditions (q_0, p_0) on a KAM torus remains forever confined to the torus. Any orbit with initial conditions (q_0, p_0) belonging to the *complement*, in phase space, with respect to the set of all KAM tori, remains confined to the $(2n-1)$-dimensional manifold defined by the orbit's constant energy value $\mathcal{M}_E := \{(q, p) \in \mathbb{R}^{2n} : H(q, p) = E = H(q_0, p_0)\}$. Take first $n = 2$. Thus, $\dim(\mathcal{M}_E) = 3$. Suppose there is a KAM torus $\mathcal{T}$ embedded in the same energy manifold. We have $\dim(\mathcal{T}) = 2$. Since the torus's dimension differs just by one from the dimension of the energy manifold, $\mathcal{T}$ divides $\mathcal{M}_E$ into two parts, which can be called the 'interior' and the 'exterior' of the torus. Furthermore, since $H(q, p)$ is autonomous, there can be no trajectory going from the interior to the exterior of the torus; such a trajectory would necessarily have to cross transversally the torus at a point $(q(t_c), p(t_c)) \in \mathcal{T}$ at some time t_c, but this is impossible since the flow on the torus is invariant, i.e., the initial condition $q = q(t_c)$, $p = p(t_c)$ would lead necessarily to a trajectory confined on the torus. We roughly refer to this as the 'dividing property' of KAM tori in systems with $n = 2$ degrees of freedom. On the other hand, there is no dividing property of the KAM tori when $n \geq 3$, since, in that case $\dim(\mathcal{M}_E) - \dim(\mathcal{T}) \geq 2$. For example, when $n = 3$ we have $\dim(\mathcal{M}) = 5$, and $\dim(\mathcal{T}) = 3$, thus $\mathcal{T}$ cannot divide $\mathcal{M}$ into disconnected sets. To visualize this just lower all dimensions in the above examples by one: hence, a circle (dimension 1) divides a plane (dimension 2) to the interior and the exterior of the circle, while a circle embedded in Euclidean space (dimension 3) cannot divide the latter into disconnected sets.

The non-existence of topological barriers when $n \geq 3$ renders a priori possible to have long excursions of the chaotic orbits throughout the whole constant energy manifold. However, two questions become immediately relevant: (i) can we prove that the chaotic orbits *do really undergo* those (topologically allowed) arbitrarily long chaotic excursions? (ii) is the timescale involved short enough to make the phenomenon relevant and worth of further study as regards applications in physical systems (including, for the purposes of the present course, systems of interest in celestial mechanics or astrodynamics)?

We refer to question (i) above as the problem of the *existence of Arnold diffusion*. In the words of Lochak's influential review [49], it is the problem of demonstrating that "topological transitivity on the energy surface generically takes place". We refer, instead, to question (ii) as the problem of how to quantitatively estimate the *speed of Arnold diffusion*. Addressing this question in the context of particular problems encountered in physics and astronomy requires a (partly heuristic) use of compu-

tational techniques, as discussed in detail, for example, in a well known review by Chirikov [13]. It is worth mentioning that, after about 60 years of research, only partial answers are available today regarding both questions. In particular, the existence of Arnold diffusion has been rigorously established in various cases of so-called *a priori unstable* systems (see [10, 11, 22, 32]). Instead, it remains an open problem in the far more difficult case of *a priori stable* systems (see Sect. 3 for definitions). In the latter case, we avail, however, ample numerical evidence of the global drift of the trajectories within the so-called Arnold's *web of resonances*, as visualized in a series of beautiful numerical studies ([30, 37, 38, 44]; see [45] for a review). In fact, the visualization of the Arnold web in a priori stable systems was made possible by the use of techniques allowing to carefully choose initial conditions along the thin resonance layers in phase space marked by the web of resonances. The Fast Lyapunov Indicator (FLI, [28]) is an example of such technique.

As emphasized by Lochak [49], a demonstration that the Arnold diffusion really takes place requires establishing the existence of a *mechanism of transport* for the weakly chaotic orbits within the Arnold web. Arnold's original example [2] actually describes such a mechanism. This is based on proving the existence of heteroclinic intersections between the stable and unstable manifolds emanating from a set of nearby partially hyperbolic low-dimensional tori arranged in a so-called 'transition chain'. One initially demonstrates that two nearby tori, of a small distance, say, $\mathcal{O}(\delta)$, where δ is a small parameter, exhibit a 'splitting of the separatrices' (their stable and unstable manifolds) such that these manifolds develop heteroclinic intersections. Let τ_i, $i = 1, 2, \ldots$ be a sequence of tori, τ_i being neighbor to τ_{i-1}, τ_{i+1}. Assume we know that the unstable manifold emanating from τ_i has a heteroclinic intersection with the stable manifold ending at τ_{i+1}. Then, there is a 'doubly asymptotic' orbit which tends to τ_i as $t \to -\infty$, while it tends to τ_{i+1} forward in time as $t \to \infty$. Such orbits can be established for any pair τ_i, τ_{i+1}, $i = 1, 2, \ldots$, but of course they cannot themselves be the orbits of Arnold diffusion, since they never go very far either from τ_i or τ_{i+1}. On the other hand, invoking a so-called 'shadowing lemma' (see [20] for a review), one demonstrates that there are true orbits of the system which shadow the whole chain of heteroclinic orbits established in the above way. Thus, these shadowing orbits undergo Arnold diffusion. A quick estimate of the speed of diffusion is obtained as follows: upon completion of one cycle of the transition mechanism, the trajectory has traveled a distance $S = \mathcal{O}(\delta)$ in a time $T_{i,i+1} \approx T_s$, which roughly coincides with the time required to cover one homoclinic loop close to the separatrix of the resonance associated with the unstable tori τ_i (see Sect. 2 below). Hence, the local speed of Arnold diffusion is $V_{AD} \approx \delta / T_s$, where both parameters δ and T_s depend on the small parameters of the problem under study. Of course, this is an oversimplified estimate. Estimates of practical interest are rather hard to obtain, as explained in the sections to follow. On the other hand, the topic of how to describe itself the one-step transition of the chaotic trajectories far from, and then back to the asymptotic ends of the intersecting manifolds has been developed substantially in recent years, leading to the concept of the so-called 'scattering map' (see [19, 21]). Applications of the scattering map technique in Celestial Mechanics are discussed, in particular, by [6] (see also [8] and references therein).

Regarding numerical investigations of Arnold diffusion, since this is a slow phenomenon its revelation requires a rather high computing power and the capacity to numerical propagate large sets of trajectories over long integration times. Owing to its complexity, the numerical investigation of the weakly chaotic diffusion has so far been limited to few DOF dynamical systems, including several systems of particular interest for dynamical astronomy (see an extensive, but only indicative, list of references in Sect. 4 of [25]). However, it is unclear whether the notion of Arnold diffusion can be useful for the analysis of the diffusive processes in all those models. On the other hand, there are cases, in particular around normally hyperbolic invariant objects in the restricted three-body problem, where Arnold diffusion has been explicitly demonstrated to apply (see, for example, [6–8, 27, 52]).

The present tutorial is organized as follows: Sect. 2 presents in some detail the original example discussed in [2], serving to introduce most elements of the conceptual framework for the discussion of Arnold diffusion. Section 3 deals with the case of a priori stable systems and with the connection of Arnold diffusion with Nekhoroshev theory. Finally, Sect. 4 discusses various semi-analytical approaches to the quantification of the speed of Arnold diffusion.

2 Arnold's Example

The Hamiltonian model presented by Arnold in [2] is

$$H(q, \phi_1, t, p, J_1) = \frac{1}{2}p^2 + \frac{1}{2}J_1^2 + \epsilon\,(\cos q - 1)\;(1 + \mu\,(\sin\phi_1 + \cos t))\;, \quad (1)$$

It is a model of a pendulum (variables (q, p)) coupled with a rotator (variables (ϕ_1, J_1)) via the time-dependent term $\epsilon\mu \cos q \cos t$. We will assume $\epsilon > 0$ and fixed, while varying μ, with $|\mu| << \epsilon$. The Hamiltonian can be formally extended to 3DOF autonomous by introducing the angle $\phi_2 = t$ conjugated to a dummy action J_2:

$$H \to H(q, \phi_1, \phi_2, p, J_1, J_2) = \frac{1}{2}p^2 + \frac{1}{2}J_1^2 + J_2 + \epsilon\,(\cos q - 1)\;(1 + \mu\,(\sin\phi_1 + \cos\phi_2))\;. \quad (2)$$

For $\mu = 0$, we have $\dot{J}_1 = \dot{J}_2 = 0$, thus the actions remain invariant along the trajectories. For any values (J_1, J_2), the angles ϕ_1, ϕ_2 evolve linearly with frequencies $\omega_1 = J_1, \omega_2 = 1$. Thus, changing the value of J_1, we can obtain any desired frequency ratio $\omega_1/\omega_2 = J_1$ (the dummy action J_2 can be set initially to any value (e.g. $J_2(0) = 0$) without consequences for the dynamics).

Consider now the case $\mu \neq 0$. For generic trajectories, we obtain $\dot{J}_1 \neq 0$, $\dot{J}_2 \neq 0$. However, there is a particular set of initial conditions for which the trajectories preserve the actions:

$$\tau(J_1, J_2) = \left\{ q = p = 0, J_1 = const, J_2 = const, (\phi_1, \phi_2) \in \mathbb{T}^2 \right\} \quad . \tag{3}$$

Taking Hamilton's equations for the complete system:

$$\begin{array}{ll} \dot{q} = p, & \dot{p} = -\epsilon \sin q (1 + \mu(\sin\phi_1 + \cos\phi_2)) \\ \dot{\phi}_1 = J_1, & \dot{J}_1 = \ \epsilon\mu(\cos q - 1)\cos\phi_1 \\ \dot{\phi}_2 = 1, & \dot{J}_2 = -\epsilon\mu(\cos q - 1)\sin\phi_2 \end{array} \tag{4}$$

we immediately find that any initial condition in the set $\tau(J_1, J_2)$ leads to $\dot{q} = \dot{p} = 0 = \dot{J}_1 = \dot{J}_2 = 0$, while $\dot{\phi}_1 = J_1(t) = const$, $\dot{\phi}_2 = J_2(t) = const$. Thus, $\tau(J_1, J_2)$ is invariant under the flow and homeomorphic to the 2D-torus $(\phi_1, \phi_2) \in \mathbb{T}^2$. We will denote by $\mathcal{T}$ the invariant set formed by the family of all the tori $\tau(J_1, J_2)$ $(J_1, J_2) \in \mathbb{R}^2$.

The invariance of the tori $\tau(J_1, J_2)$ crucially relies on having set (q, p) as $(q, p) = (0, 0)$. We now wish to explore what will happen if, instead, we choose the initial condition (q_0, p_0) close to, but not equal to $(0, 0)$. For example, we can set $q_0 = 0$, $p_0 \neq 0$, with $|p_0| < D$ and D small, and $(J_{1,0}, J_{2,0}, \phi_{1,0}, \phi_{2,0})$ chosen at will. We then want to understand the future evolution, in particular of the actions $J_1(t)$, $J_2(t)$, as a consequence of choosing initial conditions in the neighborhood of, but not exactly on the torus $\tau(J_{10}, J_{20})$. Addressing this question requires the use of a mixture of analytical as well as geometric arguments. Let us summarize some basic ones:

2.1 Existence of KAM Tori

We can demonstrate the existence of Kolmogorov–Arnold–Moser (KAM) tori for a Cantor set (of non-zero measure) of initial conditions p_0 along the line $q = 0$. Decomposing the Hamiltonian as:

$$H(q, \phi_1, \phi_2, p, J_1, J_2) = H_0(p, J_1, J_2) + \epsilon H_1(q, \phi_1, \phi_2, p, J_1, J_2; \mu) \tag{5}$$

where $H_0 = \frac{1}{2}(p^2 + J_1^2) + J_2$, $H_1 = \epsilon\,(\cos q - 1)\ (1 + \mu\,(\sin\phi_1 + \cos\phi_2))$, the Hamiltonian H_0 satisfies the iso-energetic non-degeneracy condition:

$$\det \begin{pmatrix} \mathrm{Hess}(H_0) & \nabla_I(H_0) \\ (\nabla_I(H_0))^T & 0 \end{pmatrix} = 0 \tag{6}$$

where $\mathrm{Hess}(H_0)$ is the 3×3 Hessian matrix of H_0 with respect to $I \equiv (p, J_1, J_2)$. Thus, the necessary conditions for the Kolmogorov theorem [43] hold, namely:

Theorem (Kolmogorov 1954) *There exist positive constants* ϵ_0, γ, τ *such that, for* $|\epsilon| < \epsilon_0$, *and* (p_0, J_{10}) *such that the frequencies* $\omega_p = (\partial H_0/\partial p)_{p=p_0} = p_0, \omega_1 = J_1$, $\omega_2 = 1$ *satisfy the Diophantine condition*

$$|k_p\omega_p + k_1\omega_1 + k_2\omega_2| > \frac{\gamma}{k^\tau} \tag{7}$$

where $k = |k_p| + |k_1| + |k_2|$, *the trajectory with initial conditions* $p(0) = p_0$, $J_1(0) = J_{10}$), $q(0) = 0$, $J_2(0) = J_{20} \in \mathbb{R}$, *as well as* $(\phi_1(0), \phi_2(0)) \in \mathbb{T}^2$ *lies in a three-dimensional torus, where all phase-space co-ordinates evolve quasi-periodically with the frequencies* $(\omega_p, \omega_1, \omega_2)$.

The above theorem can be proven by the construction of the so-called *Kolmogorov normal form* in the neighborhood of the chosen initial conditions. The value of γ restricts the measure of initial conditions satisfying the Diophantine condition. By number-theoretical arguments we find $|p_0| > D = \mathcal{O}(\gamma)$, hence motions very close to $p_0 = 0$ cannot be quasi-periodic.

2.2 Semi-analytical ('Melnikov') Approach

In order to deal with non-quasiperiodic motions, very close to the torus $p_0 = 0$, we can try to approximate the evolution of the variables $(\phi_1, \phi_2, J_1, J_2)$ by a model in which the evolution in the variables $(q(t), p(t))$ is a priori modeled via some 'near-separatrix' analytical approximation $(q_s(t; \varepsilon_s), p_s(t; \varepsilon_s))$ based on the pendulum model (or, in general, the model of resonance giving rise to a particular form of the separatrix). This strategy is explored heuristically in a well known review on Arnold diffusion by Chirikov [13] and set in a rigorous base in [42]. It is based on the remark that choosing (q_0, p_0) very close to the values (0, 0) leads to a motion in the variables $(q(t), p(t))$ which can be modeled as a sequence of stochastic alterations between pendulum librations or rotations, each with nearly conserved pendulum energy

$$h_s(q, p) = \varepsilon_s = \frac{1}{2}p^2 + \epsilon(\cos q - 1) \tag{8}$$

with $\varepsilon_s \approx \varepsilon_{s,0} = 0$ (corresponding to the invariant torus $(q, p) = (0, 0)$). Figure 1 exemplifies this approach. The figure shows the evolution of the trajectory with initial conditions $q(0) = 0$, $\phi_1(0) = 0$, $\phi_2(0) = 0$, $p(0) = 5 \times 10^{-5}$, $J_1(0) = 0.3\sqrt{2}$, $J_2 = 0$, under the *complete* flow (2), with $\epsilon = 0.03$ and $\mu = 0.01$. Since the coupling term between pendulum and the rest of the system has size $\mathcal{O}(\mu\epsilon)$, with $\mu = 0.01$ this term is two orders of magnitude smaller than the $\epsilon \cos q$ term defining the pendulum separatrix. As a consequence, the 'splitting' of the separatrix will be quite small. This means that there will be only a small error in approximating the evolution of $(q(t), p(t))$ *as if it was governed only by the pendulum Hamiltonian* $h_s(q, p)$ (Eq. (8)). Figure 1 indicates that this is essentially correct. Denote by $R(+)$, $R(-)$ a pen-

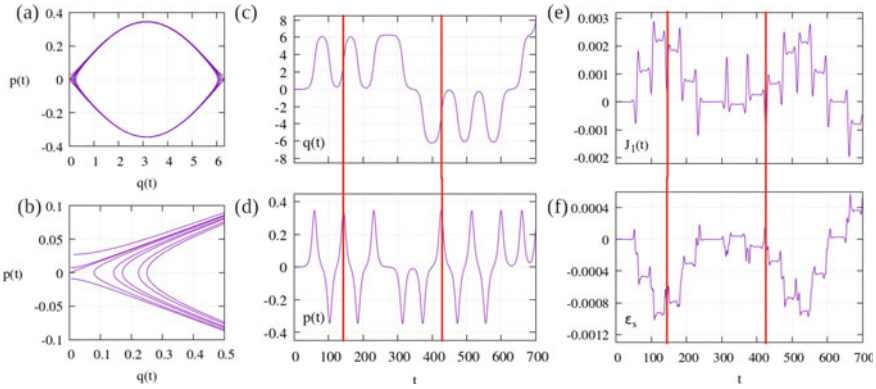


Fig. 1 Evolution of the orbit with initial conditions $\phi_1(0) = 0$, $\phi_2(0) = 0$, $p(0) = 5 \times 10^{-5}$, $J_1(0) = 0.3\sqrt{2}$, $J_2 = 1$, under the flow of the Hamiltonian (2) with $\epsilon = 0.03$ and $\mu = 0.01$: $p(t)$ versus $q(t)$ in panel (**a**) and (**b**), $q(t)$ in panel (**c**),$p(t)$ in panel (**d**), $J_1(t)$ in panel (**e**), in the time interval $t \in [0, 700]$. **f** Evolution of the pendulum energy ε_s for the same orbit. In **a** the angle $q(t)$ is shown modulo 2π. The two red vertical lines in panels (**c**) to (**f**) are helping guides to the eye: they indicate two different moments where the trajectory passes from the uppermost point of the separatrix. All jumps in $J_1(t)$ occur at such passages

dulum rotation with the Hamiltonian $h_s(q, p)$ and with $p > 0$ or $p < 0$ respectively, and by $L(+)$, $L(-)$ the upper and lower parts (again $p > 0$ or $p < 0$) of a librational curve in the same Hamiltonian. Then, the evolution of $p(t), q(t)$ in Fig. 1 can be represented as a sequence of segments of pendulum librational or rotational curves. Up to $t = 700$ we have

$$R(+), L(-), L(+), L(-), L(+), R(-), R(-), L(+), L(-), L(+), L(-), R(+), R(+), \ldots$$

Denoting by $T_{s,i}, i = 1, 2, \ldots$ the time it takes to accomplish one segment, the times $T_{s,i}$ can be estimated as the times between two successive local extrema in Fig. 1c. We find that $T_{s,i}$ has value nearly always around $T_s \lesssim 100$. Also, using the values $q(t_i)$, $p(t_i)$ at the times t_i of the local extrema of the curve $q(t)$, we can compute a sequence of corresponding pendulum energies $\varepsilon_i = h_s(q(t_i), p(t_i))$ characteristic of each segment.

Chirikov [13] proposed a model to study the qualitative properties of the mapping $(q(t_i), p(t_i)) \to (q(t_{i+1}), p(t_{i+1}))$, or, equivalently, $\varepsilon_i \to \varepsilon_{i+1}, t_i \to t_{i+1}$, called, by him the *whisker mapping* ('whiskers' meaning the separatrices of the torus $(q, p) = (0, 0)$). Figure 1f shows the first few transitions in the energy values ε_s. In every step, $\varepsilon_s(t)$ takes nearly constant value in a 'plateau', separated from the next plateau by a rapid oscillation. These oscillations take place mid-way along each homoclinic transition far from and back to the neighborhood of the torus $(q, p) = (0, 0)$.

We now discuss how to exploit the above empirical information in order to model the evolution in the remaining variables J_1, J_2, ϕ_1, ϕ_2 along such homoclinic transitions. The so-called 'Melnikov approach' consists essentially of the following

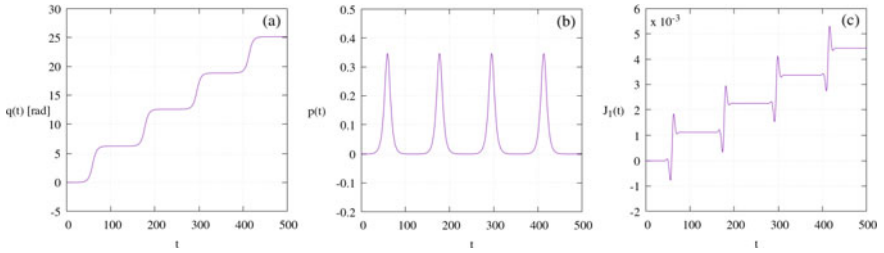


Fig. 2 Evolution of the pendulum solution **a** $q_s(t)$, **b** $p_s(t)$, for the same initial condition as in Fig. 1, namely $q_s(0) = 0$, $p_s(0) = 5 \times 10^{-5}$, but following the pendulum equations (Eq. (10)). **c** Evolution of the action $J_1(t)$ under the equations of the Melnikov approximation (Eq. (9)). We observe that $J_1(t)$ exhibits jumps in time which are qualitatively similar as those of the complete model, shown in panel (**e**) of Fig. 1

approximation: in the interval $t_i < t < t_{i+1}$, we will evolve the remaining variables according to the approximate system

$$\begin{aligned} \dot{\phi}_1 &= J_1, \quad & \dot{J}_1 &= \epsilon\mu(\cos q_s(t) - 1)\cos\phi_1 \\ \dot{\phi}_2 &= 1, \quad & \dot{J}_2 &= -\epsilon\mu(\cos q_s(t) - 1)\sin\phi_2 \end{aligned} \tag{9}$$

which is the same as the original system but with $q(t)$, $p(t)$ substituted with by the solutions $q_s(t)$, $p_s(t)$ of the pendulum equations

$$\dot{q}_s = p_s, \quad \dot{p}_s = -\epsilon \sin q_s \tag{10}$$

with initial conditions $q_s = q(t_i)$, $p_s = p(t_i)$.

Figure 2 shows the evolution under the approximate Eqs. (10) and (9), starting with the same initial condition as in Fig. 1, which belongs to the upper rotation domain of the pendulum ($q(0) = 0$, $p(0) > 0$). Since we now integrate the exact pendulum equations we obtain a periodic evolution of the angle q completing a circle at the period $T(\varepsilon_s)$ given by

$$T_s(\varepsilon_s) \simeq \int_0^{2\pi} \frac{dq}{\sqrt{2\,(\varepsilon_s - \epsilon(\cos q - 1))}} = \frac{32}{\sqrt{\epsilon}} \ln\left(\frac{|\varepsilon_s|}{\epsilon}\right) . \tag{11}$$

However, the action variable $J_1(t)$ (Fig. 2c) undergoes abrupt jumps of size 10^{-3} every time when the pendulum variables are mid-way along accomplishing one homoclinic transition.

The jumps in Fig. 2c are qualitatively quite similar to the jumps seen in the real orbit (Fig. 1e). In fact, the real jumps can be easily modeled by one further simplification: since all along the depicted solution $J_1(t)$ undergoes only a small ($\mathcal{O}(10^{-3})$) variation around the initial value $J_{10} = 0.3\sqrt{2}$, we can approximate the solution of the angular equation $\dot{\phi}_1(t) = J_1(t)$ by $\phi_1(t) = \phi_{1,0} + J_{10}(t - t_0)$, where $\phi_{1,0}$ is the

value of the angle ϕ_1 at the starting time t_0 of one homoclinic transition. We also approximate the solution $q_s(t)$ by the one holding along the pendulum separatrix:

$$q_s(t) \approx 4 \arctan\left(e^{\sqrt{\epsilon}(t-t_0-T_s/2)}\right) \quad , \tag{12}$$

with T_s still given by Eq. (11) (this last approximation is not really needed, but makes the computation easier with respect to the pendulum solution for the exact initial conditions given in terms of elliptic functions). As shown in Fig. 3a, the separatrix solution (12) fits the evolution of $q(t)$ along the first homoclinic transition as obtained numerically by the complete model (4) up to a time $t \approx 80$, where the real orbit starts its second homoclinic transition. Using the above approximations, all quantities in the r.h.s. of the differential equation for J_1 in the system (9) becomes explicit functions of the time t. Then, the approximative solution $J_1(t)$ can be obtained by quadratures:

$$J_1^{(M)}(t) = J_1(0) + \epsilon\mu \int_0^t \left(\cos(4 \arctan(\exp(\sqrt{\epsilon}(t' - T_s/2)))) - 1\right) \cos(\phi_{10} + J_{10}t')dt' \tag{13}$$

An integral of the form (13) is called a 'Melnikov integral'. It has the distinguishing feature that the integrand contains trigonometric functions $\cos\phi$, with $\phi = m_q q + m_1\phi_1 + m_1\phi_2$, $(m_q, m_1, m_2) \in \mathbb{Z}^3$, for some of which the evolution is not linear in time, as for example, the angle q which follows the near-separatrix pendulum solution (13). Figure 3b shows the comparison between the 'Melnikov' model $J_1^{(M)}(t)$ and the real evolution of the same variable up to the end of the first homoclinic transition, showing an excellent fit for the observed jump of the action $J_1(t)$.

How can we understand this success of the 'Melnikov approximation'? Of course the answer is hidden in the properties of the quadrature (13). As a coarse remark, by the equation for $\dot{J}_1(t)$ in (9)), the evolution of $J_1(t)$ is determined by the terms $\cos(q + \phi_1)$, $\cos(q - \phi_1)$ and $\cos\phi_1$. We saw that ϕ_1 evolves nearly linearly $\phi_1(t) \approx \phi_1(0) + J_{10}t$, so the integral $\int_0^t \cos(\phi_1(t'))dt' \approx \frac{1}{J_{10}} \sin(\phi_1(0) + J_{10}t)$ will only produce some rapid oscillation in the evolution of $J_1(t)$. The remaining terms, however, $\cos(q + \phi_1)$, $\cos(q - \phi_1)$ depend on the angle q, which evolves approximately by the pendulum trajectory of Eq. (12) (as shown in Fig. 3a). Now, the pendulum trajectory spends most of the time near the unstable origin, hence we have $\dot{q} \approx 0$ there. On the other hand the speed $\dot{q}$ in the middle of the homoclinic transition can be estimated as $\dot{q}(t) \approx 2\sqrt{\epsilon}$ (equal to $\dot{q}_s(t = T_s/2)$ in Eq. (12)). Thus, the curve $q(t)$ consists, essentially, of three parts, marked in Fig. 3a by A, B, and C respectively. In the domains A,C the curve is nearly horizontal, and $\cos(q \pm \phi_1) \simeq \cos(\phi_1)$, thus the integrals $\int \cos(q \pm \phi)$ yield essentially the same oscillatory behavior as for the integral $\cos\phi_1$ alone. In the domain B, instead, we have a slower evolution of the angle $q - \phi_1$: in our example we have $\dot{q} - \dot{\phi}_1 \simeq 2\sqrt{\epsilon} - J_{10} = -0.07785\ldots$ in B, compared to $\dot{q} - \dot{\phi}_1 \simeq \sqrt{2\epsilon} = 0.34\ldots$, $J_{10} = 0.4242\ldots$ in A or C. As a consequence, the curve $\cos(q(t) - \phi_1(t))$ develops an approximate 'plateau' near the time $t = T_s/2 \simeq 59$ (Fig. 3c). Since the integrand of the Melnikov integral in (13) temporarily stabilizes to a constant value, the integral will give a locally linear evo-

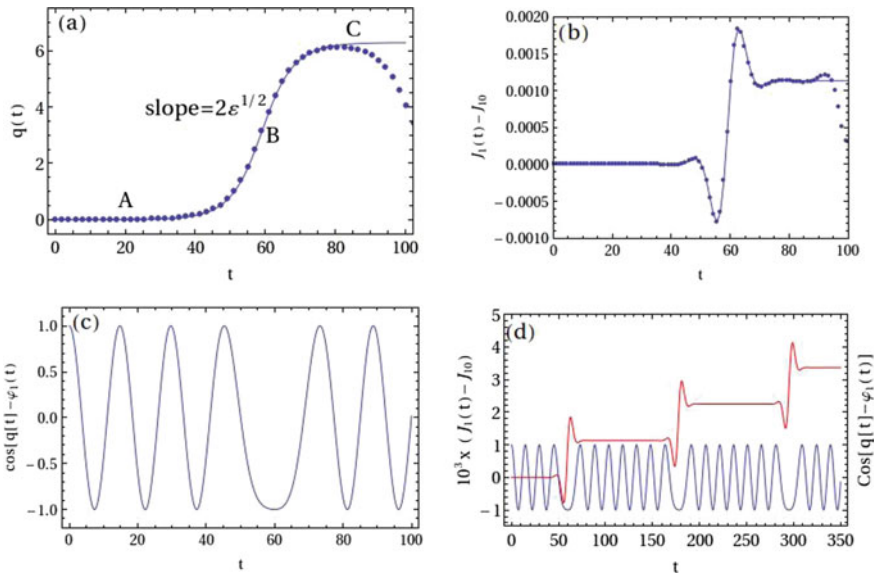


Fig. 3 **a** The evolution of the variable $q(t)$ along the first homoclinic transition, as obtained by numerical integration in the complete model (4) (points), and with the model of Eq. (12) (solid curve). **b** The first observed numerical jump in $J_1(t)$ (points) against the prediction of the model of Eq. (13). **c** The curve $\cos(q(t) - \phi_1(t))$ in the time interval corresponding to the first jump. **d** Left axis: several jumps in the variable $J_1(t)$ compared with (right axis) the evolution of $\cos(q(t) - \phi_1(t))$. The jumps take place at precisely those points where the phase $q - \phi_1$ forms a local plateau, departing from a pure oscillation

lution of $J_1(t)$, thus causing a quick jump, lasting roughly as the time duration of B. After exit from B, the $J_1(t)$ returns to an oscillatory evolution, which keeps up to the next homoclinic transition. More jumps then occur at each successive homoclinic transition, as shown in Fig. 3d.

Comparing the above picture with Fig. 1e, we do now interpret qualitatively the nature of the jumps, but we still need to understand why the jumps differ in size and/or sign. The sequences of times where jumps occur can be estimated by $t_{i+1} - t_i \approx T_s(\varepsilon_{s,i})$, with T_s given by Eq. (11). These times are of similar order, but different one from the other even for a small change in ε_s (compare the times T_s when $\varepsilon_s = 10^{-5}$ or 10^{-3}). As a consequence, at the starting point of each homoclinic transition, the orbit is at a different value of the starting angle $\phi_{1,0}$. However, as shown in Fig. 4, according to the value of $\phi_{1,0}(t_i)$ we can obtain jumps in J_1 of various sizes, positive or negative. Under the assumption that the sequence $\phi_{1,0}(t_i)$ is random ('random phase approximation'), this leads to a random walk model for the variations of $J_1(t)$. In reality, long correlations can survive in the sequence $\phi_{1,0}(t_i)$, and the diffusion in $J_1(t)$ can partly loose its normal character (typically the dynamics becomes sub-diffusive,

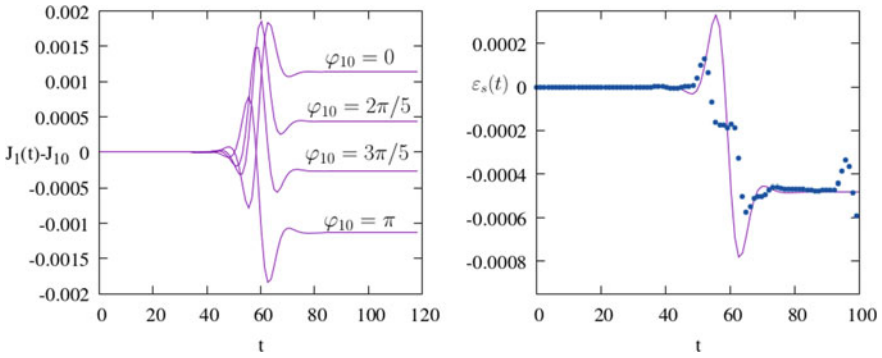


Fig. 4 *Left:* the jumps in the variable J_1 obtained through formula (13) by changing the initial angle ϕ_0 according to the values indicated in the figure. *Right:* The jump in the pendulum energy $\varepsilon_s = p^2/2 + \epsilon(\cos q - 1)$ as computed for the numerical orbit in the complete model (points) and with the 'Melnikov model' of Eq. (14)

see [50]). Also, using the Melnikov approach, we may compute a continuous in time approximation for the evolution of the energy $\varepsilon_s(t)$

$$\varepsilon_s = \varepsilon - \frac{1}{2}(J_1^{(M)})(t))^2 - J_2^{(M)}(t) - \epsilon\mu\,(\cos(q_s(t)) - 1)\,(\sin(\phi_{10} + J_{10}t) + \cos(\phi_{20} + t))\ , \tag{14}$$

where $J_2^{(M)}(t)$ is the 'Melnikov' model for the evolution of the action J_2, analogous to the model (13) for the action J_1. The right panel in Fig. 4 shows the evolution of the pendulum energy $\varepsilon_s(t)$ for the first jump in the real orbit and as obtained by the model (14), showing again a good fit. Then using all the above approximations, we can arrive at a heuristic model for Chirikov's 'whisker map'. While deterministic, in practice this model leads to nearly random sequences ε_i, $\phi_{1,i}$, that is, to a stochastic process for the evolution of the orbit in the action space. Estimating the value of the diffusion coefficient relies on some semi-analytical approaches, as discussed in Sects. 3 and 4 below.

As a final comment, one can remark that the 'plateaus' of the curve $\cos(q - \phi_1)$, responsible for the jumps Fig. 3d, are due to the tuning of the values of $\dot{q}$ and $\dot{\phi}_1 \simeq J_{10}$ at region B of Fig. 3. This tuning is rather exceptional, and was essentially imposed for illustration purposes by the choice of the initial condition J_{10}. Generic initial conditions instead (as, for example, choosing J_{10} one order of magnitude larger) will destroy such tuning. Does this imply that there is no more drift in action space by jumps as the above? As will be discussed in Sect. 4, we can make a number of steps of perturbation theory, seeking to eliminate altogether the now useless combinations $\cos(q - \phi_1)$, $\cos(q + \phi)$ and prove perpetual stability for the actions J_1 and J_2. However, doing so generates new 'dangerous' harmonics along the normalization process (see, for example, [53]). As higher order harmonics $\cos(m_q q + m_1\phi_1)$ are generated by the normalization, there will eventually appear some harmonics causing important jumps. Recalling that the jumps always take place in the domain

B of Fig. 3a, where the condition $\dot{q} \approx 2\sqrt{\epsilon}$ should hold, the tuning occurs for a harmonic satisfying $2m_q\sqrt{\epsilon} + m_1 J_{10} \approx 0$. This implies a ratio $|m_1|/|m_q| = \mathcal{O}(1/\sqrt{\epsilon})$. In Arnold's model, such a harmonics will be generated for the first time at the normalization order $s_0 = |m_1| + |m_q| = \mathcal{O}(1/\sqrt{\epsilon})$. Then, it turns out that there is an optimal normalization order beyond which the critical harmonic can no longer be removed from the Hamiltonian. Usual normal form estimates (see Sect. 4) lead to $s_{opt} = \mathcal{O}(1/\mu^b)$, for a positive exponent b. The size of the harmonic at optimal order will be $\mathcal{O}(\exp(1/\mu^b))$, i.e., *exponentially small* in $1/\mu$. This, yields, in general, an exponentially small drift velocity in action space.

An important remark regarding the precise estimates on the speed of Arnold diffusion is that the latter depend crucially on whether a system is *a priori stable* or *a priori unstable* (see also Sect. 3 below). This distinction has been emphasized in a central paper on the subject by [11] (hereafter CG). That paper provides a rigorous proof of the occurrence of Arnold diffusion in a priori unstable systems and also along the simple resonances of a priori stable systems. It also discusses lower bounds on the times necessary for making $\mathcal{O}(1)$ excursions in action space. These bounds are estimated as exponentially small in $1/\mu^2$.[1]

2.3 Geometric Approach

The arguments exposed so far justify local variations in the values of the actions J_1 and J_2, but provide no theory for the long ($\mathcal{O}(1)$) excursions of the trajectories in the action space. Demonstration that such excursions are possible requires, instead, the use of some geometric method. A standard method relies on the existence of orbits shadowing the heteroclinic intersections between the stable and unstable invariant manifolds emanating from the family of hyperbolic tori lying in the phase space of the system under study.

In Arnold's example, these are the tori $\tau(J_1, J_2)$ defined in Eq. (3), which are quite distinct from the 3-dimensional KAM tori referred to Sect. 2.1. In particular, along the tori $\tau(J_1, J_2)$ we always have the invariance $q(t) = p(t) = 0$, corresponding to the hyperbolic fixed point of the pendulum. However, contrary to what we saw in the previous subsection, in the geometric method we seek to characterize the motions in the neighborhood of a hyperbolic torus $\tau(J_1, J_2)$ via the study of the invariant *asymptotic manifolds* emanating from the torus.

Consider first the case $\mu = 0$. We define the stable and unstable manifolds, $\mathcal{W}^U_{(0,0)}$, $\mathcal{W}^S_{(0,0)}$ of the unstable fixed point of the pendulum as the set of all initial condi-

[1] Despite the appearances, the paper by CG contains several parts accessible to physicists and astrodynamicists. As an exercise, readers are invited to study the analogy between several rigorous definitions given in CG and the corresponding heuristic definitions given in [13], which is addressed to physicists. For example, pendulum motions close to the upper and lower branches of the pendulum separatrix correspond to the 'separatrix swings' in CG, the region B where the jumps occur is called 'origin of the separatrix', the phase sequences $\phi_{1,i}$, $i = 1, 2, \ldots$ of the whisker map are called 'phase shifts' (CG Sect. 4, etc.).

tions (q_0, p_0) whose time evolution leads to orbits $(q(t; q_0, p_0), p(t; q_0, p_0)$ tending asymptotically to the unstable point $(0, 0)$ as $t \to -\infty$ (for the unstable manifold) or $t \to \infty$ (for the stable manifold):

$$\mathcal{W}^U_{(0,0)} = \left\{(q_0, p_0) \in \mathbb{T} \times \mathbb{R} : \lim_{t\to-\infty} (q(t; q_0, p_0), p(t; q_0, p_0)) = (0, 0)\right\} \quad (15)$$
$$\mathcal{W}^S_{(0,0)} = \left\{(q_0, p_0) \in \mathbb{T} \times \mathbb{R} : \lim_{t\to\infty} (q(t; q_0, p_0), p(t; q_0, p_0)) = (0, 0)\right\} \ .$$

For $\mu = 0$ the sets $\mathcal{W}^U_{(0,0)}, \mathcal{W}^S_{(0,0)}$ coincide, as they both correspond to the pendulum separatrix. Consider, now, the following set of initial conditions of the full problem:

$$\begin{aligned} &\mathcal{Q}_0 : J_1(0) = J_{10}, \quad J_2(0) = J_{20}, \quad \phi_1(0) = \phi_{10}, \quad \phi_2(0) = \phi_{20} \\ &(q(0) = q_0, p(0) = p_0) \in \mathcal{W}^S_{(0,0)} \ . \end{aligned} \quad (16)$$

Since $\mu = 0$ the variables (q, p) evolve independently from the variables (ϕ, J). Since $(q_0, p_0) \in \mathcal{W}^S_{(0,0)}$, $(q(t), p(t))$ will tend to $(0, 0)$ as $t \to \infty$, while (ϕ, J) will have an identical evolution as in the torus $\tau(J_1, J_2)$. Hence, the trajectory tends to the torus $\tau(J_{10}, J_{20})$ as $t \to \infty$. We then define the stable and unstable manifolds of a torus $\tau(J_1, J_2)$ as:

$$\mathcal{W}^U_{\tau(J_1,J_2)} = \left\{\mathcal{Q}_0 \in \mathbb{T}^3 \times \mathbb{R}^3 : \lim_{t\to-\infty} \text{dist}\,(\mathcal{Q}(t; \mathcal{Q}_0), \tau(J_1, J_2)) = 0\right\} \quad (17)$$
$$\mathcal{W}^S_{\tau(J_1,J_2)} = \left\{\mathcal{Q}_0 \in \mathbb{T}^3 \times \mathbb{R}^3 : \lim_{t\to\infty} \text{dist}\,(\mathcal{Q}(t; \mathcal{Q}_0), \tau(J_1, J_2)) = 0\right\}$$

where $\mathcal{Q}(t; \mathcal{Q}_0) \in \mathbb{T} \times \mathbb{R}^3$ denotes the trajectory (in all six variables) corresponding to the initial condition $\mathcal{Q}_0$.

We saw that the invariant tori $\tau(J_1, J_2)$ (with $q = p = 0$) continue to exist when $\mu \neq 0$. Is it, however, possible to find initial conditions $\mathcal{Q}_0$ satisfying the definition of the stable and unstable manifolds $\mathcal{W}^S_{\tau(J_1,J_2)}, \mathcal{W}^U_{\tau(J_1,J_2)}$ when $\mu \neq 0$? The answer to this question is affirmative. In fact, a local normal form around the torus $\tau(J_1, J_2)$ allows to give in parametric form initial conditions in the neighborhood of the torus which satisfy the manifold definition. Then, propagating these local initial conditions backwards of forwards in time, respectively, we can unfold the whole set of initial conditions belonging to the manifolds $\mathcal{W}^S_{\tau(J_1,J_2)}, \mathcal{W}^U_{\tau(J_1,J_2)}$ in the perturbed case as well. However, as argued by Arnold ([2]; see also [11]), the manifolds emanating from different tori in the perturbed system $\mu \neq 0$ have a property not holding when $\mu = 0$, namely, manifolds of tori corresponding to the same energy but being sufficiently close to each other can *intersect heteroclinically*, i.e. the unstable manifold of one torus can interest with the stable manifold of a nearby torus and vice versa. Figure 5 shows schematically what happens with the manifolds of the tori $\tau(J_1, J_2)$ in Arnold's model (2): Consider a fixed value of the energy E. On one such torus we have $q = p = 0$, thus $E = J_1^2/2 + J_2$. For every initial condition with $J_1 = J_{1,0}$ we can

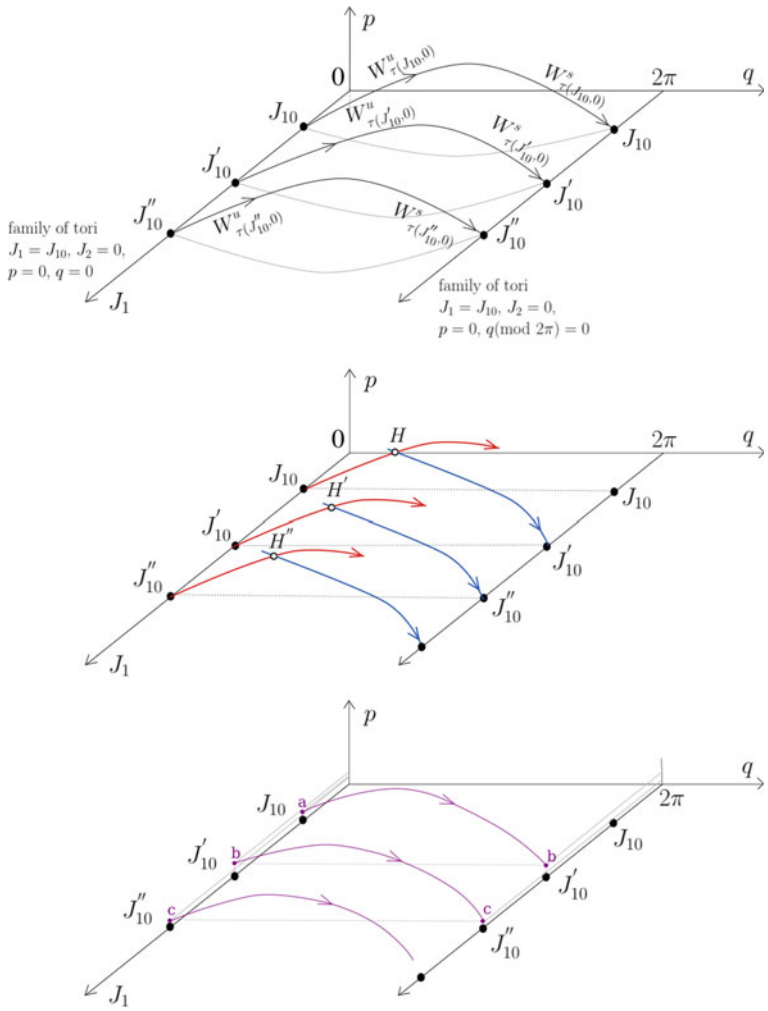


Fig. 5 Schematic representation of Arnold's mechanism: *Top:* When $\mu = 0$, the 'whiskers' (stable and unstable manifolds) of three nearby hyperbolic 2D tori labeled by the actions J_{10}, J'_{10} and J''_{10} are joined smoothly as pendulum separatrices. *Middle:* For $\mu \neq 0$, the unstable manifolds (red) of one torus intersect heteroclinically with the stable manifolds (blue) of a nearby torus. This establishes a 'chain' of heteroclinic connections. *Bottom:* There is a true orbit (purple) 'shadowing' the above chain, that is, undergoing Arnold diffusion

specify $J_2 = E - J_1^2/2$, and thus define the torus $\tau(J_1 = J_{10}, J_2 = E - J_1^2/2$. In reality, since J_2 is a dummy action variable measuring the change of energy in the non-autonomous system (1), which is equivalent to the system (2), only the action J_{10} truly labels different tori. Hence, for different values of J_{10} we obtain a family of tori, denoted by $\tau(J_{10}, 0)$, for different values of the constant J_{10}. The top panel of Fig. 5 shows three such tori, $\tau(J_{10}, 0), \tau(J'_{10}, 0), \tau(J''_{10}, 0)$, corresponding to three points on the axis J_1 of the figure. In reality, the tori are not points, but they are parameterized by the angles ϕ_1, ϕ_2 given by all possible trajectories $\phi_1(t) = \phi_{10} + J_{10}t$, $\phi_2 = t$. These angular variables are not included in the schematic Fig. 5.

Now, from every torus $\tau(J_{10}, 0)$ emanate the stable and unstable manifolds $\mathcal{W}^S_{\tau(J_{10},0)}$, $\mathcal{W}^U_{\tau(J_{10},0)}$. In the case $\mu = 0$, we saw that these manifolds join each other smoothly, as they actually coincide with the pendulum separatrix. Hence, as shown in the top panel of Fig. 5, the manifolds of different tori cannot intersect, i.e., $\mathcal{W}^U_{\tau(J_{10},0)}$ cannot intersect with $\mathcal{W}^S_{\tau(J'_{10},0)}$, $\mathcal{W}^U_{\tau(J'_{10},0)}$ cannot intersect with $\mathcal{W}^S_{\tau(J''_{10},0)}$, etc., no matter how close the tori $\tau(J_{10}, 0)$, $\tau(J'_{10}, 0)$, $\tau(J''_{10}, 0)$ are one to the other. However, this changes when $\mu \neq 0$, and it can be demonstrated that if $\tau(J_{10}, 0)$ is taken sufficiently close to $\tau(J'_{10}, 0)$, the manifolds $\mathcal{W}^U_{\tau(J_{10},0)}$ and $\mathcal{W}^S_{\tau(J'_{10},0)}$ can intersect. The middle panel of Fig. 5 shows such an intersection, at the point H, called a heteroclinic point. The sequence of the heteroclinic points H, H', H" of the middle panel of Fig. 5 will be called a 'heteroclinic chain'. The sequence of tori whose manifolds yield the points H, H', H" are known with various names, namely, the Arnold chain of 'whiskered tori' (the manifolds are the 'whiskers'), or the 'diffusion path' (see [11]).

Consider, finally, the past and future trajectories with initial conditions corresponding to the points H, H', H", etc. The trajectory from H belongs to both the invariant manifolds $\mathcal{W}^U_{\tau(J_{10},0)}$ and $\mathcal{W}^S_{\tau(J'_{10},0)}$. Thus, in the limit $t \to \infty$ the trajectory tends to the torus $\tau(J_{10}, 0)$, while, in the limit $t \to \infty$ the trajectory tends to the torus $\tau(J'_{10}, 0)$. This implies that this particular trajectory undergoes no large excursion in the action space, since its past and future is confined between two nearby asymptotic limits. Similarly, the past and future from the heteroclinic point H' connect the tori $\tau(J'_{10}, 0)$ with $\tau(J''_{10}, 0)$, those from the heteroclinic point H" connect the tori $\tau(J''_{10}, 0)$ with $\tau(J'''_{10}, 0)$, etc., but the corresponding trajectories make only bounded excursions in the action space. However, employing a so-called *shadowing lemma*, it is possible to demonstrate that there is one continuous in time trajectory of the system which remains piece-wise close (i.e. 'shadows') any one of the distinct heteroclinic trajectories from the points H, H', H'',...Such a trajectory is shown schematically in the last panel of Fig. 5. It is precisely this trajectory which materializes the 'Arnold's mechanism' referred to in the introduction. Extending the heteroclinic chain H, H', H'', ..., $H^{(n)}$, ... to include more heteroclinic points, one can find a trajectory connecting the neighborhoods of the initial torus $\tau(J_{10}, 0)$ and another torus $\tau(J^{(n)}_{10}, 0)$ located at arbitrarily large distance from $\tau(J_{10}, 0)$ (possibly limited only by the requirement of the two tori being isoenergetic).

Does the 'Arnold mechanism' interpret the long-term evolution of the numerical trajectory used in our example in the previous subsection? Figure 6 suggests this to be so, provided that the trajectory is integrated for times much longer than those referred

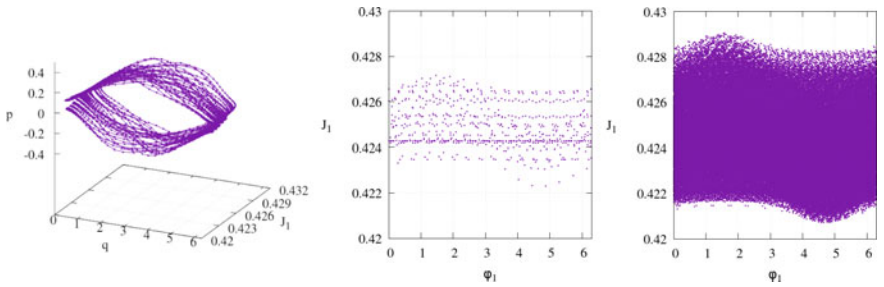


Fig. 6 *Left:* Real (non-schematic) orbit shadowing the intersecting manifolds of nearby tori in Arnold's model, obtained by plotting in the $(p, q) \times J_1$ space the same orbit as in Fig. 1, for the integration time $t = 700$. *Center* and *Right:* the projection of the orbit on the (ϕ_1, J_1) plane at two different integration times, $t = 700$ and $t = 150000$

to in the previous subsection. The left panel shows how the trajectory produced by integration of the complete model (2), and with the same initial conditions as in Fig. 1 shadows the whiskers of nearby tori $\tau(J_1, J_2)$. The middle and right panels show the projection of the trajectory in the plane (ϕ_1, J_1). Clearly, the trajectory remains piece-wise close to various rotational tori (corresponding to different values of J_1), however, as the integration time extends from $t = 700$ to $t = 1.5 \times 10^5$ the excursion in J_1 extends from a total size $\sim 10^{-2}$ to nearly $\sim 10^{-1}$. Note that as the trajectory reaches domains further and further away from this particularly selected initial condition, the drift in action space actually gets slower (see last paragraph of Sect. 2.2).

As a final remark, the above geometric picture of intersecting manifolds can be extended, from the chain of nearby tori, to include the whole invariant set $\mathcal{T}$ of the tori $\tau(J_1, J_2)$. This is a four-dimensional subset of $\mathbb{R}^2 \times \mathbb{T}^2$, which is *normally hyperbolic* (see [21] for definitions). Normal hyperbolicity implies the existence of a stable and unstable manifold for the whole invariant set $\mathcal{T}$. Since, in Arnold's example, $\mathcal{T}$ is just foliated by the tori $\tau(J_1, J_2)$, the manifolds $\mathcal{W}^U_{\mathcal{T}}$, $\mathcal{W}^S_{\mathcal{T}}$ are just the union of the unstable and stable manifolds of all the tori. Homoclinic orbits can then be described by a 'scattering map' indicating how a point on $\mathcal{T}$ is mapped asymptotically in time to another point on $\mathcal{T}$ via a doubly-asymptotic orbit.

3 A Priori Stable Systems—Nekhoroshev Theory

Consider the following Hamiltonian in action-angle variables, which, according to Poincaré [56], represents the *"fundamental problem of dynamics"*:

$$H(\phi, I) = H_0(I) + \epsilon H_1(\phi, I) \tag{18}$$

with $\phi \in \mathbb{T}^n$, $I \in \mathbb{R}^n$.

For $\epsilon = 0$ the system is integrable $H = H_0(I)$ and the phase space is foliated by invariant tori labeled by the constant actions I. On each torus the angles evolve linearly with the frequencies $\omega(I) = \nabla_I H_0(I)$. Periodic orbits, or, in general, tori of dimension $n' < n$ correspond to values of the actions I for which the frequencies $\omega(I)$ satisfy $n - n'$ commensurability conditions. However, all these low-dimensional objects are neutral in stability, and there are no separatrices or any other type of asymptotic manifolds ('whiskers') associated to them. In other words, there is no in-built hyperbolicity in the Hamiltonian $H_0(I)$. Hence, invariant objects of (partially) hyperbolic character can only be born by setting $\epsilon \neq 0$. Such systems were thus called (by CG) 'a priori stable'.

The lack of invariant phase space objects with inherent hyperbolicity generates several challenging new questions regarding Arnold diffusion. We now summarize some of these questions as well as known results related to Arnold diffusion in a priori stable systems.

3.1 Nekhoroshev Theory and Exponential Stability

Whatever the mechanism possible to cause Arnold diffusion in an a priori stable system, the speed of the drift in action space in such a system is bounded before all by the *Nekhoroshev theorem* [3, 4, 48, 55, 57]:

Nekhoroshev theorem: *Assume a Hamiltonian of the form* (18) *with* $\epsilon > 0$, *with* H *analytic in a complex extension* $\mathcal{D}$ *of the set* $D \times \mathbb{T}^n$, *where* $D \subset \mathbb{R}^n$ *is open, and* H_1 *bounded. Assume that* H_0 *satisfies suitable steepness conditions. Then, there are positive constants* a, b, ϵ_0 *such that, for* $\epsilon < \epsilon_0$ *and for all initial conditions in* $\mathcal{D}$, *under the flow of the Hamiltonian* H *we have:*

$$|J(t) - J(0)| < \epsilon^a \quad \text{for all times } t < T_N \text{ with } T_N = \mathcal{O}\left(\frac{\epsilon_0}{\epsilon} \exp((\epsilon_0/\epsilon)^b)\right) \tag{19}$$

We refer to T_N as the 'Nekhoroshev time'. A detailed discussion of the meaning and importance of 'steepness' in the above theorem is made in [12, 39, 60]. We briefly refer to steepness in Sect. 3.2 below.

Demonstration of the Nekhoroshev Theorem (see [34] for a tutorial) requires combining an *analytical* with a *geometric* part. The analytical part deals with the local construction of a 'Nekhoroshev normal form', whose remainder at the optimal normalization order turns to be exponentially small. On the other hand, the geometric part deals with the construction of a set of subdomains $D_1, D_2, \ldots \subset \mathcal{D}$ defined so that: (i) a different local normal form with exponentially small remainder can be constructed in each domain, and (ii) the union of all domains provides a covering of $\mathcal{D}$. The structure of *resonant manifolds* (see below), depending on the form of the integrable part $H_0(I)$ of the Hamiltonian, as well as the size of the analyticity domain around each manifold, determined by the form of $H_1(\phi, I)$, are crucial

factors in the appropriate definition of the domains D_i. In particular, the domains D_i must have size depending algebraically on ϵ, i.e. $\mathrm{diam}(D_i) = \mathcal{O}(\epsilon^{a_i})$, $a_i > 0$. One then demonstrates that this dependence allows to obtain a covering of $\mathcal{D}$ by combining many such domains when ϵ is arbitrarily small (see [54] for a heuristic argument). Now, the size of the optimal remainder of each local normal form scales as $||R|| = O\left(-\exp((\epsilon_{0,i}/\epsilon)^{b_i})\right)$, for some positive constant $\epsilon_{0,i}$ and positive exponent b_i. Choosing the worst possible combination a_i, b_i and $\epsilon_{i,0}$ from those holding in each domain allows to arrive at the global bound (19). In practice, locally we can obtain better bounds using the local parameters a_i, b_i, $\epsilon_{0,i}$. It turns out that the exponents a, b depend on (i) the number of degrees of freedom n, (ii) the so-called steepness indices holding within the domain (see [39] for definitions) and, finally, (iii) the *multiplicity* of the local resonance considered (see below).

It is noteworthy that, while in the proof of the theorem the analytical part plays a minimal role, the *actual construction* of the Nekhoroshev normal form in any explicit application implies reaching a very high order of normalization, involving typically millions of operations that can only be carried out with the aid of a computer-algebraic program. Starting from the sixties [15, 16, 35, 36], such programs dealt first with the simpler case of systems with elliptic equilibria, such as the celebrated Hénon-Heiles system [41]. In such systems, exponential estimates can be obtained without the need of a geometric construction as the one of the Nekhoroshev theorem. Well known applications in Celestial Mechanics have been given, referring, for example, to the long term stability of the Trojan asteroids of Jupiter [9, 23, 33, 47], the spin-orbit problem [59], and the J_2 problem of satellite motions [18, 61]. On the other hand, computing the optimal Nekhoroshev normal form in a generic Hamiltonian of the form (18) has been possible so far only in simple models with $n = 3$ degrees of freedom [14, 24, 26, 40]. Such computations allow for a direct comparison between 'semi-analytical' (i.e. by the remainder of the Nekhoroshev normal form) and numerical results on the speed of Arnold diffusion, as well as on the adiabatic evolution of the action variables in a priori stable systems. Most notable among the numerical experiments are those carried over the years by the group of Froeschlé, Guzzo and Lega [30, 37, 39, 44], which have given clear evidence of the occurrence of Arnold diffusion in a priori stable systems. A comparison of the exponents a, b found by the Nekhoroshev normal form construction and by the numerical experiments has shown a very good agreement. This has extended also to estimates on the coefficient of Arnold diffusion as well as to the modeling of the jumps carried by the adiabatic action variables along the heteroclinic transitions taking place in single resonance domains. In the sequel we give a summary of the above results with the help (as in the previous section) of a simple example of a priori stable system with $n = 3$ degrees of freedom.

3.2 A Simple Example

Consider the 3DOF Hamiltonian in action-angle variables:

$$H = H_0 + \epsilon H_1 = \frac{I_1^2}{2} - \frac{I_2^2}{2} + \frac{I_2^3}{3\pi} + 2\pi I_3 + \frac{\epsilon}{4 + \cos\phi_1 + \cos\phi_2 + \cos\phi_3} \ . \quad (20)$$

The Hamiltonian (20) has been used in [40] in the study of the evolution of the adiabatic action variables. An analogous 4D symplectic mapping was used in [39] for the study of the effects of steepness on the stability of the orbits.

The flow corresponding to the integrable part of (20)

$$H_0 = \frac{I_1^2}{2} - \frac{I_2^2}{2} + \frac{I_2^3}{3\pi} + 2\pi I_3 \ . \quad (21)$$

is given by $\dot{I}_i = 0$, $i = 1, 2, 3$ and $\dot{\phi}_1 = \omega_{0,1} = I_1$, $\dot{\phi}_2 = \omega_{0,2} = -I_2 + \frac{1}{\pi} I_2^2$, $\dot{\phi}_3 = \omega_{0,3} = 2\pi$. Thus, all trajectories lie on invariant tori labeled by the actions I_i or the corresponding frequencies $\omega_{0,i}$.

Let $k \equiv (k_1, k_2, k_3) \in \mathbb{Z}^3$. We call *resonant manifold* $\mathcal{RM}(k_1, k_2, k_3)$ associated to the Hamiltonian H_0 the two-dimensional manifold

$$\mathcal{RM}(k_1, k_2, k_3) := \Big\{ (I_1, I_2, I_3) \in \mathbb{R}^3 : \quad (22)$$
$$k \cdot \omega_0(I) = k_1 I_1 + k_2(-I_2 + \frac{1}{\pi} I_2^2) + k_3 2\pi = 0 \Big\} \ .$$

We call *energy manifold* $\mathcal{E}(E)$ the two-dimensional manifold

$$\mathcal{E}(E) := \left\{ (I_1, I_2, I_3) \in \mathbb{R}^3 : \ H_0(I) = \frac{1}{2}(I_1^2 - I_2^2) + \frac{I_2^3}{3\pi} + 2\pi I_3 = E \right\} \ . \quad (23)$$

Figure 7a shows a part of the energy manifold $\mathcal{E}(E)$ for $E = 1$ as well as parts of the two resonant manifolds $\mathcal{RM}(1, 1, 0)$ and $\mathcal{RM}(4, -1, -1)$. The set of all curves formed by the intersection of all resonant manifolds $\mathcal{RM}(k)$, $k \in \mathbb{Z}^3$, $|k| \neq 0$ with the energy manifold $\mathcal{E}(E)$ is called the *Arnold web* (or 'web of resonances'). In our example, the definition of the resonant manifolds via Eq. (22) does not depend on I_3. Thus all resonant manifolds intersect normally the plane (I_1, I_2) at curves given by Eq. (22). Figure 7b shows some of these resonant curves marked with the corresponding integers (k_1, k_2, k_3).

The set of the resonant curves defined by all possible $(k_1, k_2, k_3) \in \mathbb{Z}^3$, $|k| \neq 0$ is dense in the square $S(I_1, I_2)$ depicted in Fig. 7b: for any open, small whatsoever, neighborhood $S_i \subset S(I_1, I_2)$ there exist integers (k_1, k_2, k_3) such that the corresponding resonant curve crosses S_i. However, not all these resonances are equally important for dynamics. This is evidenced by computing a *stability map* in the same square

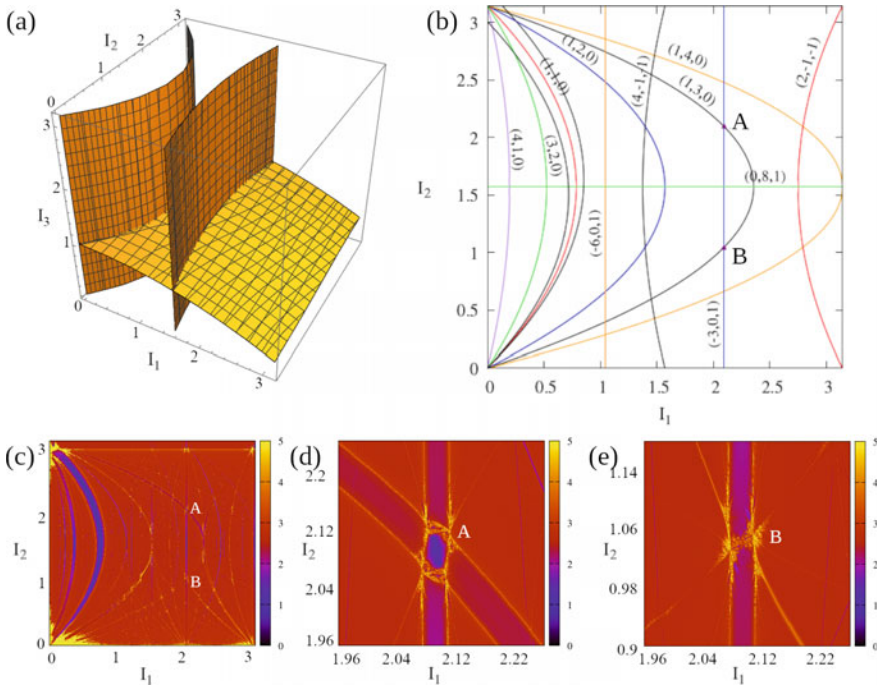


Fig. 7 **a** Part of the energy manifold $\mathcal{E}(E)$ in the model (21) for $E = 1$ (yellow), intersected by parts of the resonant manifolds $\mathcal{RM}(1, 1, 0)$ and $\mathcal{RM}(4, -1, -1)$ (orange). **b** Projection of the Arnold web of resonances on the (I_1, I_2) plane. For the resonance $(1, 1, 0)$ the corresponding separatrix borders are also displayed as computed theoretically for $\epsilon = 0.05$ (see text). **c** FLI stability map for the Hamiltonian (20) with $\epsilon = 0.05$. The web of resonances is visualized through the detection of weakly chaotic orbits at the borders of each resonance. **d** and **e** Details of figure (**c**) around the resonance junctions A and B, respectively, where the resonant manifolds $\mathcal{RM}(1, 3, 0)$ and $\mathcal{RM}(3, 0, -1)$ intersect

$S(I_1, I_2)$ via the use of a chaotic indicator. Figure 7c shows the stability map computed by the Fast Lyapunov Indicator (FLI, [28]) in a grid of initial conditions for (I_1, I_2), setting initially $I_3 = \phi_1 = \phi_2 = \phi_3 = 0$, and for an integration time $t = 1000$. We immediately note that the FLI map in Fig. 7c is able to depict the structure of the Arnold web in great detail. This fact, first found in [29] has played a crucial role in the numerical study of Arnold diffusion in a priori stable systems.

In Fig. 7c we see that the most prominent structures are related to low order resonances ($|k| = |k_1| + |k_2| + |k_3|$ small). Also, we notice that, for some resonances (e.g. $(1, 1, 0)$), the FLI map shows a double set of curves going nearly parallel one to the other along the resonance, with a blue zone between the curves. Other resonances, instead, are identified by a single line (yellow). This distinction depends on the sign of the Fourier coefficient of the corresponding resonant harmonics in the function H_1 of Eq. (20). We have:

$$\frac{1}{4+\cos\phi_1+\cos\phi_2+\cos\phi_3}=\sum_{k_1,k_2,k_3=-\infty}^{\infty} h_{k_1,k_2,k_3}\cos(k_1\phi_1+k_2\cos\phi_2+k_3\phi_3)$$

where h_{k_1,k_2,k_3} can be easily computed expanding the denominator in Taylor series and using the trigonometric reduction formulas. Consider a toy Hamiltonian in which only one harmonic is isolated:

$$H_{res}=\frac{I_1^2}{2}-\frac{I_2^2}{2}+\frac{I_2^3}{3\pi}+2\pi I_3+\epsilon h_{k_1,k_2,k_3}\cos(k_1\phi_1+k_2\cos\phi_2+k_3\phi_3)\ . \quad (24)$$

Such a model will be obtained by just performing one step of perturbation theory eliminating from the Hamiltonian (20) all other harmonics except for the resonant one (see Sect. 4). Now, the Hamiltonian (24) is integrable. To show this, assume (without loss of generality) $k_1\neq 0$. Consider two linearly independent integer vectors $m,n\in\mathbb{Z}^3$ such that $m\cdot k=n\cdot k=0$ (for example $m=(k_2,-k_1,0)$, $n=(k_3,0,-k_1)$). Consider the canonical transformation $(\phi_1,\phi_2,\phi_3,I_1,I_2,I_3)\to(\phi_R,\phi_{F1},\phi_{F2},I_R,I_{F1},I_{F2})$ defined by

$$\phi_R=k\cdot\phi,\ \phi_{F1}=m\cdot\phi,\ \phi_{F2}=n\cdot\phi\ ,$$

as well as the inverse of the equations

$$\begin{aligned} I_1&=k_1I_R+m_1I_{F1}+n_1I_{F2},\\ I_2&=k_2I_R+m_2I_{F1}+n_2I_{F2},\\ I_3&=k_3I_R+m_3I_{F1}+n_3I_{F2}\ . \end{aligned} \quad (25)$$

Substituting these expressions into (24) we arrive at:

$$H_{res}=H_0(I_R,I_{F1},I_{F2})+\epsilon h_{k_1,k_2,k_3}\cos(\phi_R)\ . \quad (26)$$

Since the angles ϕ_{F1},ϕ_{F2} are ignorable, the above model has two integrals of motion I_{F1},I_{F2} besides the energy. We are interested in studying the behavior of the model H_{res} in a neighborhood around values (I_{1*},I_{2*},I_{3*}) which satisfy the resonance exactly. Setting $I_i=I_{i*}+J_i$, $i=1,2,3$ and substituting into (24) we arrive at:

$$\begin{aligned} H_0(J)=H_0(I_*)+\nabla_I H_0(I_*)\cdot J+\frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{\partial^2 H_0(I_*)}{\partial I_i\partial I_j}J_iJ_j \quad (27)\\ +\frac{1}{6}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{l=1}^{3}\frac{\partial^3 H_0(I_*)}{\partial I_i\partial I_j\partial I_l}J_iJ_jJ_l+\cdots\ . \end{aligned}$$

The constant term $H_0(I_*)$ can be omitted. The term $\nabla_I H_0(I_*)\cdot J$ has the form

$$\nabla_I H_0(I_*) \cdot J = (k \cdot \omega_*) J_R + (m \cdot \omega_*) J_{F1} + (n \cdot \omega_*) J_{F2} = (m \cdot \omega_*) J_{F1} + (n \cdot \omega_*) J_{F2} \ .$$

where ω_* denotes the vector of the resonant frequencies $\omega_{i*} = \omega_i(I_*)$, and the variables J_R, J_{F1}, J_{F2} are defined as $J_R = I_R - I_{R*}$, $J_{F1} = I_{F1} - I_{F1*}$, $J_{F2} = I_{F2} - I_{F2*}$ with

$$\begin{pmatrix} I_{R*} \\ I_{F1*} \\ I_{F2*} \end{pmatrix} = \begin{pmatrix} k_1 \ m_1 \ n_1 \\ k_2 \ m_2 \ n_2 \\ k_3 \ m_3 \ n_3 \end{pmatrix}^{-1} \begin{pmatrix} I_{1*} \\ I_{2*} \\ I_{3*} \end{pmatrix}$$

The frequencies ω_* satisfy $k \cdot \omega_* = 0$, hence the transformed Hamiltonian contains linear terms only for the 'fast' action variables J_{F1}, J_{F2}. Instead, the resonant action J_R appears in the Hamiltonian only in quadratic terms (or of higher degree) in the actions. Setting the integrals as $J_{F1} = 0$, $J_{F2} = 0$ implies the relations $I_{F1} = I_{F1*}$, $I_{F2} = I_{F2*}$, that is:

$$\begin{aligned} I_1 &= k_1 I_R + m_1 I_{F1*} + n_1 I_{F2*} \\ I_2 &= k_2 I_R + m_2 I_{F1*} + n_2 I_{F2*} \\ I_3 &= k_3 I_R + m_3 I_{F1*} + n_3 I_{F2*} \end{aligned} \tag{28}$$

Thus, the motion in all three action variables under the flow of the model Hamiltonian (24) is determined by the only evolving action, namely I_R, and it is confined along a line $\mathcal{L}(I_*)$ in the space (I_1, I_2, I_3) defined parametrically by Eq. (28). The projection of the line $\mathcal{L}(I_*)$ on the plane (I_1, I_2) is given by

$$I_2 = \frac{1}{k_1} \left(-k_2 I_I + (k_2 m_1 - k_1 m_2) I_{F1*} + (k_2 n_1 - k_1 n_2) I_{F2*} \right) \tag{29}$$

Also, the only non-ignorable angle in the model Hamiltonian of the resonance is $\phi_R \in \mathbb{T}$. The set $\mathcal{P}_F(I_*) = \mathcal{L}(I_*) \times \mathbb{T}$ is called *plane of fast drift*. On this plane the motion is described by a pendulum-like Hamiltonian in the local variables (ϕ_R, J_R). The Eqs. (28) imply $J_R = (k \cdot J)/(k \cdot k)$. Then, the quadratic term in the actions in (27) takes the form:

$$\frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial^2 H_0(I_*)}{\partial I_i \partial I_j} J_i J_j = \frac{1}{2} \beta(I_*) J_R^2 \ \text{ with } \ \beta(I_*) = \frac{1}{k^2} (M(I_*) k) \cdot k \tag{30}$$

where $M(I_*)$ is the 3×3 Hessian of the Hamiltonian H_0 calculated at the point I_*

$$M_{ij}(I_*) = \left(\frac{\partial^2 H_0}{\partial I_i \partial I_j} \right)_{I = I_*}$$

Similarly, the cubic term in the actions takes the form $(1/3)\gamma(I_*) J_R^3$ with

$$\gamma(I_*) = \frac{1}{2|k|^{3/2}} \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{l=1}^{3} \left(\frac{\partial^3 H_0(I_*)}{\partial I_i \partial I_j \partial I_l} \right)_{I=I_*} k_i k_j k_l \tag{31}$$

Hence, apart from constants we have

$$H_{res} = \frac{1}{2}\beta(I_*) J_R^2 + \frac{1}{3}\gamma(I_*) J_R^3 + \epsilon h_k \cos(\phi_R) \tag{32}$$

where, in the model (21) we get:

$$\beta(I_*) = k_1^2 + k_2^2 \left(\frac{2I_{2*}}{\pi} - 1 \right), \quad \gamma(I_*) = \frac{k_2^3}{\pi} \tag{33}$$

Except for the case $k_1 = k_2$ and $I_{2*} \to 0$, the coefficient $\beta(I_*)$ is in general a $\mathcal{O}(1)$ quantity. Then, taking J_R in a domain of size $\mathcal{O}(\epsilon^{1/2})$, the term $\frac{1}{2}\beta(I_*) J_R^2$ is more important than the term $\frac{1}{3}\gamma(I_*) J_R^3$ in H_{res}. This means that H_{res} (ignoring cubic terms) becomes a pendulum Hamiltonian with separatrices extending in a domain $J_{R,min} \leq J_R \leq J_{R,max}$ estimated by:

$$J_{R,min} \simeq -2 \left(\frac{\epsilon}{|\beta(I_*)|} \right)^{1/2}, \quad J_{R,max} \simeq 2 \left(\frac{\epsilon}{|\beta(I_*)|} \right)^{1/2} . \tag{34}$$

In reality, the motion very close to the separatrix will be weakly chaotic, due to the fact that, as discussed below, the remaining resonances can be eliminated only up to an exponentially small remainder, and hence there is some degree of chaos due to the interaction of these resonances with the principal one (k_1, k_2, k_3). The motion along the separatrix-like thin chaotic layer of the resonance can be projected also on the plane (I_1, I_2). The projection is constrained in a segment along the line $\mathcal{L}(I_*)$, which represents the intersection of the plane of fast drift with the plane (I_1, I_2). In particular, the motion along the separatrix layer projects to a linear segment given by Eq. (28), setting $I_R = I_{R*} + J_R$, and varying J_R in the limits $J_{R,min} \leq J_R \leq J_{R,max}$.

We are now able to understand the structure of the FLI map shown in Fig. 7c. Let I_* be one point along the resonance (k_1, k_2, k_3). Since in the computation of the FLI we have set the initial conditions $\phi_i = 0, i = 1, 2, 3$, the FLI map intersects the plane of fast drift crossing the point I_* at the value $\phi_R = 0$. Whenever the coefficients $\beta(I_*)$ and h_{k_1,k_2,k_3} have the same sign, the point ϕ_R represents the unstable equilibrium point of the Hamiltonian H_{res}. One has $J_R = 0$ there, thus, by Eq. (28) we get a unique point on the FLI map, given by $I_1 = I_{1*}$, $I_2 = I_{2*}$. On the contrary, when $\beta(I_*)$ and h_{k_1,k_2,k_3} have opposite signs, the point ϕ_R corresponds to the stable equilibrium point of the Hamiltonian H_{res}. Then, the line $\phi_R = 0$ on the fast drift plane crosses the separatrix layer approximately at the values $J_R = J_{R,min}$ and $J_R = J_{R,max}$. Thus, by Eq. (28) we get two point on the FLI map, given by $I_1 = I_{1*} + k_1 J_{R,min}$, $I_2 = I_{2*} + k_2 J_{R,min}$, and $I_1 = I_{1*} + k_1 J_{R,max}$, $I_2 = I_{2*} + k_2 J_{R,max}$. Joining the two families of points representing the separatrices for different points I_* along the same resonance yields

two curves on the plane (I_1, I_2) which follow nearly parallelly the curve of the resonance, having between themselves a $\mathcal{O}(\epsilon^{1/2})$ distance. Figure 7 shows the two curves marking the borders of the resonance (1, 1, 0), as computed by the above formulas. This fits very well the borders found by the FLI map of Fig. 7c. The blue zone between the two borders corresponds to regular orbits, which are the libration orbits of the pendulum for initial conditions inside the separatrices.

In general, fixing a certain model H_0, we have $\text{sign}[\beta(I_*)] = \text{sign}[(M(I_*)k) \cdot k]$. When the quadratic form $(M(I_*)k) \cdot k$ is positive definite, $\beta(I_*)$ has always the same sign, independently of the resonant vector k. In this case, whether the separatrices intersect with the chosen section at a single or double curve depends only on the sign of the coefficient h_k of the Fourier harmonic $\cos(k \cdot \phi)$ in H_1. On the contrary, if the Hessian matrix $M(I_*)$ is not positive definite, the sign of $\beta(I_*)$ depends on the value of I_* and on the choice of resonance, i.e., of the vector k. In the model (21), we readily find that $M(I_*)$ is positive definite in the semi-plane $I_{2*} > \pi/2$, while it is not in the semi-plane $I_{2*} < \pi/2$. In the latter one, the sign of β depends on the particular choice of resonance. For example, for the resonance $k = (1, 1, 0)$ there is no change of sign of $\beta(I_*)$ across the two semi-planes. For all other resonances $k = (1, k_2, 0)$, $k_2 > 1$, $\beta(I_*)$ changes sign, instead, at the value $I_{2*} = (\pi/2)(1 - k_2^2/k_1^2)$, a fact easily verified by carefully inspecting the FLI map of Fig. 7.

Besides graphical consequences for the FLI maps, positive-definiteness (or not) of the Hessian matrix $M(I_*)$ affects several aspects of the dynamics: an important aspect regards the dynamics around *resonance junctions*. In the case with $n = 3$ DOF, we consider points I_* for which there exist two linearly independent non-zero integer vectors $k^{(1)}, k^{(2)}$ satisfying:

$$k^{(1)} \cdot \omega(I_*) = 0, \quad k^{(2)} \cdot \omega(I_*) = 0 \ . \tag{35}$$

Such points I_* are said to belong to resonant junctions of multiplicity 2: this is a curve, in the 3D action space, where all resonant manifolds $\mathcal{RM}(\lambda_1 k^{(1)} + \lambda_2 k^{(2)})$ defined by the two linearly independent vectors $k^{(1)}, k^{(2)}$ and by $\lambda_1, \lambda_2 \in \mathbb{Z}$ intersect each other. For $n = 3$ a resonant junction can only be of multiplicity 2. For $n > 3$, instead, resonance junctions can be of multiplicity $2 \leq mult \leq n - 1)$, and the corresponding resonant junctions are manifolds of dimension $n - mult$.

Figure 7d and e show the FLI maps around the resonance junctions formed by the crossing of the resonances $(1, 3, 0)$ and $(3, 0, -1)$ at the points A and B. We immediately notice the difference in structure of the resonance crossings at these two points. Briefly, this can be understood as follows (see [25] for details): let I_* be a doubly resonant point. Define the vector $m = k^{(1)} \times k^{(2)}$ as well as the canonical transformation:

$$\begin{aligned} J_i &= k_i^{(1)} J_{R1} + k_i^{(2)} J_{R2} + m_i J_F, \quad i = 1, 2, 3 \\ \phi_{R1} &= k^{(1)} \cdot \phi, \quad \phi_{R2} = k^{(2)} \cdot \phi, \quad \phi_F = m \cdot \phi \end{aligned} \tag{36}$$

where, as before, $J_i = I_i - I_{i*}$. By Eq. (27) up to quadratic terms we now get (apart from a constant)

$$H_0 = \omega_F J_F + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}(I_*)(k_i^{(1)} J_{R1} + k_i^{(2)} J_{R2} + m_i J_F)(k_j^{(1)} J_{R1} + k_j^{(2)} J_{R2} + m_j J_F) \tag{37}$$

The frequency $\omega_F = m \cdot \omega$ yields the rate of change of the unique 'fast angle' of the problem $\phi_F = m \cdot \phi$ (conjugate to J_F). As before, we can assume computing a resonant normal form which eliminates all harmonics in the problem except $\cos((\lambda_1 k^1 + \lambda_2 k^{(2)}) \cdot \phi)$. Thus, an appropriate toy model for the double resonance is

$$H_{doubleres} = H_0(J_{F1}, J_{F2}, J_F) + \epsilon \sum_{l_1,l_2} g_{l_1,l_2} \cos(l_1\phi_{R1} + l_2\phi_{R2}) \ . \tag{38}$$

The coefficients g_{l_1,l_2} are expressed in terms of the original Fourier coefficients h_k. Now, contrary to the case of single resonance, $H_{doubleres}$ has only one ignorable angle (ϕ_F), hence, besides the energy, only the action J_F is integral of motion. Then, considering J_F as a parameter, the dynamics of $H_{doubleres}$ corresponds to a *non-integrable* system with two degrees of freedom. This is a general property of multiple resonances, for which the Nekhoroshev normal form induces a non-integrable dynamics. Availing no other restrictions than those imposed by energy conservation, the dynamics near the junction can be very chaotic (see, for example, [26, 31]). However, as discussed in [4, 57], energy conservation can still be used in many cases to constrain the orbits consistently with the Nekhoroshev theorem. As in the case of simple resonance, consider, without loss of generality, the normal form dynamics induced by the Hamiltonian Eq. (38) for (constant) $J_F = 0$. The normal form energy $E = H_{doubleres}$ is a constant of motion. Thus, the quantity $H_0(J_{R1}, J_{R2}, 0)$ can only undergo $\mathcal{O}(\epsilon)$ oscillations around the value $E = H_0(J_{R1}, J_{R2}, 0)$. We then seek for conditions on H_0 such that the manifold $E = H_0(J_{R1}, J_{R2}, 0)$ be *bounded*, i.e. that none of J_{R1}, J_{R2} can take $\mathcal{O}(1)$ values while the energy $E = H_0(J_{R1}, J_{R2}, 0)$ still remains in the interval $E - \mathcal{O}\epsilon < H_0 < E + \mathcal{O}(\epsilon)$. Subtracting an irrelevant constant, consider values of the energy $E = \mathcal{O}(\epsilon)$. We have (for $J_F = 0$):

$$\begin{aligned} E &= \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}(I_*)(k_i^{(1)} J_{R1} + k_i^{(2)} J_{R2})(k_i^{(1)} J_{R1} + k_i^{(2)} J_{R2}) \\ &= \zeta_2 = (J_{R1}, J_{R2}) Y (J_{R1}, J_{R2})^T \end{aligned} \tag{39}$$

where Y is the 2×2 matrix

$$Y = k^{(1,2)} M(I_*) (k^{(1,2)})^T$$

with

$$k^{1,2} = \begin{pmatrix} k_1^{(1)} & k_2^{(1)} & k_3^{(1)} \\ k_1^{(2)} & k_2^{(2)} & k_3^{(2)} \end{pmatrix}$$

The quadratic form (39) is positive definite when $M(I_*)$ has three non-zero eigenvalues of equal sign, or two eigenvalues of equal sign and one equal to zero. In the first case, the Hamiltonian H_0 will be called *convex*, and in the second *quasi-convex*. In general, we give the following definitions:

Convexity: The n-degrees of freedom Hamiltonian H_0 is convex at the point I_* if there is a positive constant M such that for any $x \in R^n, x \neq 0$ we have $|(M(I_*)x) \cdot x| \geq M$.

Quasi-convexity: The Hamiltonian H_0 is quasi-convex at the point I_* if $\omega(I_*) \neq 0$ and the only solution to the system $\omega(I_*) \cdot x = 0$ and $(M(I_*)x) \cdot x = 0$ is $x = 0$.

We leave to the reader as an exercise to demonstrate that when H_0 is (quasi)convex at the point I_*, the 2×2 matrix Y of Eq. (39) is positive definite (see also equation (171) in [25]). Then, the equation $\zeta_2(J_{R1}, J_{R2}) = E$ is the equation of an ellipse. For fixed value of $E = \mathcal{O}(\epsilon)$, both actions J_{R1}, J_{R2} are bounded by the fixed size (say, the semi-major axis) of the ellipse. The latter is of order $\sqrt{\epsilon}$, hence the actions J_{R1}, J_{R2} are bounded in a domain of size $\mathcal{O}(\sqrt{\epsilon})$. On the contrary, at points I_* where (quasi-)convexity is not satisfied, the matrix Y can be positive-definite or not, depending on the particular resonant vectors $k^{(1)}, k^{(2)}$. Correspondingly, the equation $\zeta_2(J_{R1}, J_{R2}) = E$ gives either an ellipse or a hyperbola. At those junctions where we have hyperbolas, the actions J_{R1}, J_{R2} are unbounded along the asymptotes of the hyperbolas.[2]

In this case, however, a bound for the actions J_{R1}, J_{R2} via the requirement $|H_0(J_{R1}, J_{R2}, 0)| < \mathcal{O}(\epsilon)$ can still be obtained using the cubic terms in the formula for H_0 (Eq. (27)). Without entering into details, we only mention that such a bound exists when the Hamiltonian H_0 satisfies the *three-jet* condition:

Three-jet: at the point I_* we have $\omega(I_*) \neq 0$ and the only solution to the system of equations

$$\omega(I_*) \cdot x = 0, \quad (M(I_*)x) \cdot x = 0, \quad \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} \left(\frac{\partial^3 H_0}{\partial I_i \partial I_j \partial I_l} \right)_{I=I_*} x_i x_j x_l = 0 \tag{40}$$

[2] For example: $H_0 = (I_1^2 - I_2^2)/2 + I_3$. Then, $\omega_1 = I_1, \omega_2 = -I_2, \omega_3 = 1$, and $(Mk) \cdot k = k_1^2 - k_2^2$ which is not positive definite. Take the point $I_* = (1, 1, 0)$ corresponding to the double resonance $k^{(1)} = (1, 1, 0)$, $k^{(2)} = (1, 0, -1)$. We obtain $J_1 = J_{R1} + J_{R2} - J_F$, $J_2 = J_{R1} + J_F$, $J_3 = -J_{R2} - J_F$, implying $H_0 = J_{R1} J_{R2} + J_{R_2}^2$. Then, the equation $E = J_{R1} J_{R2} + J_{R_2}^2 = \frac{1}{2}(J_{R1} + J_{R2})^2 - J_{R1}^2$ represents hyperbolas with the asymptotes $J_{R2} = 0$ and $J_{R2} = -2J_{R1}$. Thus, even with energy $E = 0$, the actions can move freely along the asymptotes without violating the constant energy condition.

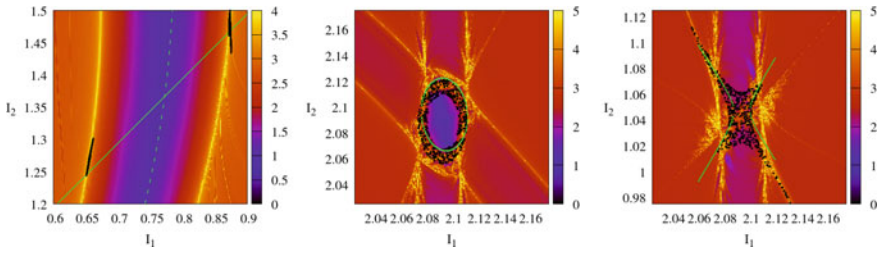


Fig. 8 **a** *Left:* Arnold diffusion along a simple resonance in the model (20) for $\epsilon = 0.1$ (see text). *Center:* Diffusion around the resonance junction A (quasi-convex domain). The ellipse represents the constant energy condition of Eq. (39). *Right:* Same as previously, but for the resonant junction B (non-convex, steep). The constant energy condition (39) now yields hyperbolas

is $x = 0$. In the case $n = 3$ the three-jet condition is generically satisfied, as only coincidentally we can find a model H_0 in which all three Eqs. (40) be satisfied for some $x \neq 0$. However, when $n > 3$ the fulfillment of the condition depends on the choice of H_0 has to be checked case by case (see [60]).

Returning to the example of Fig. 7d, e, we can easily check the above conditions at the junctions A, B. We have $A = (I_{*1}, I_{*2}, I_{*3}) = (2\pi/3, 2\pi/3, 0)$, $B = (I_{*1}, I_{*2}, I_{*3}) = (2\pi/3, \pi/3, 0)$. We saw already that The Hessian matrix of H_0 is positive definite if $I_{*2} \geq \pi/2$. Thus H_0 is convex in the case A. At B, instead, we have $k_1 = (3, 1, 0)$, $k_2 = (3, 0, -1)$, thus

$$Y_B = \begin{pmatrix} -2 & 3 \\ 3 & 9 \end{pmatrix}$$

with opposite sign eigenvalues $\lambda_{1,2} = \frac{1}{2}(7 \pm \sqrt{157})$. This means that the quadratic form of Eq. (39) yields hyperbolas (see Fig. 8).

3.3 *Diffusion in the Web of Resonances*

We mentioned in Sect. 2 that it is possible to prove the existence of Arnold diffusion along the *simple* resonances of a priori stable systems (see CG). The first numerical example of Arnold diffusion in an a priori stable system similar to the one treated in the examples above (but with $H_0 = (I_1^2 + I_2^2) + I_3$ satisfying everywhere the quasi-convexity condition) was provided by [44]. Several more examples, including a spectacular demonstration of the drift of the trajectories throughout the entire Arnold web, were provided in [37].

Figure 8 (left) gives an example of the slow drift along the resonance (1, 1, 0) in the model (20) around the point I_* with $I_{1*} = 0.77211\ldots$, $I_{2*} = 1.3665$, $I_{3*} = 0$, for $\epsilon = 0.1$. Using the FLI map, we first compute the borders of the resonance (yellow). We then compute the plane of fast drift crossing the chosen point I_* (Eq. (29), thin

line in Fig. 8). Computing the FLI (for $t = 1000$) for initial conditions along this line, we obtain two points (on each separatrix layer) where the FLI has a local maximum. The point of maximum on the top right of the figure has co-ordinates $I_1 = 0.87166$, $I_2 = 1.466054$. Taking trajectories in a very small square (of size 10^{-5} in our case) around this point, and forward propagating these trajectories, allows to observe their slow drift along the separatrix layers of the resonance. The points in black in Fig. 8 correspond to only four such trajectories, integrated up to a time $t = 10^9$. The trajectories are shown only when returning to the same angular section $(\phi_1 + \phi_2) \mod 2\pi = 0$, and $\phi_3 \mod 2\pi = 0$ as the one for which the FLI was computed (with a numerical tolerance 10^{-2}). We notice that the trajectories make an overall excursion in the action space of length $\sim$0.5 after this long integration time. Due to the selected section, the trajectories yield points near the extrema of both branches of the theoretical separatrix of the resonance (see previous subsection), corresponding to the left and right groups of points in Fig. 8, which are both produced by the *same* trajectories. Besides the fast change in the resonant action I_R (Eq. (25)), we observe that the trajectories undergo a slow change of the value of the adiabatic actions I_{F1}, I_{F2}, a fact making them to jump from one to a nearby plane of fast drift, with all these planes parallel to the one shown in Fig. 8. How to quantify these jumps will be discussed in the next section.

The center and right panels of Fig. 8 refer now to chaotic trajectories around the resonant junctions A and B. We saw that the quadratic form of the constant energy condition of Eq. (39) yields ellipses in the case of the point A, while it yields hyperbolas in the case of the point B. Clearly, the chaotic trajectories around the junction are governed by this difference. In the case of the junction A, the normal form dynamics impedes the chaotic trajectories to move beyond a layer of thickness $\mathcal{O}(e)$ around each ellipse. In the case of the junction B, instead, the chaotic trajectories can have larger excursions by following a path close to the asymptotes of the hyperbolas. In that case, the trajectories are still limited around the resonant junction due to the cubic terms in the Hamiltonian (21).

On the other hand, all predictions made by the normal form models are valid up to an error determined by the exponentially small *remainder* of the normal form. More specifically, the Nekhoroshev normal form has the form:

$$H_N = Z_N + R_N \tag{41}$$

where Z_N is the normal form part and R_N the remainder, with

$$||R_N|| = \mathcal{O}\left(\exp((\epsilon_0/\epsilon)^b)\right) \quad .$$

Let I_{Fi} (the 'adiabatic actions') be the integrals of Z_N (in the 3DOF case, $i = 1, 2$ in the case of a simple resonance, and $i = 1$ in the case of the double resonance). We have

$$\dot{I}_{Fi} = -\partial R_N / \partial \phi_{Fi}, \quad i = 1, 2 \tag{42}$$

For the derivatives we have the estimate $||\dot{I}_{F_i}|| = \mathcal{O}\left(\exp((\epsilon_0/\epsilon)^b)\right)$. From this, we can conclude that, although the actions I_{Fi} cease to be integrals of motion in the complete Hamiltonian, up to a given time t the actions I_{F_i} can have excursions of length *bounded from above* by $\Delta I_{F_i} < \mathcal{O}\left(t \exp((\epsilon_0/\epsilon)^b)\right)$. This estimate yields the local speed of Arnold diffusion, which can hence be measured using the norm $||R_N||$. Another numerical test regards the comparison between the numerically computed (by ensembles of trajectories) value of the diffusion coefficient D, and the size of the remainder $||R_N||$. Empirical fitting has given the law $D \sim ||R_N||^3$ in the case of simple resonances, and $D \sim ||R_N||^2$ in the case of double resonances. Implementing the theory of Chirikov, instead, leads to the estimate $D \propto ||R_N||^{2+\alpha}$, where the correction $0 < \alpha < 1$ depends locally (in a simply-resonant domain) on the detailed structure of the 'layer resonances' determining the remainder of the local Nekhoroshev normal form [14].

To unveil the detailed evolution of the variables $I_{F,i}(t)$ for any trajectory one needs to solve the initial value problem for the differential equations (41) up to any desired time t. It turns out that, even availing the explicit expressions for a high order truncation of the remainder R_N, in practice it is hard to try to integrate the differential equations (42) directly in the computer. A good number of reasons impede us on this task, starting from the fact that the remainder R_N is actually a series, whose representation in the computer is given by a truncated trigonometric polynomial typically containing millions of terms. This is an expression hard to deal with not only numerically, but also in any theoretical attempt to establish the existence of phase space objects (e.g. manifolds like the ones of Fig. 5) having the role of drivers of Arnold diffusion.[3]

On the other hand, we can always attempt to *model* the dynamics of itself the remainder R_N. As discussed in the sequel, such a modeling is possible and leads to a way more tractable expression $R_N^{(model)}$. Using $R_N^{(model)}$ we can then probe and visualize most phenomena related to Arnold diffusion. In particular, we can unravel the 'jumps' in action space (similar as in Fig. 1d) undergone by the weakly chaotic trajectories within the layers of a selected resonance. We can also predict and model the size of these jumps. Finally, we can identify the fastest drifting trajectories and monitor how close their speed is to the theoretical upper bound provided by the Nekhoroshev theorem $\Delta I_{F_i}(max) = t \sup_{\mathcal{D}_*} |\partial R_N/\partial \phi_{Fi}|$ (see examples in the next section).

[3] While drifting along a simple resonance, a chaotic trajectory will eventually reach a multiple resonance domain. For some time, the trajectory then behaves as shown in the middle and right panels of Fig. 8. To demonstrate Arnold diffusion requires, however, showing that the trajectory will eventually exit from the multiple resonance, continuing to drift along the same exit simple resonance as the entry one, or choosing a different exit resonance. The lack of proof, in a priori stable systems, of the existence of a mechanism guaranteeing that these transitions will take place, is known as the 'large gap problem' [19, 22]. The existence of orbits undergoing long excursions in a priori stable systems, but far from double resonances, is demonstrated in [5].

4 Construction of the Nekhoroshev Normal Form: Semi-analytical Estimates

4.1 Construction of the Nekhoroshev Normal Form

It was mentioned before that most semi-analytical results on the quantification of the Arnold diffusion follow after the appropriate construction of a local Nekhoroshev normal form in a selected domain $\mathcal{D}_*$ around some point $I_* \in \mathbb{R}^n$ of the action space of the problem. We here summarize the method implemented in [14, 24, 26, 40], for an efficient computation of the Nekhoroshev normal form. We assume a n-DOF system with Hamiltonian

$$H(I, \phi) = H_0(I) + \epsilon H_1(I, \phi) \quad , \tag{43}$$

satisfying the properties enumerated below.

4.1.1 Analyticity

We assume that there is an open domain $\mathcal{I} \subset \mathbb{R}^3$ and real constants $\rho > 0$, $\sigma > 0$ such that for all points $I_* \in \mathcal{I}$ and all complex quantities $J_i \in \mathbb{C}$, $i = 1, \ldots n$ satisfying $|J_i| < \rho$ the following properties hold true:

(i) the function H_0 can be expanded as a convergent Taylor series

$$H_0 = H_{0*} + \omega_* \cdot J + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij*} J_i J_j + \cdots \tag{44}$$

where $\omega_* = \nabla_I H_0(I_*)$ and M_{ij*} are the elements of the Hessian matrix of H_0 at I_*, denoted by M_*.

(ii) For all $I_* \in \mathcal{I}$, H_1 admits a Fourier expansion

$$H_1 = \sum_k h_k(I_* + J) \exp(ik \cdot \phi) \tag{45}$$

analytic in the domain

$$\mathcal{D}(I_*) = \{I_i = I_{i*} + J_i, |J_i| < \rho, \Re(\phi_i) \in \mathbb{T}, |\Im(\phi_i)| < \sigma, i = 1, \ldots n\} \quad . \tag{46}$$

The analyticity of the function H_1 in the domain $\mathcal{D}$ implies that all the coefficients h_k can be expanded in convergent Taylor series around I_* as

$$h_k = h_{k*} + \nabla_{I_*} h_k \cdot J + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} h_{k,ij*} J_i J_j + \cdots \tag{47}$$

4.1.2 Book-Keeping

Due to the analyticity of H_1, the Fourier coefficients h_k in the domain $\mathcal{D}(I_*)$ decay exponentially, that is, there are positive constants A, σ such that

$$\sup_{\mathcal{D}_*} |h_k(I)| < Ae^{-|k|\sigma} \tag{48}$$

Taking the exponential decay into account, we then split the Fourier harmonics in groups with the wave number satisfying $(s-1) \leq |k| < sK - 1$, $s = 1, 2, \ldots$, and

$$K = -\frac{1}{\sigma} \log(\rho_0) \ , \tag{49}$$

where ρ_0 is the size of the domain around the point I_* where the normal form is to be computed, i.e., $|J_i| < \rho_0$. For resonant constructions of any multiplicity it is convenient to take $\rho_0 = \mathcal{O}(\sqrt{\epsilon})$. Introducing a 'book-keeping' symbol λ, with numerical value $\lambda = 1$, the Hamiltonian can then be split in ascending powers of λ:

$$H = H^{(0)}(J, \phi) = Z_0 + \sum_{s=1}^{\infty} \lambda^s H_s^{(0)}(J, \phi; \epsilon) \tag{50}$$

where

$$Z_0 = \omega_* \cdot J$$

and

$$H_s^{(0)} = \sum_{\mu=1}^{s} \sum_{k=K'(s-\mu)}^{K'(s-\mu+1)-1} H_{\mu,k}^{(0)}(J) \exp(ik \cdot \phi) \tag{51}$$

where $H_{\mu,k}^{(0)}(J)$ are polynomials containing terms of degree $\mu - 1$ or μ in the action variables J. In the $n = 3$ cases dealt with in the numerical examples of this article, we have, in particular:

$$H_{\mu,k}^{(0)}(J) = \epsilon \sum_{\mu_1=0}^{\mu-1} \sum_{\mu_2=0}^{\mu-1-\mu_1} \sum_{\mu_3=0}^{\mu-1-\mu_1-\mu_2} \frac{1}{\mu_1!\mu_2!\mu_3!} \frac{\partial^{\mu-1} h_{1,k}(I_*)}{\partial^{\mu_1} I_1 \partial^{\mu_2} I_2 \partial^{\mu_3} I_3} J_1^{\mu_1} J_2^{\mu_2} J_3^{\mu_3}$$

if $|k| > 0$, or

$$H_{\mu,k}^{(0)}(J) = \sum_{\mu_1=0}^{\mu} \sum_{\mu_2=0}^{\mu-\mu_1} \sum_{\mu_3=0}^{\mu-\mu_1-\mu_2} \frac{1}{\mu_1!\mu_2!\mu_3!} \frac{\partial^{\mu} H_0(I_*)}{\partial^{\mu_1} I_1 \partial^{\mu_2} I_2 \partial^{\mu_3} I_3} J_1^{\mu_1} J_2^{\mu_2} J_3^{\mu_3}$$

$$+ \epsilon \sum_{\mu_1=0}^{\mu-1} \sum_{\mu_2=0}^{\mu-1-\mu_1} \sum_{\mu_3=0}^{\mu-1-\mu_1-\mu_2} \frac{1}{\mu_1!\mu_2!\mu_3!} \frac{\partial^{\mu-1} h_{1,0}(I_*)}{\partial^{\mu_1} I_1 \partial^{\mu_2} I_2 \partial^{\mu_3} I_3} J_1^{\mu_1} J_2^{\mu_2} J_3^{\mu_3}$$

if $k = 0$. In all the above expressions, the superscript (0) means 'the starting Hamiltonian of the iterative normalization process'. This is simply the original Hamiltonian re-organized in powers of the book-keeping symbol λ. Subscripts (as e.g. s in the functions $H_s^{(0)}(J, \phi; \epsilon)$) mean terms book-kept with the power λ^s. In physical terms, this can be interpreted as 'terms of the sth order of smallness'. All expressions in the initial and in subsequent normalization steps are finite, i.e., they are trigonometric polynomials easily represented in the computer's memory via an indexing function. The maximum 'book-keeping' order N_{tr} adopted in the normalization algorithm is called the truncation order.

4.1.3 Resonant Module

Following the definitions given in Sect. 4.2, the point I_*, and its corresponding frequency vector $\omega_* = \omega(I_*)$, are called 'M-tuple resonant' (with $0 \leq M \leq n-1$) if there can be found M linearly independent non-zero integer vectors $k^{(i)}, i = 1, \ldots M$ such that $k^{(i)} \cdot \omega_* = k^{(i)} \cdot \omega(I_*) = 0$ for all $i = 1, \ldots, M$. When a point I_* is M-tuple resonant, there are many harmonics $\cos(k \cdot \phi)$ with $|k| \neq 0$ in the Hamiltonian which cannot be normalized since their elimination would involve a divisor exactly equal to zero. The set of all possible wavevectors k such that $k \cdot \omega_* = 0$ is called the *resonant module* at the point I_*. Since checking numerically the condition $k \cdot \omega_* = 0$, with $\omega_* \in \mathbb{R}^n$, is sensitive to round-off errors, a convenient way to define the resonant module, which involves only operations among integer numbers, is by use of the concept of 'pseudo-frequency' vector. This is defined as follows: if ω_* is M-tuple resonant with $M \geq 1$, choose M non-zero linearly independent integer vectors $k^{(i)}$, $i = 1, \ldots, M$ such that $k^{(i)} \cdot \omega(I_*) = 0$. Then, there exist $n - M$ non-zero integer vectors $m^{(j)}$, $j = 1, \ldots, n - M$ such that $k^{(i)} \cdot m^{(j)} = 0$ for all possible pairs i, j. To define these vectors, solve the $n - M$ systems of linear equations given by

$$\begin{aligned} & k_1^{(1)} q_1^{(j)} + k_2^{(1)} q_2^{(j)} + \cdots + k_M^{(1)} q_M^{(j)} = -k_{M+j}^{(1)} \\ & k_1^{(2)} q_1^{(j)} + k_2^{(2)} q_2^{(j)} + \cdots + k_M^{(2)} q_M^{(j)} = -k_{M+j}^{(2)} \\ & \cdots \\ & k_1^{(M)} q_1^{(j)} + k_2^{(M)} q_2^{(j)} + \cdots + k_M^{(M)} q_M^{(j)} = -k_{M+j}^{(M)} \end{aligned} \tag{52}$$

for $j = 1, \ldots, n - M$. The solutions give vectors $q^{(j)} = (q_1^{(j)}, \ldots, q_M^{(j)}, \delta_{M+1,M+j}, \ldots, \delta_{n,M+j})$ with rational components. Multiplying the vector $q^{(j)}$ with the maximal common divisor of all its components yields the jth pseudo-frequency vector $m^{(j)}$.

We can now determine which harmonics $\cos(k \cdot \phi)$ to be excluded from the normalization process. The set of all integer vectors k corresponding to the excluded harmonics is called the *resonant module* $\mathcal{M}(k^{(1)}, \ldots, k^{(M)})$ defined as:

$$\mathcal{M}(k^{(1)},\dots,k^{(M)}) = \begin{cases} \{k=(0,0,\dots,0)\} & \text{if } M=0 \\ \{k\in\mathbb{Z}^n : k\cdot m^{(j)}=0 \quad \text{for all } j=1,\dots,n-M\} & \text{if } M>0 \end{cases} \tag{53}$$

where $m^{(j)}$, $j=1,\dots,n-M$ are the pseudo-frequency vectors determined through Eq. (52).

Note that, even when the origin of the expansion I_* is non-resonant, i.e., when $M=0$, arbitrarily close to it there can be found M-tuple resonant points of any multiplicity $M>0$. This is a consequence of the fact that resonances are dense in the action space (see the examples in [25]). Whenever the non-resonant vector ω_* is 'close' to a low-order M-tuple resonant vector Ω, in the sense that $|\omega_*-\Omega|<\alpha$ with α small, and the wavevectors k satisfying $k\cdot\Omega$ are of order $|k|$ smaller than the 'cut-off' order (see below), we say to be in a 'near-resonance' case. In this case too, we may wish to avoid the presence in the series of those divisors $k\cdot\omega_*$ for which $k\cdot\Omega=0$. We then define the resonant module as above, but using Ω in the place of ω_*.

4.1.4 Hamiltonian Normalization

We consider a sequence of normalizing canonical transformations

$$(\phi,J)\equiv(\phi^{(0)},J^{(0)})\to(\phi^{(1)},J^{(1)})\to(\phi^{(2)},J^{(2)})\to\cdots$$

leading to re-express the Hamiltonian, after r normalization steps, in new canonical variables $(\phi^{(r)},J^{(r)})$ such that

$$H(\phi((\phi^{(r)},J^{(r)}))=Z^{(r)}((\phi^{(r)},J^{(r)});\lambda,\epsilon)+R^{(r)}(\phi^{(r)},J^{(r)};\lambda,\epsilon)\ . \tag{54}$$

The functions $Z^{(r)}(J^{(r)},\phi^{(r)};\lambda,\epsilon)$ and $R^{(r)}(J^{(r)},\phi^{(r)};\lambda,\epsilon)$ are called the normal form and the remainder respectively. The normal form is a finite expression which contains terms up to order r in the book-keeping parameter λ. By definition, these are terms belonging to the resonant module $\mathcal{M}(k^{(1)},\dots,k^{(M)})$. The remainder, instead, is a convergent series containing terms of order λ^{r+1}, including all possible harmonics.

To compute the normalizing transformation, we use the composition of Lie series with generating functions $\chi_1,\dots,\chi_r$. Denote $Q=(\phi,J)\equiv Q^{(0)}$. The normalizing transformation is:

$$Q^{(r)}=\exp(-L_{\chi_1})\exp(-L_{\chi_2})\dots\exp(-L_{\chi_r})Q \tag{55}$$

The generating functions are determined recursively, by solving, for $n_r=0,\dots,r-1$ the homological equations:

$$\{\omega_*\cdot J^{(n_r+1)},\chi_{n_r+1}\}+\lambda^{n_r+1}\tilde{H}^{(n_r)}_{n_r+1}(J^{(n_r+1)},\phi^{(n_r+1)})=0 \tag{56}$$

where

$$H^{(n_r)} = \exp(L_{\chi_{n_r}}) H^{(n_r-1)} \quad . \tag{57}$$

4.1.5 Optimal Remainder

Basic normal form theory (see [25]) establishes that the above normalization process has an *asymptotic* character. Namely, (i) the domain of convergence of the remainder series $R^{(r)}$ shrinks as the normalization order r increases, and (ii) the size $||R^{(r)}||$ of $R^{(r)}$, where $||\cdot||$ is a properly defined norm in the space of trigonometric polynomials, initially decreases, as r increases, up to an optimal order r_{opt} beyond which $||R^{(r)}||$ increases with r. In the *Nekhoroshev regime*, one has $||Z^{(r_{opt})}|| >> ||R^{(r_{opt})}||$. Hence, the normal form obtained at the order r_{opt} best unravels the dynamics, which is given essentially by the Hamiltonian flow of $Z^{(r_{opt})}$ slightly perturbed by $R^{(r_{opt})}$. Furthermore, the optimal normalization order r_{opt} depends on ϵ via an inverse power-law [24, 26], namely

$$r_{opt} \sim \epsilon^{-a} \quad , \tag{58}$$

for some positive exponent a depending on the multiplicity of the resonance around which the normal form is computed. The leading terms in the optimal remainder function are $O(\lambda^{r_{opt}+1})$. Due to the book-keeping relation (49), the terms of order $\lambda^{r_{opt}}$ have size estimated as $e^{-\sigma K_{opt}}$, where

$$K_{opt}(\epsilon) = K' r_{opt}(\epsilon) \tag{59}$$

is called the *Nekhoroshev cut-off* order. Then, $K_{opt} \sim K'\epsilon^{-a}$, implying:

$$||R^{(r_{opt})}|| \sim \epsilon^{1/2} \exp\left(\frac{-K'\sigma}{\epsilon^a}\right) \tag{60}$$

i.e., the remainder at the optimal normalization order is exponentially small in $1/\epsilon$.

In practice, to specify the optimal normalization order, after performing all the above symbolic computations with the aid of a computer program, we proceed as follows: we set the truncation order N_t to be several orders larger than the maximum reached normalization order r. Then, we compute the truncated-norm estimates

$$||R^{(r)}||_{W^{(r)}} = \sum_{s=r+1}^{N_t} \sum_{m} \sup |R_s^{(r)}|_{W^{(r)}} \tag{61}$$

where $\sup |R_s^{(r)}|$ means the sup norm of the sth book-keeping term of the truncated remainder over a domain of interest $Q^{(r)} \in W^{(r)}$ where the rth step canonical variables. To this end, we first probe numerically that $W^{(r)}$ is smaller than the convergence domain for the rth step normalization. We then verify the asymptotic character of the sequence $||R^{(r)}||_{W^{(r)}}$, for $r = 1, 2, 3, \ldots$. That is, for ϵ sufficiently small, initially (at

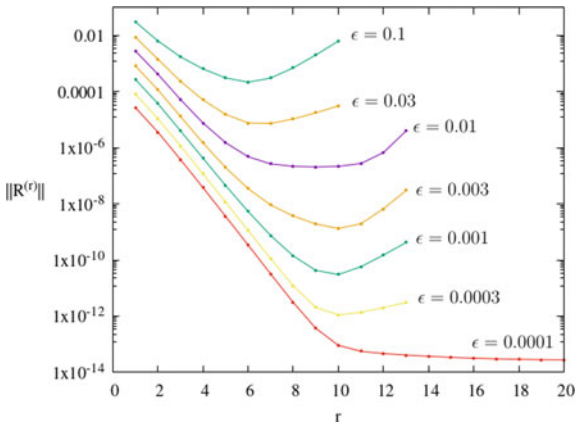


Fig. 9 Size of the remainder as a function of the normalization order r for various values of ϵ. The value of r at the minimum of each curve corresponds to the optimal normalization order. Note that the optimal order is higher than 20 in the case $\epsilon = 0.0001$

low orders) $||R^{(r)}||_{W^{(r)}}$ decreases as r increases, up to the optimal order r_{opt} at which $||R^{(r_{opt})}||_{W^{(r_{opt})}}$ reaches a minimum. Then, for $r > r_{opt}$, $||R^{(r)}||_{W^{(r)}}$ increases with r. This behavior is exemplified in Fig. 9, referring to the normal form computed for the data of the simple resonance corresponding to the left panel of Fig. 8.

4.2 Removal of Deformation Effects

We have seen that, at the optimal order, the adiabatic actions $I_{Fi}^{(r_{opt})}$ are integrals of the normal form dynamics, while in the full Hamiltonian they undergo exponentially small time variations due to the exponentially small optimal remainder. One important effect, which impedes to measure the real speed of the variations of the adiabatic action variables is *deformation*. Consider the inverse of the transformation (55) at optimal order:

$$Q = \exp(L_{\chi_r}) \exp(L_{\chi_{r-1}}) \dots \exp(L_{\chi_1}) Q^{(r_{opt})} \tag{62}$$

Due to the relation $\exp(L_{\chi_1}) Q^{(r_{opt})} = Q^{(r_{opt})} + \{Q^{(r_{opt})}, \chi_1\} + \cdots$, as well as the fact that $\chi_s = \mathcal{O}(\rho_0^s)$, we have that

$$Q = Q^{(r_{opt})} + \mathcal{O}(\rho_0) \tag{63}$$

Furthermore, for resonant normal forms, we saw that $\rho_0 = \mathcal{O}(\epsilon^{1/2})$. Thus, we find that even while the adiabatic actions $I_{Fi}^{(r_{opt})}$ undergo a very slow time variation (including drift), in the original variables this variation is completely hidden in a $\mathcal{O}(\epsilon^{1/2})$ oscillation, due entirely to the canonical transformation linking old with new variables. Since, without knowledge of the normalizing transformation, we are forced to deduce all the information on the behavior of the system by numerical experiments performed using the original variables, this implies that we have to recover

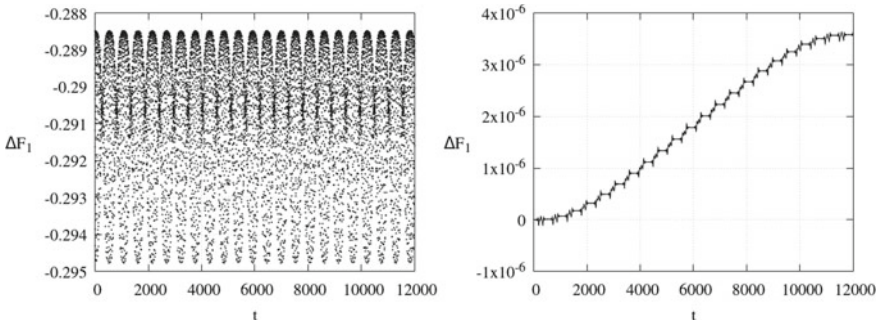


Fig. 10 Evolution of the adiabatic action J_F along a simple resonance in the model (20). *Left*: numerical trajectory. *Right:* the same trajectory, but plotted in the optimal canonical variable $J_F^{(r_{opt})}$ (see text)

the drift by removing all the noise induced by these large amplitude, but irrelevant for dynamics, oscillations.

Being able to compute the optimal normalizing transformation, allows, instead to spectacularly remove the deformation effect and easily obtain (and measure) the underlying drift of the adiabatic action variables. Figure 10 shows the removal of the deformation in the case of a trajectory undergoing Arnold diffusion in the model (20) and with initial condition as in Fig. 8. Recall that to visualize the drift using the original variables in that case has required an extremely long integration time $t = 10^9$. For quite shorter times, instead, ($t = 10^4$ in Fig. 10) the drift of the unique adiabatic action of the problem, measured by $\Delta F, 1 = I_F(t) - I_F(0)$ is completely hidden in a oscillation of size 0.2 (Fig. 10, left), and thus impossible to measure with numerical experiments up to the time $t = 10^4$. If, instead, we pass all the numerical data $Q(t)$ of the trajectory through the optimal normalizing transformation (Eq. (55)), we obtain the evolution of the optimal variable $\Delta F, 1^{(r_{opt})} = I_F^{(r_{opt})}(t) - I_F^{(r_{opt})}(0)$, shown in Fig. 10, right. Now, the drift is clearly demonstrated, and its local velocity can be measured by a simple fitting to the data. In fact, as discussed in the next subsection, the drift in the action space is not necessarily monotone, and $\Delta F, 1^{(r_{opt})}(t)$ may exhibit both an increase or decrease at different intervals of time. At any rate, the ability to remove the deformation effect can be exploited in the modeling of the evolution of the adiabatic action variables, as discussed in the next subsection.

4.3 Modeling the Jumps in the Adiabatic Action Variables

We have mentioned that it is possible to prove the occurrence of the Arnold mechanism in a priori stable systems only in the case of simple resonances (CG). We will now discuss how to model the evolution of the adiabatic action variables, including the jumps similar in nature as those of the original Arnold model, using, however,

the information encapsulated in the remainder at the optimal normalization order. Consider an optimal Hamiltonian of the form (54) obtained by normalization around a simply-resonant point I_*.

Following [40], to simplify all notations, denote as H^N (the 'Nekhoroshev normal form') the Hamiltonian $H^{(r_{opt})}$, depending on the resonant action-angle variables $(S, \sigma) \equiv \left(J_R^{(r_{opt})}, \phi_R^{(r_{opt})}\right)$ and the $n-1$ adiabatic action variables conjugate to fast angles $(F, \phi) \equiv \left(J_F^{(r_{opt})}, \phi_F^{(r_{opt})}\right)$ (see Sect. 3). With the new notation, we have

$$H^N = h(F, S) + \epsilon f^N(F, S, \sigma) + r^N(F, S, \sigma, \phi) \quad . \tag{64}$$

The (simply-resonant) normal form is

$$\overline{H}^N = h(F, S) + \epsilon f^N(F, S, \sigma) \, . \tag{65}$$

The remainder r^N is provided as a Taylor-Fourier series:

$$r^N = \sum_{m \geq 0} \sum_{\nu \in \mathbb{Z}^d} \sum_{k \in \mathbb{Z}^{n-d}} r^m_{\nu,k}(F)(S - S_*)^m e^{i\nu \cdot \sigma + ik \cdot \phi} \tag{66}$$

expanded at a suitable S_*, with computer-evaluated truncations involving a large number (typically 10^7 to 10^8) terms.

To define the resonant normal form dynamics, as in Sect. 3 we first expand $\overline{H}^N$ at the values of the actions (F_*, S_*) identifying the center of the resonance, where

$$\frac{\partial h}{\partial S}(S_*, F_*) = 0. \tag{67}$$

Then

$$\overline{H} = \overline{H}_0 + \cdots \quad , \quad \overline{H}_0 = \omega_* \cdot \hat{F} + \frac{A}{2}\hat{S}^2 + \hat{S}B \cdot \hat{F} + \frac{1}{2}C\hat{F} \cdot \hat{F} + \epsilon v(\sigma) \tag{68}$$

where $\hat{F} = F - F_*$, $\hat{S} = S - S_*$, $A \in \mathbb{R}$, $\omega_*, B \in \mathbb{R}^{n-1}$, C is a $(n-1) \times (n-1)$ square matrix and $v(\sigma)$ is a trigonometric function depending parametrically on $S_*(I_*)$, $F_*(I_*)$. The actions $\hat{F}$ are the constants of motion for the Hamiltonian flow of $\overline{H}_0$.

Consider, now, the family of curves $\hat{S}(u; \alpha)$, for different a, given by

$$\hat{S} = \sqrt{\epsilon} s_\alpha(\sigma) = \pm\sqrt{\epsilon}\sqrt{\frac{2}{|A|}(M(1+\alpha) - v(\sigma))} \tag{69}$$

where $M = \max_{\sigma \in [0, 2\pi]} v(\sigma)$, and α is the energy of the pendulum Hamiltonian (equal to $\overline{H}_0$ for $\hat{F} = 0$):

$$a = \frac{A}{2}\hat{S}^2 + \epsilon v(\sigma) = \epsilon M(1+\alpha) \quad . \tag{70}$$

Since $\overline{H}_0$ has the structure of a pendulum Hamiltonian, we can attempt to implement the Melnikov approximation, introduced in Sect. 2, in order to compute the jumps in the variables F over one complete homoclinic transition of the variables $(\hat{S}, \sigma)$, assigning to the remainder r^N (Eq. (66)) the role of the coupling term between the resonant variables (S, σ) and the remaining variables (F, ϕ). Since $\dot{F}_j = -\partial r^N/\partial\phi_j$, the Melnikov approximation will then consist of estimating the variation $\Delta F_j(T) = F_j(t) - F_j(0)$ after a time T via the integral

$$\Delta F_j(T) = -\sum_{m,\nu,k} \int_0^T ik_j r^m_{\nu,k}(F(t))\hat{S}(t)^m e^{i\nu\sigma(t)+ik\cdot\phi(t)}dt := \sum_{m,\nu,k} \Delta F^{m,\nu,k}_{j,T} \quad . \tag{71}$$

where the true solution $(F(t), S(t), \sigma(t), \phi(t))$ in the r.h.s of the integrals (71) will be substituted by the approximate solution under the flow of the normal form $\overline{H}_0$

$$(F_*, S^0(t), \sigma^0(t), \phi^0(t)) = (F_*, S_*, 0, 0) + (0, \hat{S}^0(t), \sigma^0(t), \phi^0(t))$$

where $(0, \hat{S}^0(t), \sigma^0(t), \phi^0(t))$ is a solution of Hamilton's equations of $\overline{H}_0$.

Contrary to the simple model of Sect. 2, it is important to recall that the number of Melnikov integrals to compute in (71) are of the same order as the number of remainder terms (10^7 to 10^8), thus the computation is hardly tractable in practice. However, we get an enormous simplification of the problem noticing that, out of all these integrals, only few ($\sim 10^3$) really contribute to the result. To this end, we first observe that representing $\hat{S}^0(t)$ parametrically as a function of $\sigma^0(t)$, for fixed α, allows to change the integration variable in (71) from t to σ:

$$\Delta F^{m,\nu,k}_{j,T}(T) \simeq \Delta^0 F^{m,\nu,k}_{j,T}(T) = -ik_j \frac{r^m_{\nu,k}(F_*)\epsilon^{\frac{m-1}{2}}}{A} e^{ik\cdot\phi(0)} \int_0^{\sigma^0(T)} [s_\alpha(\sigma)]^{m-1} e^{i\theta(\sigma)} d\sigma \tag{72}$$

where the phase $\theta(\sigma)$ is defined by:

$$\theta(\sigma) = \mathcal{N}\sigma + \frac{\Omega}{A\sqrt{\epsilon}} \int_0^\sigma \frac{dx}{s_\alpha(x)}$$

with

$$\mathcal{N} = \nu + k \cdot B/A, \quad \Omega = k \cdot \omega_* \quad . \tag{73}$$

Then, invoking the principle of stationary phase, it is clear that only integrals involving a slow variation of the phase $\theta(\sigma)$ over a time T_α, representing the period of one homoclinic transition, will be important in the computation of the jumps via the Eq. (72).

To make this argument more explicit, assume that the lowermost order terms in the resonant normal form (for $\hat{F} = 0$) have the form of the pendulum Hamiltonian:

$$H_{pend} = \frac{|A|}{2}\hat{S}^2 + \epsilon\beta\cos\sigma + \cdots \tag{74}$$

where, for simplicity, we set $\epsilon, \beta > 0$. Consider a remainder term labeled by the integers (m, ν, k) in Eq. (72). Using the approximation (74), and setting $\alpha = 0$ (separatrix solution), the function $\theta(\sigma)$ for the term in question can be approximated by:

$$\theta(\sigma) \approx \theta_0 + \mathcal{N}\sigma + \mathcal{W}\ln\tan(\sigma/4), \quad \mathcal{W} = \frac{\Omega}{\sqrt{|A|\beta\epsilon}} \ . \tag{75}$$

where $\mathcal{N}$ and Ω are given by Eq. (73), hence, they depend only on the term labels ν, k. From Eq. (75), we obtain

$$\theta'(\sigma) \approx \mathcal{N} + \frac{\mathcal{W}}{2}\frac{1}{\sin(\sigma/2)} \tag{76}$$

Therefore, one has $\lim_{\sigma\to 0} W\theta'(\sigma) = \lim_{\sigma\to 2\pi} W\theta'(\sigma) = +\infty$, and since $\theta'(\sigma)$ is a function symmetric with respect to π and monotonically decreasing (increasing) in $[0, \pi)$ ($(\pi, 2\pi]$), there exists a minimum of the function at $\sigma = \pi$ of value $\theta'(\sigma) = \mathcal{N} + \mathcal{W}/2$. Thus, $\theta'(\sigma)$ has zeroes (stationary points) $\sigma_c = \pi \pm \Delta\sigma_c$, with $0 < \Delta\sigma_c < \pi$, if and only if the minimum value $\theta'(\pi)$ is negative. This lead to the following condition:

$$\text{The term defined by } (m, \nu, k) \text{ is stationary } \iff \mathcal{N}\cdot\mathcal{W} < 0 \text{ and } |\mathcal{N}| > \frac{|\mathcal{W}|}{2} \tag{77}$$

In case the condition (77) is not satisfied, we still have to check for the existence of terms (m, ν, k) which, albeit non-stationary, exhibit only a small variation of the phase $\theta(\sigma)$ over the period of the homoclinic transition. Such terms will be called *quasi-stationary* and they can be selected from the remainder by the following procedure: neglecting the slowly varying factor $[s_a(\sigma)]^{m-1}$ in Eq. (72), and factoring out a constant phase $e^{i(\theta_0+\mathcal{N}\pi)}$, important quasi-stationary terms are those for which the integral

$$\Delta\mathcal{I} = \int_0^{2\pi} \cos\left(\mathcal{N}(\sigma - \pi) + \mathcal{W}\ln\tan(\sigma/4)\right) \tag{78}$$

has absolute value above a small (arbitrarily chosen) threshold μ_0. Consider for a moment the approximation $\mathcal{W} \simeq const.$. Since the inspected term is assumed not to be stationary (not selected by the condition (77)), we have that $\mathcal{N}$ varies according to $\mathcal{N} \geq -\mathcal{W}/2$ for $\mathcal{W} > 0$, or $\mathcal{N} \leq -\mathcal{W}/2$ for $\mathcal{W} < 0$. Different values of $\mathcal{N}$ generate different behaviors for $\theta(\sigma)$, symmetric with respect to $\sigma = \pi$, as shown in Fig. 11a. Figure 11b shows the functions $\cos\left(\mathcal{N}(\sigma - \pi) + \mathcal{W}\ln\tan(\sigma/4)\right)$, for the same frequencies σ of panel (a). From the comparison of the two plots, we

see that the nearly flat domains of the curve $\theta(\sigma)$ near $\sigma = \pi$, along with the sigmoid variations at the two ends (in panel (a)) imply the formation of a plateau of the curves in (b) accompanied by fast lopsided oscillations, which nearly cancel each other in the integral (78). The flatter the function $\theta(\sigma)$ in the vicinity of $\sigma = \pi$, the wider is the plateau of $\cos(\theta(\sigma))$. Since the dominant contribution in $\Delta\mathcal{I}$ comes from the central plateau of $\cos(\theta(\sigma))$, the maximum absolute value of $\Delta\mathcal{I}$ occurs when the slope $\theta'(\sigma)$ becomes zero at $\sigma = \pi$. Hence, from Eq. (76), the maximum occurs when $\mathcal{N} = -\mathcal{W}/2$. The length of the plateau is given by $\Delta\sigma_p = 2\sigma_p$, where $\theta(\pi \pm \sigma_p) = \pi/2$. From Eq. (75), we find $\sigma_p \simeq (24\pi/\mathcal{W})^{1/3}$, and hence $\Delta\mathcal{I}_{\mathcal{N}=-\mathcal{W}/2} \propto \mathcal{W}^{-1/3}$, an estimate verified numerically (Fig. 11e).

On the other hand, if $\mathcal{N}$ is 'detuned' from the maximum value $-\mathcal{W}/2$, the associated plateaus attenuate, leading to a decrease of $\Delta\mathcal{I}$. Yet, some of these contributions can be larger than minimum threshold considered for Eq. (78). Setting $\mathcal{N} = (\delta - 1)\mathcal{W}/2$, Fig. 11f shows the attenuation as function of the detuning δ for fixed $\mathcal{W}$. For small δ, the attenuation is nearly a linear function of δ with negative slope, $\Delta\mathcal{I} \propto \Delta\sigma_p \approx (24\pi)^{1/3}\mathcal{W}^{-1/3} - (64/3\pi)^{1/3}\mathcal{W}^{1/3}\delta$. If we extend the straight line with negative slope in Fig. 11f up to the point where the line intersects the axis $\Delta I = 0$ we find a critical detuning $\delta_c \approx (3\pi/2\sqrt{2})^{2/3}\mathcal{W}^{-2/3}$ beyond which the term can no longer be characterized as quasi-stationary. Actually, δ_c computed as above underestimates the true value of the detuning, since (i) the curve $\Delta\mathcal{I}$ has a tail extending only asymptotically to zero (i.e. as small as it may be, the contribution of a quasi-stationary terms is never exactly zero) and (ii) the slope found by linear fitting of the left part of the curves ΔI versus δ for various values of $\mathcal{W}$ shows that the power law estimate of the slope $\propto \mathcal{W}^p$ yields an exponent substantially larger than $1/3$ for values of $\mathcal{W}$ well below unity (Fig. 11g). On the other hand, a numerical evaluation of the dependence of the critical detuning δ_c as function of $\mathcal{W}$ (Fig. 11h) yields a law $\delta_c \propto \mathcal{W}^{-q}$, with $q \approx 0.8$, i.e., slightly larger than the theoretical estimate $q = 2/3$. Taking into account all these considerations, we formulate a heuristic criterion for quasi-stationarity, namely:

$$\begin{aligned} &\text{The term defined by } (m, \nu, k) \text{ is quasi} - \text{stationary} \\ &\iff \mathcal{N} \cdot \mathcal{W} < 0 \text{ and } |\mathcal{N}| < (1 - \delta_c)\frac{|\mathcal{W}|}{2} \end{aligned} \tag{79}$$

with $\delta_c = \delta_{c0}|\mathcal{W}|^{-0.8}$, where, by numerical fitting, $\delta_{c0} \simeq 3$ for an adopted attenuation factor 0.1, or $\delta_{c0} = 4.2$ for an adopted attenuation factor $\sim$0.01.

The conditions (77) and (79) are derived by considering the upper branch of the separatrix solution $\theta(\sigma)$. For the lower branch we have, instead, $\theta(\sigma) = \mathcal{N}\sigma - \mathcal{W}\ln\tan(\sigma/4)$, hence we obtain the same conditions for stationary of quasi-stationary terms, but with the inequality $\mathcal{N} \cdot \mathcal{W} > 0$ instead of $\mathcal{N} \cdot \mathcal{W} < 0$. Also, the above analysis, based solely on the behavior of the phase $\theta(\sigma)$, allows to identify stationary or quasi-stationary terms for $|\mathcal{W}|$ arbitrarily large. It is important to recognize that the quantity $\Omega = k.\omega_* = \mathcal{W}(|A|\beta\epsilon)^{1/2}$, represents the divisor associated with the remainder term (m, ν, k). Thus, we may further restrict the selection of remainder

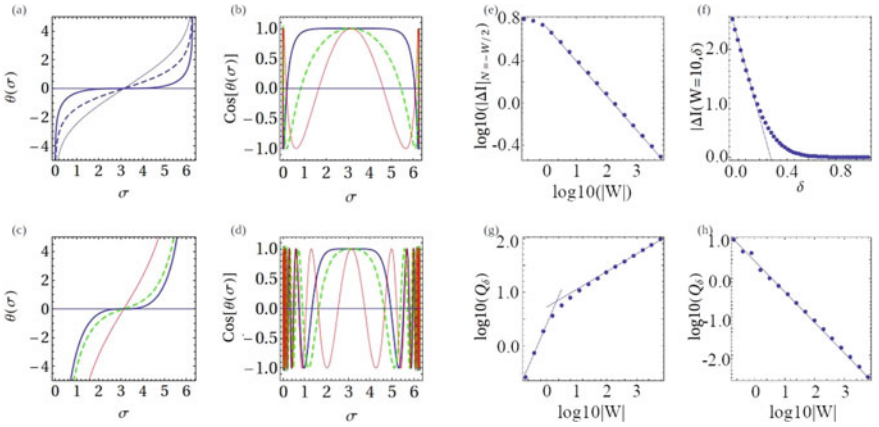


Fig. 11 **a** The function $\theta(\sigma)$ (Eq. 75) written as $\theta(\sigma) = \mathcal{W}[0.5(1-\delta)](\sigma - \pi) + \ln\tan(\sigma/4))$ for $\mathcal{W} = 1$ and $\delta = 0$ (thick blue), $\delta = 1$ (dashed green), or $\delta = 2$ (thin red). The corresponding curves $\cos(\theta(\sigma))$ are shown in (**b**). The extent of the 'plateau' is reduced for larger δ. Similar curves are shown in **c** and **d** for $\mathcal{W} = 10$, and $\delta = 0$, 0.1 and 0.5. **e**, the integral ΔI (Eq. 78) for $\delta = 0$, as a function of $|\mathcal{W}|$. (f) The attenuation of the integral ΔI with respect to its value for $\delta = 0$ as δ increases, for fixed $\mathcal{W} = 10$. The linear part of the curve, for small δ can be fitted with a line of negative slope Q_δ. (g) The slope $|Q_\delta|$ as a function of $|\mathcal{W}|$. (h) The critical value δ_c for which the integral ΔI attenuates to 10% its value at $\delta = 0$, as a function of $|\mathcal{W}|$

terms by retaining only those passing the stationary or quasi-stationary criterion, and simultaneously satisfying an upper threshold for the divisor value, say $|\Omega| < 1$.

Figure 12 shows the main result obtained by selecting only the few terms ($\sim$1000) of the remainder passing the criteria of stationarity or quasi-stationarity. Swarms of 100 trajectories with initial conditions very close to the hyperbolic torus at the simply resonant point I_* same as in Fig. 8, but for $\epsilon = 0.003$ (top left) or $\epsilon = 0.01$ (top right) for a very small time ($T = 1200$ and $T = 700$ respectively), corresponding to the time required for the orbits to complete the first homoclinic transition along the pendulum, according to the approximative formula:

$$T_\alpha = \frac{1}{\sqrt{A\epsilon\beta}} \ln(32A\epsilon\beta/||R^{opt}||) \tag{80}$$

This formula is the same as Eq. (11) used in Sect. 2, setting the pendulum energy as $\varepsilon = A\epsilon\beta$, with the coefficients A and β obtained from the simply-resonant normal form (Sect. 3). As discussed before, showing numerically computed original values of the adiabatic action $J_F(t)$ for these trajectories provides no information, due to the deformation effect. Showing, however, the same variable at optimal order by use of the transformation (55) makes clear the jumps along homoclinic transitions exhibited by these trajectories. In particular, we distinguish how the random distribution of the initial phases results in a stochastic spreading of the actions $F_1(t) = J_F^{(r_{opt})}(t)$ (with $r_{opt} = 10$ in the left panel, and $r_{opt} = 7$ in the right panel), in a way qualitatively

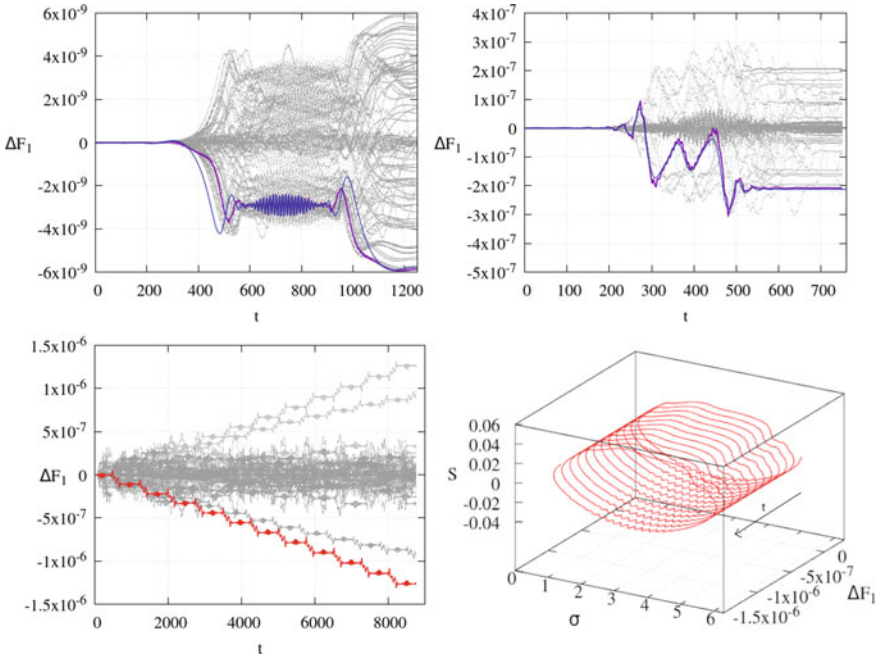


Fig. 12 *Top:* Evolution of $F_1(t)$ for a swarm of 100 trajectories with initial conditions very close to the hyperbolic torus at the simply resonant point I_* same as in Fig. 8, but for $\epsilon = 0.003$ (left) or $\epsilon = 0.01$ (right). The blue curves show the fitting to one trajectory of the swarm using the Melnikov integrals (71) for only those remainder terms selected as stationary or quasi-stationary. *Bottom:* A ballistic orbit, drifting continuously in the same direction along the resonance, as depicted for the evolution $F_1(t)$ (left) or $F_1 \times$ the pendulum variables S, σ (right)

similar to the one observed at Fig. 4 in Arnold's model. The bottom left panel extends the calculation in the case $\epsilon = 0.01$ up to a time $t = 9000$. At this time the trajectories have undergone 13 transitions. The jump ΔF_1 in every transition shows the behavior of a random walk with size 10^{-7}. Thus, most trajectories spread over an interval $(-(13)^{1/2}10^{-7}, (13)^{1/2}10^{-7})$. However, we distinguish also rare trajectories which move in 'ballistic' motion, i.e., drifting systematically in the same direction. These are the fastest moving trajectories, with speed bounded by an estimate which is the closest possible to the absolute bound provided by the Nekhoroshev theorem.[4] Note, finally, the excellent representation of the jumps by the semi-analytical (Melnikov) approximation (blue curves) using only the remainder terms selected by the stationarity or quasi-stationarity criteria.

For more rigorous statements on the (quasi-)stationary phase approximation method see [40].

[4] A variational method to compute such fastest drifting trajectories in a priori unstable systems is proposed in [62].

Acknowledgements The authors acknowledge the project MIUR-PRIN 20178CJA2B "New frontiers of Celestial Mechanics: theory and applications". C.E. acknowledges the support to this project by the H2020 MSCA ETN Stardust-R (GA 813644).

References

1. Arnold, V.I.: Proof of a theorem of A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. Usp. Mat. Nauk. **18**, 13; Russ. Math. Surv. **18**, 9 (1963)
2. Arnold, V.I.: Instability of dynamical systems with several degrees of freedom. Sov. Math. Dokl. **6**, 581 (1964)
3. Benettin, G., Galgani, L., Giorgilli, A.: A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems. Celest. Mech. Dyn. Astron. **37**, 1 (1985)
4. Benettin, G., Gallavotti, G.: Stability of motions near resonances in quasi-integrable Hamiltonian systems. J. Stat. Phys. **44**, 293 (1986)
5. Bernard, P., Kaloshin, V., Zhang, Ke.: Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders. Acta Math. **217**, 1 (2016)
6. Canalias, E., Delshams, A., Masdemont, J., Roldán González, P.: The scattering map in the planar restricted three body problem. Cel. Mech. Dyn. Astron **95**, 155 (2006)
7. Capinski, M.J., Zgliczyski, P.: Transition tori in the planar restricted elliptic three-body problem. Nonlinearity **24**(5), 1395 (2011)
8. Capinski, M.J., Gidea, M., de la Llave, R.: Arnold diffusion in the planar elliptic restricted three-body problem: mechanics and numerical verification. Nonlinearity **30**(1), 329 (2017)
9. Celletti, A., Giorgilli, A.: On the stability of the Lagrangian points in the spatial restricted problem of three bodies. Cel. Mech. Dyn. Astron. **50**, 31 (1991)
10. Cheng, C.Q., Yan, J.: Existence of diffusion orbits in a priori unstable Hamiltonian systems. J. Diff. Geom. **67**(3), 457 (2004)
11. Chierchia, L., Gallavotti, G.: Drift and diffusion in phase space. Annales de l'I. H. P. A **60**(1), 1 (1994)
12. Chierchia, L., Faraggiana, M.A., Guzzo, M.: On steepness of 3-jet non degenerate functions. Annali di Matematica **198**(6), 2151 (2019)
13. Chirikov, B.V.: A universal instability of many-dimensional oscillator systems. Phys. Rep. **52**, 263 (1979)
14. Cincotta, P., Efthymiopoulos, C., Giordano, C., Mestre, M.: Chirikov and Nekhoroshev diffusion estimates: bridging the two sides of the river. Physica D **266**, 49 (2014)
15. Contopoulos, G.: A third integral of motion in a galaxy. Z. Astrophys. **49**, 273 (1960)
16. Contopoulos, G., Moutsoulas, M.: Resonance cases and small divisors in a third integral of motion. Astron. J. **70**, 817 (1965)
17. Contopoulos, G.: Order and Chaos in Dynamical Astronomy. Springer, Berlin (2002)
18. De Blasi, I., Celletti, A., Efthymiopoulos, C.: Semi-analytical estimates for the orbital stability of Earth's satellites (2021). arXiv:2101.0534
19. Delshams, A., de la Llave, R., Seara, T.M.: A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. Mem. Amer. Math. Soc. **179**(844), viii+141 (2006)
20. Delshams, A., Gidea, M., de la Llave, R., Seara, T.M.: Geometric approaches to the problem of instability in Hamiltonian systems. An informal presentation. In: Craig, W. (ed), Hamiltonian Dynamical Systems and Applications, NATO SPSS-B Series, p 285 (2008)
21. Delshams, A., de la Llave, R., Seara, T.M.: Geometric properties of the scattering map of a normally hyperbolic invariant manifold. Adv. Math. **217**(3), 1096 (2008)
22. Delshams, A., De la Llave, R., Seara, T.M.: Instability of high dimensional Hamiltonian systems: multiple resonances do not impede diffusion. Adv. Math. **294**, 689 (2016)

23. Efthymiopoulos, C., Sándor, Z.: Optimized Nekhoroshev stability estimates for the Trojan asteroids with a symplectic mapping model of co-orbital motion. Mon. Not. R. Astron. Soc. **364**, 253 (2005)
24. Efthymiopoulos, C.: On the connection between the Nekhoroshev theorem and arnold diffusion cel. Mech. Dyn. Astron. **102**, 49 (2008)
25. Efthymiopoulos, C.: Canonical perturbation theory, stability and diffusion in Hamiltonian systems: applications in dynamical astronomy. In: Cincotta, P., Giordano, C., Efthymiopoulos, C. (eds.), 3rd La Plata International School on Astronomy and Geophysics "Chaos, Diffusion and Non-integrability in Hamiltonian Systems - Aplications to Astronomy, 1st edn, p 1. Uni. Nac. de la Plata, La Plata (2012)
26. Efthymiopoulos, C., Harsoula, M.: The speed of Arnold diffusion. Physica D **251**, 19 (2013)
27. Féjoz, J., Guàrdia, M., Kaloshin, V., PRoldán, P.: Kirkwood gaps and diffusion along mean motion resonances in the restricted planar three-body problem. J. Eur. Math. Soc. **18**, 2313 (2016)
28. Froeschlé, C., Lega, E., Gonczi, R.: Fast Lyapunov Indicators. Application to Asteroidal Motions. Cel. Mech. Dyn. Astron. **67**, 41 (1997)
29. Froeschlé, C., Guzzo, M., Lega, E.: Graphical evolution of the Arnold web: from order to chaos. Science **289**(5487), 2108 (2000)
30. Froeschlé, C., Guzzo, M., Lega, E.: Local and global diffusion along resonant lines in discrete quasi-integrable dynamical systems. Cel. Mech. Dyn. Astron. **92**, 243 (2005)
31. Gelfreich, V., Simó, C., Vieiro, A.: Dynamics of 4D symplectic maps near a double resonance. Physica D **243**, 92 (2013)
32. Gidea, M., de la Llave, R.: Topological methods in the instability problem of Hamiltonian systems. Discrete Contin. Dyn. Syst. **14**(2), 295 (2006)
33. Giorgilli, A., Skokos, C.: On the stability of the Trojan asteroids Astron. Astrophys. **317**, 254 (1997)
34. Giorgilli, A.: Notes on exponential stability of Hamiltonian systems, in Dynamical Systems. Part I: Hamiltonian Systems and Celestial Mechanics, 1st edn. Pubblicazioni della Classe di Scienze, Scuola Normale Superiore, Centro di Ricerca Matematica "Ennio De Giorgi", Pisa (2003)
35. Giorgilli, A.: A computer program for integrals of motion Comp. Phys. Commun. **16**(3), 331 (1979)
36. Gustavson, F.G.: On constructing formal integrals of a Hamiltonian system near ail equilibrium point. Astron. J. **71**, 670 (1966)
37. Guzzo, M., Lega, E., Froeschlé, C.: First numerical evidence of Arnold diffusion in quasi-integrable systems. Discr. Con. Dyn. Sys. B **5**, 687 (2005)
38. Guzzo, M., Lega, E., Froeschlé, C.: Diffusion and stability in perturbed non convex integrable systems. Nonlinearity **19**, 1049 (2006)
39. Guzzo, M., Lega, E., Froeschlé, C.: First numerical investigation of a conjecture by N. N. Nekhoroshev about stability in quasi-integrable systems. Chaos **21**, 033101 (2011)
40. Guzzo, M., Efthymiopoulos, C., Páez, R.I.: Semi-analytic computations of the speed of Arnold diffusion along single resonances in a priori stable Hamiltonian systems. J. Nonlin. Sci. **30**, 851 (2020)
41. Hénon, M., Heiles, C.: The applicability of the third integral of motion: some numerical experiments. Astron. J. **69**, 73 (1964)
42. Holmes, P.J., Marsden, J.E.: Melnikov's method and Arnold diffusion for perturbations of integrable Hamiltonian systems. J. Math. Phys. **23**, 669 (1982)
43. Kolmogorov, A.N.: Preservation of conditionally periodic movements with small change in the Hamiltonian function. Dokl. Akad. Nauk SSSR **98**, 527 (1954)
44. Lega, E., Guzzo, M., Froeschlé, C.: Detection of Arnold diffusion in Hamiltonian systems. Physica D **182**, 179 (2003)
45. Lega, E., Froeschlé, C., Guzzo, M.: Diffusion in Hamiltonian quasi-integrable systems. In: Benest, D., Froeschle, C., Lega, E. (eds.) Topics in Gravitational Dynamics. Lecture Notes in Physics, vol. 729. Springer, Berlin (2007)

46. Lichtenberg, A., Lieberman, M.: Regular and Chaotic Dynamics, 2nd edn. Springer, New York (1994)
47. Lhotka, Ch., Efthymiopoulos, C., Dvorak, R.: Nekhoroshev stability at L4 or L5 in the elliptic-restricted three-body problem - application to Trojan asteroids. Mon. Not. R. Astron. Soc. **384**, 1165 (2008)
48. Lochak, P.: Canonical perturbation theory via simultaneous approximation. Russ. Math. Surv. **47**, 57 (1992)
49. Lochak, P.: Arnold diffusion: a compendium of remarks and questions. In: Simó, C. (ed.) Hamiltonian Systems with Three or More Degrees of Freedom. NATO ASI Series (Series C: Mathematical and Physical Sciences), vol. 533, p. 168. Springer, Dordrecht (1999)
50. Mestre, M., Cincotta, P., Giordano, C.: Diffusion measurements in a 3DoF Hamiltonian flow. In: Cincotta, P., Giordano, C., Efthymiopoulos, C. (eds.), 3rd La Plata International School on Astronomy and Geophysics "Chaos, Diffusion and Non-integrability in Hamiltonian Systems - Aplications to Astronomy", 1st edn., p 319. Uni. Nac. de la Plata, La Plata (2012)
51. Moser, J.: On invariant curves of area-preserving mappings of an annulus. Nachr. Akad. Wiss. Gött. II Math. Phys. **KI**, 1 (1962)
52. Moeckel, R.: Transition tori in the five-body problem. J. Diff. Equ. **129**(2), 290 (1996)
53. Morbidelli, A., Giorgilli, A.: On the role of high order resonances in normal forms and in separatrix splitting. Physica D **102**(3–4), 195 (1997)
54. Morbidelli, A., Guzzo, M.: The Nekhoroshev thorem and the asteroid belt dynamical system. Celest. Mech. Dyn. Astron. **65**, 107 (1997)
55. Nekhoroshev, N.N.: An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems. Russ. Math. Surv. **32**(6), 1 (1977)
56. Poincaré, H.: Méthodes Nouvelles de la Mécanique Céleste. Gauthier-Villars, Paris (1892)
57. Pöshel, J.: Nekhoroshev estimates for quasi-convex Hamiltonian systems. Math. Z. **213**, 187 (1993)
58. Rasband, S.N.: Chaotic Dynamics of Nonlinear Systems. Wiley, New York (1990)
59. Sansottera, M., Lhotka, C., Lemaitre, A.: Effective stability around the Cassini state in the spin-orbit problem. Cel. Mech. Dyn. Astron. **119**(1), 75 (2014)
60. Schirinzi, G., Guzzo, M.: On the formulation of new explicit conditions for steepness from a former result of N.N. Nekhoroshev. J. Math. Phys. **54**, 072702 (2013)
61. Steichen, D., Giorgilli, A.: Long time stability for the main problem of artificial satellites. Cel. Mech. Dyn. Astron. **69**, 317 (1997)
62. Zhang, K.: Speed of Arnold diffusion for analytic Hamiltonian systems. Invent. Math. **186**, 255 (2011)

Orbit Determination with the Keplerian Integrals

Giovanni Federico Gronchi

dedicated to Prof. Andrea Milani

Abstract We review two initial orbit determination methods for too short arcs (TSAs) of optical observations of a solar system body. These methods employ the conservation laws of Kepler's problem, and allow to attempt the linkage of TSAs referring to quite far epochs, differing by even more than one orbital period of the observed object. The first method (`Link2`) concerns the linkage of 2 TSAs, and leads to a univariate polynomial equation of degree 9. An optimal property of this polynomial is proved using Gröbner bases theory. The second method (`Link3`) is thought for the linkage of 3 TSAs, and leads to a univariate polynomial equation of degree 8. A numerical test is shown for both algorithms.

Keywords Orbit determination · Algebraic methods

1 Introduction

Modern telescopes collect a very large number of optical observations of solar system bodies, that can be usually grouped in **very short arcs** (VSAs), see [10]. A VSA is a set

$$\{(\alpha_i, \delta_i), \quad i = 1 \dots m\}, \qquad m \geq 2$$

of pairs of values of *right ascension* and *declination* of the same celestial body, referring to epochs t_i, and covering a very short path in the sky. Usually the data

G. F. Gronchi (✉)
Dipartimento di Matematica, Università di Pisa, Largo B. Pontecorvo 5, 56127 Pisa, Italy
e-mail: giovanni.federico.gronchi@unipi.it

G. Baù et al. (eds.), *New Frontiers of Celestial Mechanics: Theory and Applications*,
Springer Proceedings in Mathematics & Statistics 399,
https://doi.org/10.1007/978-3-031-13115-8_6

contained in a VSA do not allow to compute a least squares orbit: in this case we speak of a too short arc (TSA). Given a TSA, we can compute an **attributable**

$$\mathcal{A} = (\alpha, \delta, \dot{\alpha}, \dot{\delta})$$

at the mean epoch $\bar{t} = \frac{1}{m}\sum_i t_i$ of the observations by a linear or quadratic fit, see [10]. Given an attributable, the radial distance ρ and the radial velocity $\dot{\rho}$ of the observed body remain completely unknown. However, given two attributables referring to the same celestial body, we can try to put them together with the aim of computing an orbit that fits all the data. This operation is called **linkage** in the orbit determination literature, and it is often challenging: an orbit produced by linking together two TSAs usually needs a confirmation with additional data to be considered reliable. Moreover, we cannot know a priori that two TSAs refer to the same observed body, and to perform an efficient selection of pairs of TSAs to be passed to a linkage algorithm is a critical issue.

In this paper we review two recent initial orbit determination methods, introduced in [4, 5], for the linkage of two or three TSAs. These are called `Link2` and `Link3`, respectively. Some interesting algebraic aspects of these algorithms are also discussed, and a numerical test is shown for both.

2 Linkage with the Keplerian Integrals

The first integrals of Kepler's motion can be used to write polynomial equations for the linkage of 2 TSAs. The conservation laws of angular momentum and energy were proposed for the linkage problem already in [11–13]: here the authors observed that the equations could be put in polynomial form but did not use this form. A polynomial formulation of the linkage problem was considered later in a series of papers [4, 7, 8]. In [7] the angular momentum and energy conservation laws are used, as in [12]: a polynomial is obtained by squaring twice the equation of the energy conservation. After elimination of variables we get a univariate equation of degree 48 in the radial distance ρ_2. In [8] the degree is reduced to 20 by using the Laplace-Lenz vector projected along a suitable direction in place of the energy. In [4] all the algebraic conservation laws are combined so that the degree is reduced to 9: this is the algorithm that we recall here.

Remark 1 Classical preliminary orbit determination methods, e.g. the ones by Gauss, Laplace, Mossotti [1, 3, 6, 9] use the equations of motion, and Taylor series expansions around a central time of the observational arc, thus the observations must necessarily be close enough in time. We observe that using conservation laws this constraint on the time is not required.

2.1 Kepler's Problem and Its First Integrals

The equation of motion of Kepler's problem is

$$\ddot{\mathbf{r}} = -\mu \frac{\mathbf{r}}{|\mathbf{r}|^3}, \tag{1}$$

where $\mathbf{r} \in \mathbb{R}^3$ is the unknown position vector and μ is a positive constant. The dynamics defined by (1) has the following conserved quantities:

$$\begin{aligned}
&\mathbf{c} = \mathbf{r} \times \dot{\mathbf{r}}, && \textit{angular momentum}\\
&\mathcal{E} = \frac{1}{2}|\dot{\mathbf{r}}|^2 - \frac{\mu}{|\mathbf{r}|}, && \textit{energy}\\
&\mathbf{L} = \frac{1}{\mu}\dot{\mathbf{r}} \times \mathbf{c} - \frac{\mathbf{r}}{|\mathbf{r}|}, && \textit{Laplace-Lenz vector.}
\end{aligned}$$

We call these quantities the **Keplerian integrals**. Since $\mathbf{c}$ and $\mathbf{L}$ have 3 components we get 7 scalar conserved quantities: among them only 5 are independent, in fact

$$\mathbf{c} \cdot \mathbf{L} = 0, \qquad 2|\mathbf{c}|^2\mathcal{E} + \mu^2(1 - |\mathbf{L}|^2) = 0.$$

Given an attributable $\mathcal{A}$ at the epoch $\bar{t}$, we write below the Keplerian integrals as functions of the unknown radial distance and velocity ρ, $\dot{\rho}$. We start by writing

$$\begin{aligned}
\mathbf{r} &= \mathbf{q} + \rho\,\mathbf{e}^\rho,\\
\dot{\mathbf{r}} &= \dot{\mathbf{q}} + \dot{\rho}\,\mathbf{e}^\rho + \rho(\dot{\alpha}\cos\delta\mathbf{e}^\alpha + \dot{\delta}\mathbf{e}^\delta),
\end{aligned}$$

where $\mathbf{q}$, $\dot{\mathbf{q}}$ are the position and velocity of the observer at time $\bar{t}$,

$$\mathbf{e}^\rho = (\cos\delta\cos\alpha, \cos\delta\sin\alpha, \sin\delta)$$

gives the line of sight, and

$$\mathbf{e}^\alpha = (\cos\delta)^{-1}\frac{\partial \mathbf{e}^\rho}{\partial\alpha}, \qquad \mathbf{e}^\delta = \frac{\partial \mathbf{e}^\rho}{\partial\delta}.$$

The angular momentum vector can be expressed as

$$\mathbf{c}(\rho, \dot{\rho}) = \mathbf{r} \times \dot{\mathbf{r}} = \mathbf{D}\dot{\rho} + \mathbf{E}\rho^2 + \mathbf{F}\rho + \mathbf{G}.$$

The vectors $\mathbf{D}$, $\mathbf{E}$, $\mathbf{F}$, $\mathbf{G}$ depend only on the attributable $\mathcal{A}$ and on $\mathbf{q}$, $\dot{\mathbf{q}}$:

$$\begin{aligned}
&\mathbf{D} = \mathbf{q} \times \mathbf{e}^{\rho},\\
&\mathbf{E} = \dot{\alpha}\cos\delta\mathbf{e}^{\rho} \times \mathbf{e}^{\alpha} + \dot{\delta}\mathbf{e}^{\rho} \times \mathbf{e}^{\delta} = \dot{\alpha}\cos\delta\mathbf{e}^{\delta} - \dot{\delta}\mathbf{e}^{\alpha},\\
&\mathbf{F} = \dot{\alpha}\cos\delta\mathbf{q} \times \mathbf{e}^{\alpha} + \dot{\delta}\mathbf{q} \times \mathbf{e}^{\delta} + \mathbf{e}^{\rho} \times \dot{\mathbf{q}},\\
&\mathbf{G} = \mathbf{q} \times \dot{\mathbf{q}}.
\end{aligned} \tag{2}$$

The energy can be written as

$$\mathcal{E} = \frac{1}{2}|\dot{\mathbf{r}}|^2 - \frac{\mu}{|\mathbf{r}|},$$

where

$$|\mathbf{r}| = \sqrt{\rho^2 + 2(\mathbf{q}\cdot\mathbf{e}^{\rho})\rho + |\mathbf{q}|^2},$$

and

$$|\dot{\mathbf{r}}|^2 = \dot{\rho}^2 + (\dot{\alpha}^2\cos^2\delta + \dot{\delta}^2)\rho^2 + 2\dot{\mathbf{q}}\cdot\mathbf{e}^{\rho}\dot{\rho} + 2\dot{\mathbf{q}}\cdot(\dot{\alpha}\cos\delta\mathbf{e}^{\alpha} + \dot{\delta}\mathbf{e}^{\delta})\rho + |\dot{\mathbf{q}}|^2.$$

Finally, the Laplace-Lenz vector $\mathbf{L}$ is given by

$$\mu\mathbf{L}(\rho,\dot{\rho}) = \Big(|\dot{\mathbf{r}}|^2 - \frac{\mu}{|\mathbf{r}|}\Big)\mathbf{r} - (\dot{\mathbf{r}}\cdot\mathbf{r})\dot{\mathbf{r}},$$

where

$$\dot{\mathbf{r}}\cdot\mathbf{r} = \rho\dot{\rho} + \mathbf{q}\cdot\mathbf{e}^{\rho}\dot{\rho} + (\dot{\mathbf{q}}\cdot\mathbf{e}^{\rho} + \mathbf{q}\cdot\mathbf{e}^{\alpha}\dot{\alpha}\cos\delta + \mathbf{q}\cdot\mathbf{e}^{\delta}\dot{\delta})\rho + \dot{\mathbf{q}}\cdot\mathbf{q}.$$

Remark 2 The expressions of $\mathcal{E}$ and $\mathbf{L}$ are algebraic but not polynomial, due to the presence of the term $\mu/|\mathbf{r}|$. If we consider the auxiliary variable z defined by relation

$$|\mathbf{r}|z = \mu, \tag{3}$$

then the Keplerian integrals can be viewed as polynomials in the variables ρ, $\dot{\rho}$, z by writing z in place of $\mu/|\mathbf{r}|$. In this way, we obtain

$$\tilde{\mathcal{E}} = \frac{1}{2}|\dot{\mathbf{r}}|^2 - z, \qquad \mu\tilde{\mathbf{L}} = (|\dot{\mathbf{r}}|^2 - z)\mathbf{r} - (\dot{\mathbf{r}}\cdot\mathbf{r})\dot{\mathbf{r}}.$$

The relation between ρ and z can be taken into account through the polynomial equation

$$|\mathbf{r}|^2 z^2 = \mu^2. \tag{4}$$

Moreover, the following relations hold:

$$\mathbf{c}\cdot\tilde{\mathbf{L}} = 0, \qquad 2|\mathbf{c}|^2\tilde{\mathcal{E}} + \mu^2(1 - |\tilde{\mathbf{L}}|^2) = 0.$$

2.2 Polynomial Equations for the Linkage

Given two attributables $\mathcal{A}_1, \mathcal{A}_2$ at the epochs $\bar{t}_1, \bar{t}_2$, referring to the same solar system body, we consider the system

$$\mathbf{c}_1 = \mathbf{c}_2, \quad \mathbf{L}_1 = \mathbf{L}_2, \quad \mathcal{E}_1 = \mathcal{E}_2, \tag{5}$$

of 7 algebraic (but not polynomial) equations in the 4 unknowns $\rho_1, \rho_2, \dot{\rho}_1, \dot{\rho}_2$. System (5) depends on the vector of known parameters

$$(\mathcal{A}_1, \mathcal{A}_2, \mathbf{q}_1, \mathbf{q}_2, \dot{\mathbf{q}}_1, \dot{\mathbf{q}}_2),$$

and is overdetermined.

If we assume that the two-body dynamics is perfectly respected, and no error occurs in the coefficients, then the set of solutions of (5) in the complex field $\mathbb{C}$ (but also in $\mathbb{R}$) is not empty. More realistically, since these assumptions cannot hold exactly, system (5) is generically[1] inconsistent.

Combining the equations in (5) we can obtain an overdetermined polynomial system which is consistent and can be reduced by elimination to a univariate polynomial u of degree 9 in one of the radial distance, e.g. ρ_2, as will be shown below.

The conservation of the angular momentum $\mathbf{c}_1 = \mathbf{c}_2$ can be written as

$$\mathbf{D}_1\dot{\rho}_1 - \mathbf{D}_2\dot{\rho}_2 = \mathbf{J}(\rho_1, \rho_2), \tag{6}$$

where

$$\mathbf{J}(\rho_1, \rho_2) = \mathbf{E}_2\rho_2^2 - \mathbf{E}_1\rho_1^2 + \mathbf{F}_2\rho_2 - \mathbf{F}_1\rho_1 + \mathbf{G}_2 - \mathbf{G}_1, \tag{7}$$

and $\mathbf{D}_j, \mathbf{E}_j, \mathbf{F}_j, \mathbf{G}_j$ are given by relations (2) at times $\bar{t}_j$. Projecting Eqs. (6) onto the vectors

$$\mathbf{D}_1 \times \mathbf{D}_2, \quad \mathbf{D}_2 \times (\mathbf{D}_1 \times \mathbf{D}_2), \quad \mathbf{D}_1 \times (\mathbf{D}_1 \times \mathbf{D}_2),$$

where

$$\mathbf{D}_j = \mathbf{q}_j \times \mathbf{e}_j^{\rho},$$

we get

$$\mathbf{J}(\rho_1, \rho_2) \cdot (\mathbf{D}_1 \times \mathbf{D}_2) = 0,$$
$$|\mathbf{D}_1 \times \mathbf{D}_2|^2\dot{\rho}_1 - \mathbf{J}(\rho_1, \rho_2) \cdot \mathbf{D}_2 \times (\mathbf{D}_1 \times \mathbf{D}_2) = 0, \tag{8}$$
$$|\mathbf{D}_1 \times \mathbf{D}_2|^2\dot{\rho}_2 - \mathbf{J}(\rho_1, \rho_2) \cdot \mathbf{D}_1 \times (\mathbf{D}_1 \times \mathbf{D}_2) = 0. \tag{9}$$

We set

[1] i.e. such property can not be violated in a non-empty open subset of the data set $\mathbf{q}_j, \dot{\mathbf{q}}_j, \mathcal{A}_j$, $j = 1, 2$.

$$q(\rho_1, \rho_2) = \mathbf{J}(\rho_1, \rho_2) \cdot (\mathbf{D}_1 \times \mathbf{D}_2).$$

This is a quadratic polynomial, that can be written as

$$q(\rho_1, \rho_2) = q_{2,0}\rho_1^2 + q_{1,0}\rho_1 + q_{0,2}\rho_2^2 + q_{0,1}\rho_2 + q_{0,0}, \tag{10}$$

with

$$\begin{aligned} q_{2,0} &= -\mathbf{E}_1 \cdot \mathbf{D}_1 \times \mathbf{D}_2, & q_{0,2} &= \mathbf{E}_2 \cdot \mathbf{D}_1 \times \mathbf{D}_2, \\ q_{1,0} &= -\mathbf{F}_1 \cdot \mathbf{D}_1 \times \mathbf{D}_2, & q_{0,1} &= \mathbf{F}_2 \cdot \mathbf{D}_1 \times \mathbf{D}_2, \end{aligned}$$

$$q_{0,0} = (\mathbf{G}_2 - \mathbf{G}_1) \cdot \mathbf{D}_1 \times \mathbf{D}_2.$$

Remark 3 Using Eqs. (8), (9) we can write $\dot\rho_1, \dot\rho_2$ as quadratic polynomials in the variables ρ_1, ρ_2. This corresponds to using conservation of angular momentum in the plane orthogonal to $\mathbf{D}_1 \times \mathbf{D}_2$.

The equations

$$\mathbf{L}_1 = \mathbf{L}_2, \qquad \mathcal{E}_1 = \mathcal{E}_2 \tag{11}$$

are algebraic but not polynomial, due to the terms $\mu/|\mathbf{r}_j|$. We consider the equation

$$\boldsymbol{\xi} = \mathbf{0}, \tag{12}$$

with

$$\begin{aligned} \boldsymbol{\xi} &= [\mu(\mathbf{L}_1 - \mathbf{L}_2) - (\mathcal{E}_1\mathbf{r}_1 - \mathcal{E}_2\mathbf{r}_2)] \times (\mathbf{r}_1 - \mathbf{r}_2) \qquad (13) \\ &= \frac{1}{2}(|\dot{\mathbf{r}}_2|^2 - |\dot{\mathbf{r}}_1|^2)\mathbf{r}_1 \times \mathbf{r}_2 - (\dot{\mathbf{r}}_1 \cdot \mathbf{r}_1)\dot{\mathbf{r}}_1 \times (\mathbf{r}_1 - \mathbf{r}_2) + (\dot{\mathbf{r}}_2 \cdot \mathbf{r}_2)\dot{\mathbf{r}}_2 \times (\mathbf{r}_1 - \mathbf{r}_2). \end{aligned}$$

Note that in Eq. (12), which is a consequence of (11), the dependence on $\mu/|\mathbf{r}_j|$ has been canceled.

After eliminating $\dot\rho_1, \dot\rho_2$ by (8), (9), $\boldsymbol{\xi}$ becomes a bivariate vector polynomial with total degree 6, that we still denote by $\boldsymbol{\xi}$. In the following, we consider the bivariate polynomial system

$$q = 0, \qquad \boldsymbol{\xi} = \mathbf{0}, \tag{14}$$

which is a consequence of (5).

Remark 4 The monomials of $\boldsymbol{\xi}$ with the highest degree are all multiplied by $\mathbf{e}_1^\rho \times \mathbf{e}_2^\rho$. Therefore, the two projections

$$p_1 = \boldsymbol{\xi} \cdot \mathbf{e}_1^\rho, \qquad p_2 = \boldsymbol{\xi} \cdot \mathbf{e}_2^\rho \tag{15}$$

lower the degree, and give two polynomials with total degree 5.

2.3 Consistency of Eqs. (14)

We sketch the proof of the following result: full details are given in [4].

Theorem 1 *For generic values of the data, the bivariate and overdetermined polynomial system*

$$q = 0, \qquad \boldsymbol{\xi} = \mathbf{0}$$

is consistent. Moreover, it can be reduced by elimination to a system of two univariate polynomials of degree 10,

$$\mathrm{u}_1 = \mathrm{u}_2 = 0, \tag{16}$$

such that

$$\mathrm{u} = \gcd(\mathrm{u}_1, \mathrm{u}_2)$$

has degree 9.

Sketch of the proof. Generically, relation

$$(\mathbf{q}_1 - \mathbf{q}_2) \cdot \mathbf{e}_1^\rho \times \mathbf{e}_2^\rho \neq 0$$

holds, so that, from

$$\boldsymbol{\xi} \cdot (\mathbf{r}_1 - \mathbf{r}_2) = 0 \qquad \text{and} \qquad \mathbf{r}_j = \mathbf{q}_j + \rho_j \mathbf{e}_j^\rho \quad (j = 1, 2),$$

we obtain

$$\boldsymbol{\xi} = \mathbf{0} \qquad \text{if and only if} \qquad \boldsymbol{\xi} \cdot \mathbf{e}_1^\rho = \boldsymbol{\xi} \cdot \mathbf{e}_2^\rho = 0.$$

Thus, in place of (14) we can consider the bivariate and still overdetermined system

$$q = p_1 = p_2 = 0. \tag{17}$$

We note that, if (ρ_1, ρ_2) fulfills $q = 0$, the vectors $\mathbf{r}_1$, $\mathbf{r}_2$, $\dot{\mathbf{r}}_1$, $\dot{\mathbf{r}}_2$ all lie in the same plane. This remark leads to the following geometrical fact:

Property 1 For (ρ_1, ρ_2) fulfilling $q = 0$ the vector $\boldsymbol{\xi}$ is parallel to the common value $\mathbf{c} = \mathbf{c}_1 = \mathbf{c}_2$ of the angular momentum.

Each projection

$$p_j = \boldsymbol{\xi} \cdot \mathbf{e}_j^\rho, \qquad j = 1, 2$$

vanishes either if $\boldsymbol{\xi} = \mathbf{0}$, or if $\boldsymbol{\xi}$ is orthogonal to $\mathbf{e}_j^\rho$. By Property 1, when $q = 0$ relation $\boldsymbol{\xi} \cdot \mathbf{e}_j^\rho = 0$ can be checked using the angular momentum in place of $\boldsymbol{\xi}$. For this purpose we introduce the projections

$$c_{ij} = \mathbf{c}_i \cdot \mathbf{e}_j^\rho, \qquad i, j = 1, 2.$$

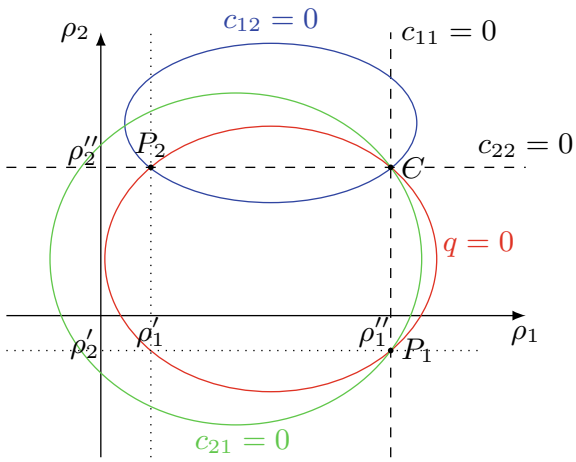


Fig. 1 Curves given by $q = 0, c_{ij} = 0$

The equations

$$c_{11}(\rho_1, \rho_2) = 0, \qquad c_{22}(\rho_1, \rho_2) = 0$$

define straight lines, while

$$c_{12}(\rho_1, \rho_2) = 0, \qquad c_{21}(\rho_1, \rho_2) = 0$$

define conic sections, see Fig. 1.

Set

$$P_1 = (\rho_1'', \rho_2'), \qquad P_2 = (\rho_1', \rho_2''), \qquad C = (\rho_1'', \rho_2''),$$

where

$$\rho_1' = \frac{\mathbf{q}_1 \times \mathbf{q}_2 \cdot \mathbf{e}_2^\rho}{\mathbf{e}_1^\rho \times \mathbf{e}_2^\rho \cdot \mathbf{q}_2}, \qquad \rho_2' = \frac{\mathbf{q}_1 \times \mathbf{q}_2 \cdot \mathbf{e}_1^\rho}{\mathbf{e}_1^\rho \times \mathbf{e}_2^\rho \cdot \mathbf{q}_1}, \tag{18}$$

$$\rho_1'' = \frac{\mathbf{q}_1 \times \dot{\mathbf{q}}_1 \cdot \mathbf{e}_1^\rho}{\mathbf{e}_1^\rho \times \mathbf{e}_1^\perp \cdot \mathbf{q}_1}, \qquad \rho_2'' = \frac{\mathbf{q}_2 \times \dot{\mathbf{q}}_2 \cdot \mathbf{e}_2^\rho}{\mathbf{e}_2^\rho \times \mathbf{e}_2^\perp \cdot \mathbf{q}_2}, \tag{19}$$

with

$$\mathbf{e}_j^\perp = \dot{\alpha}_j \cos\delta_j \mathbf{e}_j^\alpha + \dot{\delta}_j \mathbf{e}_j^\delta, \qquad j = 1, 2.$$

These points fulfill the relations

$$\begin{aligned} c_{11}(P_1) &= q(P_1) = 0, \\ c_{22}(P_2) &= q(P_2) = 0, \\ c_{11}(C) &= c_{22}(C) = q(C) = 0. \end{aligned}$$

We use the following results, that hold generically, see [4].

Lemma 1 *The point* $C = (\rho_1'', \rho_2'')$ *satisfies*

$$\mathbf{c}_1(C) = \mathbf{c}_2(C) = \mathbf{0}$$

and C *is the unique point in the plane* $\rho_1\rho_2$ *where both angular momenta vanish.*

Lemma 2 *In* C *we have* $\boldsymbol{\xi} \cdot \mathbf{e}_1^\rho \neq 0$ *and* $\boldsymbol{\xi} \cdot \mathbf{e}_2^\rho \neq 0$.

Lemma 3 *Assume* $q = 0$. *Then* $\boldsymbol{\xi} = \mathbf{0}$ *is equivalent to*

$$\begin{cases} \boldsymbol{\xi} \cdot \mathbf{e}_1^\rho = 0 \\ \mathbf{c} \cdot \mathbf{e}_1^\rho \neq 0 \end{cases} \quad or \quad \begin{cases} \boldsymbol{\xi} \cdot \mathbf{e}_2^\rho = 0 \\ \mathbf{c} \cdot \mathbf{e}_2^\rho \neq 0 \end{cases} . \tag{20}$$

Using these lemmas, first we show that system (17) has *at least* 9 solutions (in the complex field $\mathbb{C}$). By Lemma 3, system (17) is generically equivalent to

$$\begin{cases} q = p_1 = 0 \\ c_{11} \neq 0 \end{cases} \quad \text{or} \quad \begin{cases} q = p_2 = 0 \\ c_{22} \neq 0 \end{cases} . \tag{21}$$

Both systems $q = p_1 = 0$ and $q = p_2 = 0$ generically define 10 points in $\mathbb{C}^2$. Moreover, for $q = 0$, relation $c_{11} \neq 0$ discards the points P_1, C, while relation $c_{22} \neq 0$ discards P_2, C. In any case, by Lemma 2, C generically neither belongs to the curve $p_1 = 0$, nor to the curve $p_2 = 0$.

On the other hand, we can prove that

$$p_1(P_1) = p_2(P_2) = 0.$$

Let us show only that

$$p_1(P_1) = 0,$$

the proof of $p_2(P_2) = 0$ being similar. If $\boldsymbol{\xi}(P_1) = \mathbf{0}$, then the result holds trivially. Assume $\boldsymbol{\xi}(P_1) \neq \mathbf{0}$. We have $q(P_1) = 0$, therefore $\mathbf{c}_1(P_1) = \mathbf{c}_2(P_1) =: \mathbf{c}(P_1)$. Since generically $P_1 \neq C$, by Lemma 1 we have $\mathbf{c}(P_1) \neq \mathbf{0}$, and $\mathbf{c}(P_1)$ is parallel to $\boldsymbol{\xi}(P_1)$ by Property 1. From $c_{11}(P_1) = 0$ we conclude that $p_1(P_1) = 0$ because c_{11}, p_1 are the projections of $\mathbf{c}$, $\boldsymbol{\xi}$ onto $\mathbf{e}_1^\rho$.

Then, we are left with 9 solutions for both systems in (21), implying a lower bound of 9 solutions for (14).

Now we show that (17) has exactly 9 solutions. By Bezout's theorem we know that it has at most 10 solutions, because both systems $p_1 = q = 0$ and $p_2 = q = 0$ have 10 solutions each. Moreover, generically we have

$$p_1(P_2) \neq 0, \qquad p_2(P_1) \neq 0. \tag{22}$$

Using

$$p_1(P_1) = q(P_1) = 0, \qquad p_2(P_2) = q(P_2) = 0$$

and the lower bound above, we conclude that (17) has generically 9 solutions, and the two systems in (21) share the same solutions.

Consider the univariate polynomials

$$\mathfrak{u}_1 = \mathrm{res}(p_1, q, \rho_1), \qquad \mathfrak{u}_2 = \mathrm{res}(p_2, q, \rho_1)$$

given by the resultant of the pairs (p_j, q) with respect to ρ_1, see [2]. The quantities ρ_2' and ρ_2'' are roots of $\mathfrak{u}_1(\rho_2)$ and $\mathfrak{u}_2(\rho_2)$ respectively, because they are the ρ_2 components of P_1 and P_2. Generically we have

$$\mathfrak{u}_1(\rho_2'') \neq 0, \qquad \mathfrak{u}_2(\rho_2') \neq 0,$$

therefore ρ_2' and ρ_2'' do not solve (16). Then, we consider

$$\tilde{\mathfrak{u}}_1 = \frac{\mathfrak{u}_1}{\rho_2 - \rho_2'}, \qquad \tilde{\mathfrak{u}}_2 = \frac{\mathfrak{u}_2}{\rho_2 - \rho_2''}.$$

By the previous discussion we must have

$$\tilde{\mathfrak{u}}_1 = c\tilde{\mathfrak{u}}_2,$$

with c a non-zero constant, so that the univariate polynomial

$$\mathfrak{u} = \gcd(\mathfrak{u}_1, \mathfrak{u}_2) \tag{23}$$

has degree 9 and has the same roots as $\tilde{\mathfrak{u}}_1$ and $\tilde{\mathfrak{u}}_2$. This completes the proof of the theorem. □

3 An Optimal Property of the Polynomial $\mathfrak{u}$

In Remark 2 we observed that the Keplerian integrals can be viewed as polynomials in the variables $\rho, \dot{\rho}, z$ by writing z in place of $\mu/|\mathbf{r}|$. Therefore, we can consider the polynomial system

$$\mathbf{c}_1 = \mathbf{c}_2, \quad \mu\tilde{\mathbf{L}}_1 = \mu\tilde{\mathbf{L}}_2, \quad \tilde{\mathcal{E}}_1 = \tilde{\mathcal{E}}_2, \quad z_1^2|\mathbf{r}_1|^2 = \mu^2, \quad z_2^2|\mathbf{r}_2|^2 = \mu^2, \tag{24}$$

of 9 equations in the 6 unknowns

$$\rho_1, \rho_2, \dot{\rho}_1, \dot{\rho}_2, z_1, z_2.$$

In the next section we prove that dropping the last two equations in (24) we obtain a consistent polynomial system:

$$\mathbf{c}_1 = \mathbf{c}_2, \quad \mu\tilde{\mathbf{L}}_1 = \mu\tilde{\mathbf{L}}_2, \quad \tilde{\mathcal{E}}_1 = \tilde{\mathcal{E}}_2. \tag{25}$$

As a consequence of the proof we shall obtain that the univariate polynomial

$$\mathfrak{u} = \gcd(\mathfrak{u}_1, \mathfrak{u}_2)$$

of degree 9 has the *minimum degree* among the polynomials in ρ_1 or ρ_2 contained in the ideal

$$I = \langle \mathbf{c}_1 - \mathbf{c}_2,\ \mu(\tilde{\mathbf{L}}_1 - \tilde{\mathbf{L}}_2),\ \tilde{\mathcal{E}}_1 - \tilde{\mathcal{E}}_2 \rangle \subseteq \mathbb{R}[\rho_1, \rho_2, \dot{\rho}_1, \dot{\rho}_2, z_1, z_2].$$

3.1 A Gröbner Basis for the Ideal I

The following result holds true.

Theorem 2 *For generic data $\mathcal{A}_j, \mathbf{q}_j, \dot{\mathbf{q}}_j$, $j = 1, 2$, we can find a set of polynomials*

$$\{\mathfrak{g}_1, \ldots, \mathfrak{g}_6\} \subset \mathbb{R}[\rho_1, \rho_2, \dot{\rho}_1, \dot{\rho}_2, z_1, z_2]$$

that is a Gröbner basis of the ideal I for the lexicographic order with

$$\dot{\rho}_1 \succ \dot{\rho}_2 \succ z_1 \succ z_2 \succ \rho_1 \succ \rho_2, \tag{26}$$

and such that

$$\mathfrak{g}_6 = \mathfrak{u}.$$

We recall the following definition.

Definition 1 A set $\{\mathfrak{g}_1, \ldots, \mathfrak{g}_n\}$, with $n \in \mathbb{N}$, is a **Gröbner basis** of a polynomial ideal I for a fixed monomial order $\succ$ if and only if the leading term (for that order) of any element of I is divisible by the leading term of one $\mathfrak{g}_j$.

Proof For a generic choice of the data we consider the following set of generators of I:

$$
\begin{aligned}
\mathfrak{q}_1 &= (\mathbf{c}_1 - \mathbf{c}_2) \cdot \mathbf{D}_1 \times \mathbf{D}_2, \\
\mathfrak{q}_2 &= (\mathbf{c}_1 - \mathbf{c}_2) \cdot \mathbf{D}_2 \times (\mathbf{D}_1 \times \mathbf{D}_2), \\
\mathfrak{q}_3 &= (\mathbf{c}_1 - \mathbf{c}_2) \cdot \mathbf{D}_1 \times (\mathbf{D}_1 \times \mathbf{D}_2), \\
\mathfrak{q}_4 &= \mu(\tilde{\mathbf{L}}_1 - \tilde{\mathbf{L}}_2) \cdot \mathbf{e}_1^\rho \times \mathbf{e}_2^\rho, \\
\mathfrak{q}_5 &= \mu(\tilde{\mathbf{L}}_1 - \tilde{\mathbf{L}}_2) \cdot \mathbf{D}_2, \\
\mathfrak{q}_6 &= \mu(\tilde{\mathbf{L}}_1 - \tilde{\mathbf{L}}_2) \cdot \mathbf{D}_1, \\
\mathfrak{q}_7 &= \tilde{\mathcal{E}}_1 - \tilde{\mathcal{E}}_2.
\end{aligned}
$$

The first three polynomials have the form

$$
\begin{aligned}
\mathfrak{q}_1 &= q, \\
\mathfrak{q}_2 &= |\mathbf{D}_1 \times \mathbf{D}_2|^2 \dot{\rho}_1 - \mathbf{J} \cdot \mathbf{D}_2 \times (\mathbf{D}_1 \times \mathbf{D}_2), \\
\mathfrak{q}_3 &= |\mathbf{D}_1 \times \mathbf{D}_2|^2 \dot{\rho}_2 - \mathbf{J} \cdot \mathbf{D}_1 \times (\mathbf{D}_1 \times \mathbf{D}_2),
\end{aligned}
$$

where q and $\mathbf{J}$ are defined in (10) and (7). The other generators of I can be written as

$$
\begin{aligned}
\mathfrak{q}_4 &= -(\mathbf{D}_1 \cdot \mathbf{e}_2^\rho) z_1 - (\mathbf{D}_2 \cdot \mathbf{e}_1^\rho) z_2 + \mathfrak{f}_4, \\
\mathfrak{q}_5 &= -(\mathbf{D}_2 \cdot \mathbf{r}_1) z_1 + \mathfrak{f}_5, \\
\mathfrak{q}_6 &= (\mathbf{D}_1 \cdot \mathbf{r}_2) z_2 + \mathfrak{f}_6, \\
\mathfrak{q}_7 &= -z_1 + z_2 + \mathfrak{f}_7,
\end{aligned}
$$

for some polynomials $\mathfrak{f}_j = \mathfrak{f}_j(\rho_1, \rho_2, \dot{\rho}_1, \dot{\rho}_2)$. We can substitute $\mathfrak{q}_4, \ldots, \mathfrak{q}_7$ with

$$
\begin{aligned}
\mathfrak{p}_4 &= -(\mathbf{D}_2 \cdot \mathbf{e}_1^\rho)\mathfrak{q}_7 - \mathfrak{q}_4 = A z_1 + \mathfrak{a}_1, \\
\mathfrak{p}_5 &= (\mathbf{D}_1 \cdot \mathbf{e}_2^\rho)\mathfrak{q}_7 - \mathfrak{q}_4 = A z_2 + \mathfrak{a}_2, \\
\mathfrak{p}_6 &= (\mathbf{D}_1 \cdot \mathbf{r}_2)\mathfrak{p}_5 - A\mathfrak{q}_6, \\
\mathfrak{p}_7 &= (\mathbf{D}_2 \cdot \mathbf{r}_1)\mathfrak{p}_4 + A\mathfrak{q}_5,
\end{aligned}
$$

where

$$
A = \mathbf{D}_1 \cdot \mathbf{e}_2^\rho + \mathbf{D}_2 \cdot \mathbf{e}_1^\rho = (\mathbf{q}_1 - \mathbf{q}_2) \cdot \mathbf{e}_1^\rho \times \mathbf{e}_2^\rho,
$$

for some polynomials $\mathfrak{a}_j = \mathfrak{a}_j(\rho_1, \rho_2, \dot{\rho}_1, \dot{\rho}_2)$. The monomials containing z_1, z_2 cancel out in $\mathfrak{p}_6, \mathfrak{p}_7$.

Using $\mathfrak{q}_2 = \mathfrak{q}_3 = 0$, we eliminate $\dot{\rho}_1, \dot{\rho}_2$ from $\mathfrak{p}_4, \ldots, \mathfrak{p}_7$: we call $\hat{\mathfrak{p}}_4, \ldots, \hat{\mathfrak{p}}_7$ the polynomials obtained in this way. It can be shown that

$$
\hat{\mathfrak{p}}_6 = -(\mathbf{D}_1 \cdot \mathbf{e}_2^\rho) p_1, \qquad \hat{\mathfrak{p}}_7 = (\mathbf{D}_2 \cdot \mathbf{e}_1^\rho) p_2, \tag{27}
$$

where p_1, p_2 are the bivariate polynomials defined in (15).

Therefore, the elimination ideal

$$J := I \cap \mathbb{R}[\rho_1, \rho_2]$$

is generated by q, p_1, p_2:

$$J = \langle q, p_1, p_2 \rangle.$$

Let us write

$$q(\rho_1, \rho_2) = \sum_{h=0}^{2} b_h(\rho_2)\rho_1^h,$$

with

$$b_0(\rho_2) = q_{0,2}\rho_2^2 + q_{0,1}\rho_2 + q_{0,0}, \qquad b_1 = q_{1,0}, \qquad b_2 = q_{2,0}.$$

Assuming $q_{2,0} \neq 0$, that generically holds, let us set

$$\begin{aligned} &\beta_1 = 1, \qquad \beta_2 = -\frac{b_1}{b_2}, \qquad \gamma_2 = -\frac{b_0}{b_2}, \\ &\beta_{h+1} = \beta_h\beta_2 + \gamma_h, \qquad \gamma_{h+1} = \beta_h\gamma_2, \qquad h = 2, 3, 4. \end{aligned} \tag{28}$$

Moreover, we introduce the polynomials

$$\eta_h(\rho_1) = \frac{1}{b_2}\sum_{j=0}^{h-1} \beta_{h-j}\rho_1^j, \qquad h = 1, \ldots, 4. \tag{29}$$

With this notation we have

$$\rho_1^{h+1} = \eta_h q + \beta_{h+1}\rho_1 + \gamma_{h+1}, \qquad h = 1, \ldots, 4. \tag{30}$$

The generators p_1, p_2 can be written as

$$p_1(\rho_1, \rho_2) = \sum_{h=0}^{4} a_{1,h}(\rho_2)\rho_1^h, \qquad p_2(\rho_1, \rho_2) = \sum_{h=0}^{5} a_{2,h}(\rho_2)\rho_1^h,$$

for some polynomials $a_{i,j}$, so that

$$\tilde{p}_1 = p_1 - q\sum_{j=1}^{3} a_{1,j+1}\eta_j, \qquad \tilde{p}_2 = p_2 - q\sum_{j=1}^{4} a_{2,j+1}\eta_j$$

belong to the ideal J and can be written as

$$\tilde{p}_1 = \tilde{a}_{1,1}(\rho_2)\rho_1 + \tilde{a}_{1,0}(\rho_2), \qquad \tilde{p}_2 = \tilde{a}_{2,1}(\rho_2)\rho_1 + \tilde{a}_{2,0}(\rho_2),$$

with

$$\tilde{a}_{1,1} = a_{1,1} + \sum_{h=2}^{4} a_{1,h}\beta_h, \qquad \tilde{a}_{1,0} = a_{1,0} + \sum_{h=2}^{4} a_{1,h}\gamma_h,$$
$$\tilde{a}_{2,1} = a_{2,1} + \sum_{h=2}^{5} a_{2,h}\beta_h, \qquad \tilde{a}_{2,0} = a_{2,0} + \sum_{h=2}^{5} a_{2,h}\gamma_h.$$

Then, we have

$$J = \langle q, \tilde{p}_1, \tilde{p}_2 \rangle.$$

Now we set

$$J_1 = \langle \tilde{p}_1, \tilde{p}_2 \rangle$$

and prove that

$$J = J_1,$$

that is, we can generate J with two polynomials only. First we show that

$$V(J_1) = V(J), \tag{31}$$

where the variety $V(K)$ of a polynomial ideal $K \subseteq \mathbb{R}[\rho_1, \rho_2]$ is the set

$$V(K) = \{(\rho_1, \rho_2) \in \mathbb{C}^2 : p(\rho_1, \rho_2) = 0, \ \forall p \in K\}.$$

From $J_1 \subseteq J$ we have

$$V(J_1) \supseteq V(J). \tag{32}$$

To prove the opposite inclusion, we introduce the univariate polynomial

$$\mathfrak{v} = \mathrm{res}(\tilde{p}_1, \tilde{p}_2, \rho_1) = \tilde{a}_{1,1}\tilde{a}_{2,0} - \tilde{a}_{1,0}\tilde{a}_{2,1} \tag{33}$$

in the variable ρ_2. It turns out that $\mathfrak{v}$ has degree 9. We need the following results, that hold for a generic choice of the data:

(i) $\mathfrak{u}$ and $\mathfrak{v}$, defined in (23) and (33) respectively, have 9 distinct solutions in $\mathbb{C}$ (i.e. they are *square-free*),

(ii) $\tilde{a}_{1,1}$ and $\tilde{a}_{2,1}$ are relatively prime, i.e.

$$\gcd(\tilde{a}_{1,1}, \tilde{a}_{2,1}) = 1. \tag{34}$$

The proof of these results is in [5]. By (34) we can find two univariate polynomials β, γ in the variable ρ_2 such that

$$\beta\tilde{a}_{1,1} + \gamma\tilde{a}_{2,1} = 1. \tag{35}$$

Let us introduce

$$\mathfrak{w} = \beta\tilde{p}_1 + \gamma\tilde{p}_2 = \rho_1 + \mathfrak{z}(\rho_2), \tag{36}$$

where

$$\mathfrak{z} = \beta\tilde{a}_{1,0} + \gamma\tilde{a}_{2,0}.$$

We show that

$$J_1 = \langle \mathfrak{w}, \mathfrak{v} \rangle. \tag{37}$$

In fact

$$\mathfrak{v} = \tilde{a}_{1,1}\tilde{p}_2 - \tilde{a}_{2,1}\tilde{p}_1, \tag{38}$$

because

$$\tilde{a}_{1,1}\tilde{p}_2 - \tilde{a}_{2,1}\tilde{p}_1 = \tilde{a}_{1,1}(\tilde{a}_{2,1}\rho_1 + \tilde{a}_{2,0}) - \tilde{a}_{2,1}(\tilde{a}_{1,1}\rho_1 + \tilde{a}_{1,0}) = \tilde{a}_{1,1}\tilde{a}_{2,0} - \tilde{a}_{2,1}\tilde{a}_{1,0}.$$

Relations (36), (38) show that $\mathfrak{w}, \mathfrak{v} \in J_1$. On the other hand, inverting these relations we also obtain

$$\tilde{p}_1 = \tilde{a}_{1,1}\mathfrak{w} - \gamma\mathfrak{v},$$
$$\tilde{p}_2 = \tilde{a}_{2,1}\mathfrak{w} + \beta\mathfrak{v},$$

that is $\tilde{p}_1, \tilde{p}_2$ belong to the ideal generated by $\mathfrak{w}, \mathfrak{v}$.

Property (37) implies that $V(J_1)$ has 9 distinct points. In fact, for each root ρ_2 of $\mathfrak{v}$, which are all distinct because $\mathfrak{v}$ is square-free, we find from $\mathfrak{w} = 0$ a unique value of ρ_1 such that $(\rho_1, \rho_2) \in V(J_1)$.

On the other hand, generically $V(J)$ has 9 distinct points too. This can be shown using Theorem 1 and the fact that also $\mathfrak{u}$ is square-free (see [5]). Then, from (32) we have[2]

$$V(J_1) = V(J). \tag{39}$$

In particular, the polynomials $\mathfrak{v}$ and $\mathfrak{u}$ coincide up to a non-zero constant factor c:

$$\mathfrak{v} = c\mathfrak{u},$$

because their (complex) roots have the same 9 values.

Now we prove that indeed the two ideals are the same:

$$J_1 = J. \tag{40}$$

We only need to show the inclusion $J \subseteq J_1$. Assume the lexicographic order with

$$\rho_1 \succ \rho_2$$

[2] Hint: the fact that the variety of two ideals is the same does not mean that the two ideals are necessarily the same, see Hilbert's *nullstellensatz* in [2].

for the monomials in J and take any polynomial h in J. Dividing by $\mathfrak{w} = \rho_1 + \mathfrak{z}(\rho_2)$ we obtain

$$h(\rho_1, \rho_2) = h_1(\rho_1, \rho_2)\mathfrak{w}(\rho_1, \rho_2) + \mathfrak{r}(\rho_2) \tag{41}$$

for some polynomials $h_1, \mathfrak{r}$. The remainder $\mathfrak{r}$ depends only on ρ_2 because of the particular form of $\mathfrak{w}$, whose leading term is ρ_1. From $\mathfrak{w} \in J_1 \subseteq J$ and (41) we have that $\mathfrak{r} \in J$, so that the roots of $\mathfrak{r}$ must contain all the ρ_2 coordinates of the points in $V(J)$.

Using the fact that $\mathfrak{v} = c\mathfrak{u}$ is square-free we obtain that $\mathfrak{v}$ must divide $\mathfrak{r}$, i.e. $\mathfrak{r} = d\mathfrak{v}$ for some polynomal $d(\rho_2)$, which together with (41) yields

$$h = h_1\mathfrak{w} + d\mathfrak{v} \in J_1.$$

We conclude that (40) holds.

The polynomials $\mathfrak{g}_1, \ldots, \mathfrak{g}_6$, with

$$\mathfrak{g}_1 = \mathfrak{q}_2, \quad \mathfrak{g}_2 = \mathfrak{q}_3, \quad \mathfrak{g}_3 = \hat{\mathfrak{p}}_4, \quad \mathfrak{g}_4 = \hat{\mathfrak{p}}_5, \quad \mathfrak{g}_5 = \mathfrak{w}, \quad \mathfrak{g}_6 = \mathfrak{u},$$

form a Gröbner basis of the ideal I for the lexicographic order (26). To show this, we can simply check that the leading monomials of each pair $(\mathfrak{g}_i, \mathfrak{g}_j)$, with $1 \leq i < j \leq 6$, are relatively prime. This concludes the proof of the theorem. □

Remark 5 The proof above yields a *normalized* Gröbner basis for the ideal J. In fact, we can rescale by constant factors the polynomials of the basis and consider

$$\begin{aligned}
\mathfrak{g}_1 &= \dot{\rho}_1 + \mathfrak{h}_1(\rho_1, \rho_2), \\
\mathfrak{g}_2 &= \dot{\rho}_2 + \mathfrak{h}_2(\rho_1, \rho_2), \\
\mathfrak{g}_3 &= z_1 + \mathfrak{h}_3(\rho_1, \rho_2), \\
\mathfrak{g}_4 &= z_2 + \mathfrak{h}_4(\rho_1, \rho_2), \\
\mathfrak{g}_5 &= \rho_1 + \mathfrak{z}(\rho_2), \\
\mathfrak{g}_6 &= \mathfrak{u}(\rho_2),
\end{aligned}$$

with

$$\mathfrak{h}_1 = \frac{\mathbf{J} \cdot \mathbf{D}_2 \times (\mathbf{D}_1 \times \mathbf{D}_2)}{|\mathbf{D}_1 \times \mathbf{D}_2|^2}, \quad \mathfrak{h}_2 = \frac{\mathbf{J} \cdot \mathbf{D}_1 \times (\mathbf{D}_1 \times \mathbf{D}_2)}{|\mathbf{D}_1 \times \mathbf{D}_2|^2}, \quad \mathfrak{h}_3 = \frac{\mathfrak{a}_1}{A}, \quad \mathfrak{h}_4 = \frac{\mathfrak{a}_2}{A}.$$

As a consequence of Theorem 2, we obtain

Corollary 1 *The polynomial* $\mathfrak{u}$ *has the minimum degree among the univariate polynomials in the variable* ρ_2 *belonging to the ideal* I.

3.2 Selecting the Solutions

Given $\mathbf{A} = (\mathcal{A}_1, \mathcal{A}_2)$ with covariance matrix

$$\Gamma_{\mathbf{A}} = \begin{bmatrix} \Gamma_{\mathcal{A}_1} & 0 \\ 0 & \Gamma_{\mathcal{A}_2} \end{bmatrix},$$

let

$$\mathbf{R} = \mathbf{R}(\mathbf{A}) = (\mathcal{R}_1(\mathbf{A}), \mathcal{R}_2(\mathbf{A})), \qquad \mathcal{R}_i = (\rho_i, \dot{\rho}_i), \quad i = 1, 2$$

be a solution of

$$\boldsymbol{\Phi}(\mathbf{R}; \mathbf{A}) = \begin{pmatrix} \mathbf{c}_1 - \mathbf{c}_2 \\ \boldsymbol{\xi} \cdot \mathbf{e}_1^{\rho} \end{pmatrix} = \mathbf{0}, \tag{42}$$

where $\boldsymbol{\xi}$ is defined in (13), and can also be written as

$$\boldsymbol{\xi} = [\mu(\mathbf{L}_1 - \mathbf{L}_2) - (\mathcal{E}_1 - \mathcal{E}_2)\mathbf{r}_1] \times (\mathbf{r}_1 - \mathbf{r}_2). \tag{43}$$

If both $(\mathcal{A}_1, \mathcal{R}_1(\mathbf{A}))$, $(\mathcal{A}_2, \mathcal{R}_2(\mathbf{A}))$ give bounded orbits at epochs

$$\tilde{t}_i = \tilde{t}_i(\mathbf{A}) = \bar{t}_i - \frac{\rho_i(\mathbf{A})}{c}, \qquad i = 1, 2, \tag{44}$$

where aberration of light with velocity c is taken into account, then we can compute the corresponding Keplerian elements. We introduce the vector

$$\boldsymbol{\Delta}_{a,\ell} = (\Delta a, \Delta \ell),$$

representing the difference in semimajor axis and mean anomaly of the two orbits, comparing the anomalies at the same time $\tilde{t}_1$:

$$\Delta a = a_1 - a_2, \qquad \Delta \ell = \ell_1 - \ell_2 - n(a_2)(\tilde{t}_1 - \tilde{t}_2),$$

where $n(a) = \sqrt{\mu} a^{-3/2}$ is the mean motion. We consider the map

$$(\mathcal{A}_1, \mathcal{A}_2) = \mathbf{A} \mapsto \boldsymbol{\Psi}(\mathbf{A}) = \left(\mathcal{A}_1, \mathcal{R}_1, \boldsymbol{\Delta}_{a,\ell}\right),$$

giving the orbit $(\mathcal{A}_1, \mathcal{R}_1(\mathbf{A}))$ in attributable coordinates at epoch $\tilde{t}_1$, together with the vector $\boldsymbol{\Delta}_{a,\ell}(\mathbf{A})$.

We map the covariance matrix $\Gamma_{\mathbf{A}}$ of $\mathbf{A}$ into the covariance matrix of $\boldsymbol{\Psi}(\mathbf{A})$ by

$$\Gamma_{\boldsymbol{\Psi}(\mathbf{A})} = \frac{\partial \boldsymbol{\Psi}}{\partial \mathbf{A}} \, \Gamma_{\mathbf{A}} \left[\frac{\partial \boldsymbol{\Psi}}{\partial \mathbf{A}}\right]^T.$$

We can consider different ways to select the solutions. Two of them are the following.

Compatibility Conditions. We check whether the considered solution of (42) fulfills the relation

$$\boldsymbol{\Delta}_{a,\ell} = \mathbf{0}$$

within a threshold defined by $\Gamma_{\mathbf{A}}$. More precisely, consider the marginal covariance matrix

$$\Gamma_{\boldsymbol{\Delta}_{a,\ell}} = \frac{\partial \boldsymbol{\Delta}_{a,\ell}}{\partial \mathbf{A}} \Gamma_{\mathbf{A}} \left[\frac{\partial \boldsymbol{\Delta}_{a,\ell}}{\partial \mathbf{A}} \right]^T$$

of the vector $\boldsymbol{\Delta}_{a,\ell}$. The inverse matrix

$$C^{\boldsymbol{\Delta}_{a,\ell}} = \Gamma^{-1}_{\boldsymbol{\Delta}_{a,\ell}}$$

defines a norm $\| \cdot \|_\star$ allowing to test the identification of $\mathcal{A}_1, \mathcal{A}_2$:

$$\| \boldsymbol{\Delta}_{a,\ell} \|^2_\star = \boldsymbol{\Delta}_{a,\ell} C^{\boldsymbol{\Delta}_{a,\ell}} \boldsymbol{\Delta}^T_{a,\ell} \le \chi^2_{max},$$

where χ_{max} is a control parameter, that needs to be seleceted on the basis of simulations and practical tests with real data.

The orbits computed with the method of Sect. 2 are such that

$$I_1 = I_2, \qquad \Omega_1 = \Omega_2, \qquad a_1(1-e_1^2) = a_2(1-e_2^2) \tag{45}$$

because they fulfill $\mathbf{c}_1 = \mathbf{c}_2$. Assuming $a_1 = a_2$ we get $e_1 = e_2$ from the third relation in (45). Since $a_1 = a_2$ corresponds to $\mathcal{E}_1 = \mathcal{E}_2$, from $\boldsymbol{\xi} = \mathbf{0}$ we also obtain

$$\mu(\mathbf{L}_1 - \mathbf{L}_2) \times (\mathbf{r}_1 - \mathbf{r}_2) = \mathbf{0}. \tag{46}$$

The vectors $\mathbf{L}_1, \mathbf{L}_2$ have the same size because $e_1 = e_2$. Since it is quite unlikely that these vector differences are parallel, generically relation (46) implies

$$\omega_1 = \omega_2.$$

Attribution. We can try to attribute the data of $\mathcal{A}_2$ to each considered solution $\mathbf{x}_1 = (\mathcal{A}_1, \mathcal{R}_1(\mathcal{A}))$ of (42), which has the covariance matrix

$$\Gamma_{\mathbf{x}_1} = \begin{bmatrix} \Gamma_{\mathcal{A}_1} & \Gamma_{\mathcal{A}_1,\mathcal{R}_1} \\ \Gamma_{\mathcal{R}_1,\mathcal{A}_1} & \Gamma_{\mathcal{R}_1} \end{bmatrix},$$

with

$$\Gamma_{\mathcal{A}_1} = \frac{\partial \mathcal{A}_1}{\partial \mathbf{A}} \Gamma_{\mathbf{A}} \left[\frac{\partial \mathcal{A}_1}{\partial \mathbf{A}} \right]^T, \qquad \Gamma_{\mathcal{R}_1} = \frac{\partial \mathcal{R}_1}{\partial \mathbf{A}} \Gamma_{\mathbf{A}} \left[\frac{\partial \mathcal{R}_1}{\partial \mathbf{A}} \right]^T,$$

$$\Gamma_{\mathcal{A}_1,\mathcal{R}_1} = \Gamma_{\mathcal{A}_1}\left[\frac{\partial \mathcal{R}_1}{\partial \mathcal{A}_1}\right]^T, \qquad \Gamma_{\mathcal{R}_1,\mathcal{A}_1} = \Gamma^T_{\mathcal{A}_1,\mathcal{R}_1}.$$

We recall here the attribution algorithm. Assume that we have

(i) a least squares orbit $\mathbf{x}_1$ obtained from m_1 observations, with mean epoch $\bar{t}_1$, with covariance and normal matrices $\Gamma_{\mathbf{x}_1}, C_{\mathbf{x}_1}$;
(ii) an attributable $\mathcal{A}_2$ obtained from m_2 observations, with mean epoch $\bar{t}_2$, with covariance and normal matrices $\Gamma_{\mathcal{A}_2}, C_{\mathcal{A}_2}$.

Assume that

$$\mathbf{x} \mapsto \mathcal{A} = G(\mathbf{x})$$

maps orbital elements to attributables and define the prediction function

$$F(\mathbf{x}; t_0, t) = G \circ \Phi^t_{t_0}(\mathbf{x}),$$

where $\Phi^t_{t_0}(\mathbf{x})$ is the integral flow of the Kepler problem. The covariance and normal matrices of $\mathcal{A}$ are given by

$$\Gamma_{\mathcal{A}} = \left[\frac{\partial F}{\partial \mathbf{x}}\right] \Gamma_{\mathbf{x}} \left[\frac{\partial F}{\partial \mathbf{x}}\right]^T, \qquad C_{\mathcal{A}} = \Gamma^{-1}_{\mathcal{A}},$$

where $\Gamma_{\mathbf{x}}$ is the covariance matrix of $\mathbf{x}$.

Let $\mathcal{A}_2$ be an attributable and C_2 its 4×4 normal matrix. Let $\mathcal{A}_p$ be the predicted attributable at time $\bar{t}_2$, computed from the least squares orbit $\mathbf{x}_1$, and Γ_p, C_p its covariance and normal matrices.

The formulae for linear attribution in the 4-D space are the following (see [10]):

$$\begin{aligned} C_0 &= C_2 + C_p, \qquad \Gamma_0 = C_0^{-1}, \\ \mathbf{x}_0 &= \Gamma_0 \left[C_2 \mathcal{A}_2 + C_p \mathcal{A}_p\right], \\ K_4 &= (\mathcal{A}_p - \mathcal{A}_2) \cdot [C_2 - C_2\, \Gamma_0\, C_2]\ (\mathcal{A}_p - \mathcal{A}_2). \end{aligned}$$

The values of the attribution penalty K_4/m, with $m = m_1 + m_2$, is used to filter out the pairs orbit-attributable which cannot belong to the same object.

3.3 *Numerical Test with* `Link2`

We show an application of the `Link2` algorithm using 4 observations of asteroid (4542) *Mossotti* made on April 28, 2011 and 4 observations of the same asteroid made on November 4, 2013. These data have been collected by the telescope Pan-STARRS1, mount Hakeakala, Hawaii, and are displayed in Table 1. For simplicity, only a few digits are reported here.

Table 1 Values of right ascension (α) and declination (δ) used for the linkage

α (rad)	δ (rad)	t (MJD)
4.127300	−0.094246	55679.51169
4.127261	−0.094238	55679.52398
4.127221	−0.094230	55679.53664
4.127188	−0.094223	55679.54709
0.896220	0.078635	56600.43378
0.896168	0.078626	56600.44773
0.896119	0.078617	56600.46130
0.896069	0.078608	56600.47489

Table 2 Pair of preliminary orbits computed with `Link2`. The epoch $\tilde{t}$ has been corrected by aberration, see (44) in Sect. 3.2. The angles I, Ω, ω, ℓ are given in degrees

a (au)	e	I	Ω	ω	ℓ	$\tilde{t}$ (MJD)
3.03055	0.06436	11.22246	104.80204	117.44122	5.63111	55679.51899
3.02287	0.04015	11.22246	104.80204	114.03999	188.86754	56600.44185

From these observations we computed the attributables

$$\mathcal{A}_1 = (4.127242, -0.094234, -0.00316982, \quad 0.00064761),$$
$$\mathcal{A}_2 = (0.896144, \quad 0.078622, -0.00364403, -0.00065882),$$

at the mean epochs $\bar{t}_1 = 55679.52985$ MJD, $\bar{t}_2 = 56600.45442$ MJD. In the attributables $\mathcal{A}_1, \mathcal{A}_2$ the angles α, δ are given in radians and the angular rates $\dot{\alpha}, \dot{\delta}$ are given in radians/day.

After discarding solutions with non-real or non-positive values of ρ, and unbounded solutions, we are left with the radial distance pair

$$(\rho_1, \rho_2) = (1.8802, 2.1774) \text{ au},$$

leading to the pair of preliminary orbits given in Table 2.

The intersection of the curves defined by $p_1 = p_2 = q = 0$ is shown in Fig. 2.

Then we computed the rms of the preliminary orbits in Table 2 with respect to a pure Keplerian motion and selected the first orbit as the best (the one with the least rms). We propagated this orbit at the mean epoch of the observations, which is $\bar{t} = 56139.99213$, applied differential corrections and computed a least squares orbit. This orbit is shown in Table 3, together with the known orbit at the same epoch.[3]

[3] Data from AstDyS-2 (https://newton.spacedys.com/astdys/), orbit propagation with the OrbFit software (http://adams.dm.unipi.it/orbfit/).

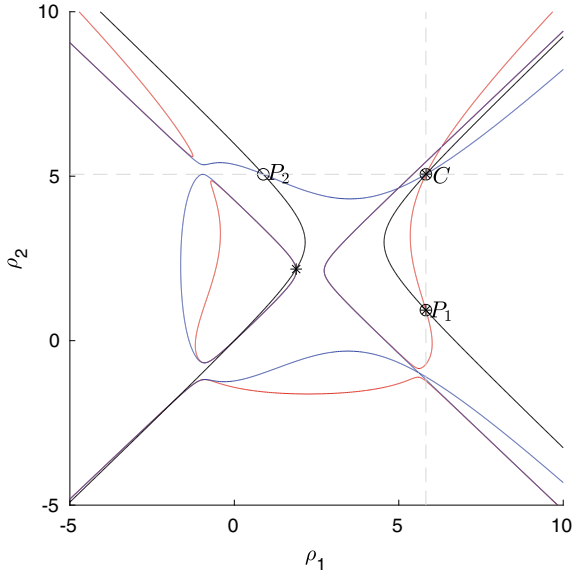


Fig. 2 Intersections of the curves $p_1 = 0$ (red), $p_2 = 0$ (blue), $q = 0$ (black) in the $\rho_1\rho_2$ plane

Table 3 Orbital elements of the least squares solution (LS) and of the known orbit. The angles I, Ω, ω, ℓ are given in degrees

	a (au)	e	I	Ω	ω	ℓ
LS	3.01802	0.05755	11.32849	104.37041	146.76038	66.54688
Known	3.00997	0.05614	11.30734	104.41991	144.01204	69.25283

4 Joining Three TSAs

Given three TSAs with attributables $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ at mean epochs $\bar{t}_1, \bar{t}_2, \bar{t}_3$, setting the conservation of angular momentum is enough to obtain a finite number of orbits. We review the following result, presented in [5]. Here the subscripts in $\mathbf{c}_i, \rho_i, \dot{\rho}_i, \mathbf{D}_i, \mathbf{E}_i, \mathbf{F}_i, \mathbf{G}_i$ refer to the three epochs.

Proposition 1 *Assume*

$$\mathbf{D}_1 \times \mathbf{D}_2 \cdot \mathbf{D}_3 \neq 0. \tag{47}$$

Then the polynomial system

$$(\mathbf{c}_1 - \mathbf{c}_2) \cdot \mathbf{D}_1 \times \mathbf{D}_2 = 0, \tag{48a}$$
$$(\mathbf{c}_1 - \mathbf{c}_2) \cdot \mathbf{D}_1 \times (\mathbf{D}_1 \times \mathbf{D}_2) = 0, \tag{48b}$$
$$(\mathbf{c}_2 - \mathbf{c}_3) \cdot \mathbf{D}_2 \times \mathbf{D}_3 = 0, \tag{48c}$$
$$(\mathbf{c}_2 - \mathbf{c}_3) \cdot \mathbf{D}_2 \times (\mathbf{D}_2 \times \mathbf{D}_3) = 0, \tag{48d}$$
$$(\mathbf{c}_3 - \mathbf{c}_1) \cdot \mathbf{D}_3 \times \mathbf{D}_1 = 0, \tag{48e}$$
$$(\mathbf{c}_3 - \mathbf{c}_1) \cdot \mathbf{D}_3 \times (\mathbf{D}_3 \times \mathbf{D}_1) = 0 \tag{48f}$$

in the 6 unknowns

$$\rho_1, \dot{\rho}_1, \rho_2, \dot{\rho}_2, \rho_3, \dot{\rho}_3$$

is equivalent to the redundant system

$$\mathbf{c}_1 = \mathbf{c}_2, \qquad \mathbf{c}_2 = \mathbf{c}_3, \qquad \mathbf{c}_3 = \mathbf{c}_1. \tag{49}$$

Proof System (49) trivially implies (48). Assume now that system (48) holds. Using relations (48e), (48f), to prove that $\mathbf{c}_3 = \mathbf{c}_1$ we only need to show that

$$(\mathbf{c}_3 - \mathbf{c}_1) \cdot \mathbf{v} = 0 \tag{50}$$

for some vector $\mathbf{v}$ that does not belong to the linear space generated by $\mathbf{D}_3 \times \mathbf{D}_1$ and $\mathbf{D}_3 \times (\mathbf{D}_3 \times \mathbf{D}_1)$. Indeed we show that we can choose

$$\mathbf{v} = \mathbf{D}_1 \times \mathbf{D}_2.$$

Note that

$$(\mathbf{D}_1 \times \mathbf{D}_2) \cdot (\mathbf{D}_2 \times \mathbf{D}_3) \times \big(\mathbf{D}_2 \times (\mathbf{D}_2 \times \mathbf{D}_3)\big) = 0,$$

that is, the vector $\mathbf{D}_1 \times \mathbf{D}_2$ belongs to the linear space generated by $\mathbf{D}_2 \times \mathbf{D}_3$ and $\mathbf{D}_2 \times (\mathbf{D}_2 \times \mathbf{D}_3)$. Moreover, $\mathbf{D}_1 \times \mathbf{D}_2$ is not generated by $\mathbf{D}_3 \times \mathbf{D}_1$ and $\mathbf{D}_3 \times (\mathbf{D}_3 \times \mathbf{D}_1)$, in fact by (47) we have

$$(\mathbf{D}_1 \times \mathbf{D}_2) \cdot (\mathbf{D}_3 \times \mathbf{D}_1) \times \big(\mathbf{D}_3 \times (\mathbf{D}_3 \times \mathbf{D}_1)\big) = |\mathbf{D}_3 \times \mathbf{D}_1|^2 \mathbf{D}_1 \times \mathbf{D}_2 \cdot \mathbf{D}_3 \neq 0.$$

Setting

$$\mathbf{v} = \mathbf{D}_1 \times \mathbf{D}_2,$$

from (48a), (48c), (48d) we obtain $(\mathbf{c}_1 - \mathbf{c}_2) \cdot \mathbf{v} = (\mathbf{c}_2 - \mathbf{c}_3) \cdot \mathbf{v} = 0$, which yield (50) and therefore we obtain $\mathbf{c}_3 = \mathbf{c}_1$. In a similar way we can prove that $\mathbf{c}_1 = \mathbf{c}_2, \mathbf{c}_2 = \mathbf{c}_3$, provided that system (48) holds.

□

Equations (49) can be written as

$$\begin{aligned}\mathbf{D}_1\dot{\rho}_1 - \mathbf{D}_2\dot{\rho}_2 &= \mathbf{J}_{12}(\rho_1, \rho_2),\\ \mathbf{D}_2\dot{\rho}_2 - \mathbf{D}_3\dot{\rho}_3 &= \mathbf{J}_{23}(\rho_2, \rho_3),\\ \mathbf{D}_3\dot{\rho}_3 - \mathbf{D}_1\dot{\rho}_1 &= \mathbf{J}_{31}(\rho_3, \rho_1),\end{aligned}$$

where

$$\begin{aligned}\mathbf{J}_{12}(\rho_1, \rho_2) &= \mathbf{E}_2\rho_2^2 - \mathbf{E}_1\rho_1^2 + \mathbf{F}_2\rho_2 - \mathbf{F}_1\rho_1 + \mathbf{G}_2 - \mathbf{G}_1,\\ \mathbf{J}_{23}(\rho_2, \rho_3) &= \mathbf{E}_3\rho_3^2 - \mathbf{E}_2\rho_2^2 + \mathbf{F}_3\rho_3 - \mathbf{F}_2\rho_2 + \mathbf{G}_3 - \mathbf{G}_2,\\ \mathbf{J}_{31}(\rho_3, \rho_1) &= \mathbf{E}_1\rho_1^2 - \mathbf{E}_3\rho_3^2 + \mathbf{F}_1\rho_1 - \mathbf{F}_3\rho_3 + \mathbf{G}_1 - \mathbf{G}_3.\end{aligned}$$

Equations (48a), (48c), (48e) depend only on the radial distances. In fact, they correspond to the system

$$\mathbf{J}_{12} \cdot \mathbf{D}_1 \times \mathbf{D}_2 = 0, \quad \mathbf{J}_{23} \cdot \mathbf{D}_2 \times \mathbf{D}_3 = 0, \quad \mathbf{J}_{31} \cdot \mathbf{D}_3 \times \mathbf{D}_1 = 0, \tag{51}$$

which can be written as

$$q_3 = a_3\rho_2^2 + b_3\rho_1^2 + c_3\rho_2 + d_3\rho_1 + e_3 = 0, \tag{52}$$
$$q_1 = a_1\rho_3^2 + b_1\rho_2^2 + c_1\rho_3 + d_1\rho_2 + e_1 = 0, \tag{53}$$
$$q_2 = a_2\rho_1^2 + b_2\rho_3^2 + c_2\rho_1 + d_2\rho_3 + e_2 = 0, \tag{54}$$

where

$$\begin{aligned}a_3 &= \mathbf{E}_2 \cdot \mathbf{D}_1 \times \mathbf{D}_2, \qquad b_3 = -\mathbf{E}_1 \cdot \mathbf{D}_1 \times \mathbf{D}_2,\\ c_3 &= \mathbf{F}_2 \cdot \mathbf{D}_1 \times \mathbf{D}_2, \qquad d_3 = -\mathbf{F}_1 \cdot \mathbf{D}_1 \times \mathbf{D}_2,\\ e_3 &= (\mathbf{G}_2 - \mathbf{G}_1) \cdot \mathbf{D}_1 \times \mathbf{D}_2,\end{aligned}$$

and the other coefficients a_j, b_j, c_j, d_j, e_j, for $j = 1, 2$, have similar expressions, obtained by cycling the indexes.

To eliminate ρ_1, ρ_3 from (51) we can first compute the resultant

$$r = \mathrm{res}(q_3, q_2, \rho_1),$$

which depends only on ρ_2, ρ_3, and then the resultant

$$\mathfrak{q} = \mathrm{res}(r, q_1, \rho_3),$$

which is a univariate polynomial of degree 8 in the variable ρ_2.

Therefore, provided that (47) holds, to get the solutions of (49) we search for the roots $\bar{\rho}_2$ of $\mathfrak{q}(\rho_2)$, compute the corresponding values $\bar{\rho}_3$ of ρ_3 from $r(\rho_3, \bar{\rho}_2) = q_1(\rho_3, \bar{\rho}_2) = 0$, and the values of ρ_1 from $q_3(\rho_1, \bar{\rho}_2) = q_2(\bar{\rho}_3, \rho_1) = 0$.

From Eqs. (48b), (48d), (48f) we can write the radial velocities $\dot\rho_j$ as functions of pairs of radial distances:

$$\dot\rho_2 = \frac{\mathbf{J}_{12}(\rho_1,\rho_2)\cdot\mathbf{D}_1\times(\mathbf{D}_1\times\mathbf{D}_2)}{|\mathbf{D}_1\times\mathbf{D}_2|^2},$$
$$\dot\rho_3 = \frac{\mathbf{J}_{23}(\rho_2,\rho_3)\cdot\mathbf{D}_2\times(\mathbf{D}_2\times\mathbf{D}_3)}{|\mathbf{D}_2\times\mathbf{D}_3|^2},$$
$$\dot\rho_1 = \frac{\mathbf{J}_{31}(\rho_3,\rho_1)\cdot\mathbf{D}_3\times(\mathbf{D}_3\times\mathbf{D}_1)}{|\mathbf{D}_3\times\mathbf{D}_1|^2}.$$

From these data we can reconstruct the orbital elements.

4.1 Straight Line Solutions

A particular solution of system (49) can be obtained by searching for values of ρ_j, $\dot\rho_j$ such that

$$\mathbf{c}_j(\rho_j,\dot\rho_j)=\mathbf{0},\qquad j=1,2,3.$$

Let us drop the index j. Relation $\mathbf{r}\times\dot{\mathbf{r}}=\mathbf{0}$ implies that there exists $\lambda\in\mathbb{R}$ such that

$$\dot\rho\mathbf{e}^\rho+\rho\boldsymbol{\eta}+\dot{\mathbf{q}}=\lambda(\rho\mathbf{e}^\rho+\mathbf{q}), \tag{55}$$

with $\boldsymbol{\eta}=\dot\alpha\cos\delta\mathbf{e}^\alpha+\dot\delta\mathbf{e}^\delta$. Setting $\sigma=\dot\rho-\lambda\rho$ we can write (55) as

$$\sigma\mathbf{e}^\rho+\rho\boldsymbol{\eta}-\lambda\mathbf{q}=-\dot{\mathbf{q}}. \tag{56}$$

We introduce the vector

$$\mathbf{u}=\mathbf{q}-(\mathbf{q}\cdot\mathbf{e}^\rho)\mathbf{e}^\rho-\frac{1}{\eta^2}(\mathbf{q}\cdot\boldsymbol{\eta})\boldsymbol{\eta},$$

which is orthogonal to both $\mathbf{e}^\rho$ and $\boldsymbol{\eta}$, where $\eta=|\boldsymbol{\eta}|$.

Thus, we can write (56) as

$$[\sigma-\lambda(\mathbf{q}\cdot\mathbf{e}^\rho)]\mathbf{e}^\rho+\left[\rho-\frac{\lambda}{\eta^2}(\mathbf{q}\cdot\boldsymbol{\eta})\right]\boldsymbol{\eta}-\lambda\mathbf{u}=-\dot{\mathbf{q}}.$$

Since $\{\mathbf{e}^\rho,\boldsymbol{\eta},\mathbf{u}\}$ is generically an orthogonal basis of $\mathbb{R}^3$, we find

$$\lambda=\frac{1}{|\mathbf{u}|^2}(\dot{\mathbf{q}}\cdot\mathbf{u}),\qquad \rho=\frac{1}{\eta^2}(\lambda\mathbf{q}-\dot{\mathbf{q}})\cdot\boldsymbol{\eta},\qquad \dot\rho=\lambda\rho+(\lambda\mathbf{q}-\dot{\mathbf{q}})\cdot\mathbf{e}^\rho.$$

In particular, we obtain the value

$$\rho = \frac{1}{\eta^2}\Big(\frac{1}{|\mathbf{u}|^2}(\dot{\mathbf{q}}\cdot\mathbf{u})(\mathbf{q}\cdot\boldsymbol{\eta}) - \dot{\mathbf{q}}\cdot\boldsymbol{\eta}\Big)$$

for the radial distance, corresponding to a solution with zero angular momentum.

4.2 Selecting the Solutions

Given $\mathbf{A} = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)$ with covariance matrices $\Gamma_{\mathcal{A}_1}, \Gamma_{\mathcal{A}_2}, \Gamma_{\mathcal{A}_3}$, let

$$\mathbf{R} = \mathbf{R}(\mathbf{A}) = \big(\mathcal{R}_1(\mathbf{A}), \mathcal{R}_2(\mathbf{A}), \mathcal{R}_3(\mathbf{A})\big), \qquad \mathcal{R}_i = (\rho_i, \dot{\rho}_i), \quad i = 1, 2, 3$$

be a solution of

$$\boldsymbol{\Phi}(\mathbf{R}; \mathbf{A}) = \mathbf{0}, \tag{57}$$

with

$$\boldsymbol{\Phi}(\mathbf{R}; \mathbf{A}) = \begin{pmatrix} (\mathbf{c}_1 - \mathbf{c}_2)\cdot\mathbf{D}_1 \times (\mathbf{D}_1 \times \mathbf{D}_2) \\ (\mathbf{c}_1 - \mathbf{c}_2)\cdot\mathbf{D}_1 \times \mathbf{D}_2 \\ (\mathbf{c}_2 - \mathbf{c}_3)\cdot\mathbf{D}_2 \times (\mathbf{D}_2 \times \mathbf{D}_3) \\ (\mathbf{c}_2 - \mathbf{c}_3)\cdot\mathbf{D}_2 \times \mathbf{D}_3 \\ (\mathbf{c}_3 - \mathbf{c}_1)\cdot\mathbf{D}_3 \times (\mathbf{D}_3 \times \mathbf{D}_1) \\ (\mathbf{c}_3 - \mathbf{c}_1)\cdot\mathbf{D}_3 \times \mathbf{D}_1 \end{pmatrix}.$$

If $(\mathcal{A}_1, \mathcal{R}_1(\mathcal{A}))$, $(\mathcal{A}_2, \mathcal{R}_2(\mathcal{A}))$, and $(\mathcal{A}_3, \mathcal{R}_3(\mathcal{A}))$ give bounded orbits at epochs

$$\tilde{t}_i = \bar{t}_i - \frac{\rho_i(\mathcal{A})}{c}, \qquad i = 1, 2, 3,$$

then we compute the corresponding Keplerian elements. We introduce the difference vectors

$$\begin{aligned} \boldsymbol{\Delta}_{12} &= \big(a_1 - a_2, \omega_1 - \omega_2, \ell_1 - \ell_2 - n(a_2)(\tilde{t}_1 - \tilde{t}_2)\big), \\ \boldsymbol{\Delta}_{32} &= \big(a_3 - a_2, \omega_3 - \omega_2, \ell_3 - \ell_2 - n(a_2)(\tilde{t}_3 - \tilde{t}_2)\big), \end{aligned}$$

where $n(a) = \sqrt{\mu}a^{-3/2}$ is the mean motion. We consider map

$$(\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3) = \mathbf{A} \mapsto \boldsymbol{\Psi}(\mathbf{A}) = (\mathcal{A}_2, \mathcal{R}_2, \boldsymbol{\Delta}_{12}, \boldsymbol{\Delta}_{32}),$$

giving the orbit $(\mathcal{A}_2, \mathcal{R}_2)$ in attributable coordinates at epoch $\tilde{t}_2$ together with the vectors $\boldsymbol{\Delta}_{12}$, $\boldsymbol{\Delta}_{32}$, which are not constrained by the angular momentum integrals.

We map the covariance matrix

$$\Gamma_{\mathbf{A}} = \begin{bmatrix} \Gamma_{\mathcal{A}_1} & 0 & 0 \\ 0 & \Gamma_{\mathcal{A}_2} & 0 \\ 0 & 0 & \Gamma_{\mathcal{A}_3} \end{bmatrix}$$

of $\mathbf{A}$ into the covariance matrix of $\boldsymbol{\Psi}(\mathbf{A})$ by the covariance propagation rule:

$$\Gamma_{\boldsymbol{\Psi}(\mathbf{A})} = \frac{\partial \boldsymbol{\Psi}}{\partial \mathbf{A}} \, \Gamma_{\mathbf{A}} \left[\frac{\partial \boldsymbol{\Psi}}{\partial \mathbf{A}} \right]^T,$$

We can check whether the considered solution of (57) fulfills the **compatibility conditions**

$$\boldsymbol{\Delta}_{12} = \boldsymbol{\Delta}_{32} = \mathbf{0}$$

within a threshold defined by $\Gamma_{\mathcal{A}}$. More precisely, consider the marginal covariance matrix $\Gamma_{\boldsymbol{\Delta}}$ of the vector

$$\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{12}, \boldsymbol{\Delta}_{32}).$$

The inverse matrix $C^{\boldsymbol{\Delta}} = \Gamma_{\boldsymbol{\Delta}}^{-1}$ defines a norm $\| \cdot \|_\star$ allowing us to test an identification between the attributables $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$: we check whether

$$\|\boldsymbol{\Delta}\|_\star^2 = \boldsymbol{\Delta} C^{\boldsymbol{\Delta}} \boldsymbol{\Delta}^T \leq \chi_{max}^2, \tag{58}$$

where χ_{max} is a control parameter.

4.3 *Numerical Test with* `Link3`

We show an application of the `Link3` algorithm using three TSAs of observations of asteroid (4628) *Laplace*, listed in Table 4. From these observations we computed the three attributables

$$\begin{aligned} \mathcal{A}_1 &= (5.497266, -0.067965, -0.00379969, -0.00072536), \\ \mathcal{A}_2 &= (0.715891, \quad 0.542071, -0.00422693, -0.00136864), \\ \mathcal{A}_3 &= (0.831367, \quad 0.390747, \quad 0.00622482, \quad 0.00054073), \end{aligned}$$

at the mean epochs $\bar{t}_1 = 55794.36935$, $\bar{t}_2 = 56226.53746$, $\bar{t}_3 = 56358.24760$, given in MJD. In the attributables $\mathcal{A}_j$ the angles are given in radians and the angular rates in radians/day. After discarding the straight-line solution, the solutions with negative values of ρ, and the unbounded ones, we are left with the radial distance triplets

Table 4 Values of right ascension (α) and declination (δ) of asteroid (4628) *Laplace* collected by the Pan-STARRS1 telescope

α (rad)	δ (rad)	t (MJD)
5.497381	−0.067942	55794.33902
5.497339	−0.067950	55794.35011
5.497195	−0.067978	55794.38807
5.497148	−0.067987	55794.40021
0.715965	0.542095	56226.52009
0.715918	0.542080	56226.53117
0.715867	0.542063	56226.54334
0.715816	0.542047	56226.55525
0.831317	0.390743	56358.23971
0.831350	0.390746	56358.24497
0.831383	0.390749	56358.25023
0.831416	0.390751	56358.25550

Table 5 Triplets of preliminary orbits computed with `Link3`. The angles I, Ω, ω, ℓ are given in degrees

	a (au)	e	I	Ω	ω	ℓ	$\tilde{t}$ (MJD)
1	2.86808	0.30942	12.13274	274.68641	172.31982	266.26844	55794.35667
	2.64520	0.13981	12.13274	274.68641	258.53770	242.07553	56226.52647
	2.59619	0.03219	12.13274	274.68641	290.50786	228.16130	56358.23074
2	2.64614	0.11646	11.78916	275.69255	249.45265	149.80066	55794.35816
	2.64562	0.11562	11.78916	275.69255	248.51598	249.78277	56226.52691
	2.64427	0.11343	11.78916	275.69255	247.58320	280.66987	56358.23093

$$(\rho_1, \rho_2, \rho_3) = (2.1955, 1.9028, 2.9200) \text{ au},$$
$$(\rho_1, \rho_2, \rho_3) = (1.9379, 1.8279, 2.8870) \text{ au},$$

leading to the triplets of preliminary orbits displayed in Table 5.

Based on the norm $\|\boldsymbol{\Delta}\|_\star$, we selected the second triplet. Checking the rms of these orbits with respect to a pure Keplerian motion we selected the first orbit of this triplet. We propagated this orbit at the mean epoch of the 12 observations in Table 4, which is $\bar{t} = 56126.38480$, applied differential corrections and computed a least squares orbit. This orbit is shown in Table 6, toghether with the known orbit at the same epoch.

Table 6 Orbital elements of the least squares solution (LS) and of the known orbit. The angles I, Ω, ω, ℓ are given in degrees

	a (au)	e	I	Ω	ω	ℓ
LS	2.64443	0.11729	11.79295	275.66956	248.45817	227.09536
Known	2.64441	0.11730	11.79294	275.66961	248.46069	227.09295

5 Conclusions and Future Work

We reviewed two initial orbit determination methods for TSAs of optical observations employing the conservation laws of Kepler's problem. Some algebraic properties of these algorithms have also been discussed and a simple test case has been presented for both. Being based on conservation laws, these methods are suitable to link TSAs quite far apart in time, even differing by more than one orbital period of the observed body. Moreover, these algorithms are very fast, because they are based on a polynomial formulation with low degree (9 for `Link2`, 8 for `Link3`). The sensitivity of these algorithms to astrometric errors is an important feature to be investigated: in fact it seems that some orbital elements are more sensitive to these errors. Moreover, it would be important to find efficient filters to discard a priori pairs of TSAs that are not likely to belong to the same observed object. Indeed, even if some filters have been proposed in [4, 7], a satisfactory solution to this problem is still missing. The mentioned problems are currently under investigation.

Acknowledgements The author wishes to thank Giulio Baù for carefully reading the manuscript, and for his useful suggestions. The author has been partially supported by the MSCA-ITN Stardust-R, Grant Agreement n. 813644 under the H2020 research and innovation program. He also acknowledges the project MIUR-PRIN 20178CJA2B "New frontiers of Celestial Mechanics: theory and applications" and the GNFM-INdAM (Gruppo Nazionale per la Fisica Matematica).

References

1. Celletti, A., Pinzari, G.: Four classical methods for determining planetary elliptic elements: a comparison. Celest. Mech. Dyn. Astron. **93**, 1–52 (2005)
2. Cox, D., Little, J., O'Shea, D.: Ideals, Varieties, and Algorithms. Springer (2005)
3. Gauss, C.F.: Theoria Motus Corporum in Sectionibus Conicis Solem Ambientium. Reprinted by Dover Publications in 1963 (1809)
4. Gronchi, G.F., Baù, G., Marò, S.: Orbit determination with the two-body integrals. III. Celest. Mech. Dyn. Astron. **123**(2), 105–122 (2015)
5. Gronchi, G.F., Baù, G., Milani, A.: Keplerian integrals, elimination theory and identification of very short arcs in a large database of optical observations. Celest. Mech. Dyn. Astron. **127**(2), 211–232 (2017)
6. Gronchi, G.F., Baù, G., Rodríguez, Ó., Jedicke, R., Moeyens, J.: A generalization of a method by Mossotti for initial orbit determination. Celest. Mech. Dyn. Astron. **133**, 41 (2021)
7. Gronchi, G.F., Dimare, L., Milani, A.: Orbit determination with the two-body integrals. Celest. Mech. Dyn. Astron. **107**(3), 299–318 (2010)

8. Gronchi, G.F., Farnocchia, D., Dimare, L.: Orbit determination with the two-body integrals. II. Celest. Mech. Dyn. Astron. **110**(3), 257–270 (2011)
9. Laplace, P.S.: Mém. Acad. R. Sci. Paris **10**, 93–146 (1780)
10. Milani, A., Gronchi, G.F.: Theory of Orbit Determination. Cambridge University Press (2010)
11. Taff, L.G.: On initial orbit determination. Astron. J. **89**(6), 1426–1428 (1984)
12. Taff, L.G., Hall, D.L.: The use of angles and angular rates. I—Initial orbit determination. Celest. Mech. **16**, 481–488 (1977)
13. Taff, L.G., Randall, P.M.S., Stansfield, S.A.: Angles-Only, Ground-Based. Initial Orbit Determination. Technical report, Lincoln Laboratory (1984)

Resonant Dynamics of Space Debris

Alessandra Celletti and Catalin Galeş

Abstract Since the launch of the Sputink 1 in 1957, a number of debris accumulated and now populates the circumterrestrial environment. These objects are found from Low Earth Orbits at altitudes of a few hundreds of kilometers to the Geostationary Earth Orbit, a region located at 42 164 km from the center of the Earth. The size of the debris runs from submillimeters to a few meters, but in view of their high velocities, they pose a serious threat for current and future satellites and are a source of hazard for human spaceflights. The sustainability of future space activities is definitely a priority for the current science and technology. In this context, it is of paramount importance to make a thorough study of the dynamics of space debris, especially in view of determining regular and chaotic stability properties. In this work we illustrate the models that can describe the dynamics at different altitudes; such models strongly depend on the location of the debris, since at low altitudes one needs to consider the Earth's atmosphere, while lunisolar effects become more important at higher altitudes. After having introduced the equations of motion both in Cartesian and Hamiltonian formalism, we analyze the dynamics of different resonances, most notably the geosynchronous and semi-synchronous resonances. We also review some results about the study of the orbits in the LEO region.

Keywords Resonances · Space debris · Chaos · Stability

A. Celletti
Department of Mathematics, University of Rome Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italy
e-mail: celletti@mat.uniroma2.it
URL: http://www.mat.uniroma2.it/celletti

C. Galeş (✉)
Department of Mathematics, Al. I. Cuza University of Iaşi, Blvd. Carol I, No. 11, 700506 Iaşi, Romania
e-mail: cgales@uaic.ro
URL: https://www.math.uaic.ro/cgales

G. Baù et al. (eds.), *New Frontiers of Celestial Mechanics: Theory and Applications*, Springer Proceedings in Mathematics & Statistics 399,
https://doi.org/10.1007/978-3-031-13115-8_7

1 Introduction

Space debris are artificial (non-operational or defunct) objects that orbit at different altitudes around the Earth. These objects are remnants of space missions, like rocket stages, old satellites, fragments from disintegrations, lost equipment, bolts, paint flakes, batteries, etc. The oldest space debris still in orbit is the satellite Vanguard 1. It was launched on an elliptic orbit on 17 March 1958 and has an orbital lifetime of about 240 years. As the number of launches continues to grow, much space junks have been produced in about 60 years of the exploration and exploitation of the circumterrestrial space. Nowadays, the Space Launch Report archive (see [94]) lists more than 6 000 launches since the Sputnik 1 on 4th October 1957, which deployed more than 10 600 spacecraft and satellites. The total mass of all space objects in Earth's orbit is estimated at 9 200 tonnes. Currently, about 6250 satellites are still in space and about 3600 are still functioning. A large portion of space debris are due to catastrophic events, either explosions or collisions, which have dramatically increased the number of objects in orbit.

ESA estimates that (as of October 2022) there are more than 36 500 debris objects with size larger than 10 cm in the Earth's orbit, about 28 000 objects being regularly tracked by the U.S. Space Surveillance Network (see [93]). Moreover, around the Earth there are about 1 million objects between 1 and 10 cm and about $1.3 \cdot 10^8$ objects between 1 mm and 1 cm. One has to bear in mind that in the LEO region a collision with a 10 cm object would entail a catastrophic fragmentation of a typical satellite, a collision with a 1 cm object would most likely disable a spacecraft, a collision with a 1 mm object could lead to destroy on board sub-systems.

Figure 1 shows the spatial density of objects as a function of the altitude, where the following terminology has been adopted:

- LEO stands for Low-Earth-Orbit ranging up to 2 000 km of altitude,
- MEO stands for Medium–Earth–Orbit with altitudes between 2 000 and 35 786 km,
- GEO is short for Geostationary–Earth–Orbit at an altitude close to 35 786 km.

Since the beginning of the space age, ESA estimates that more than 550 break-ups (explosions, collisions, or anomalous events resulting in fragmentations) have been occurred.

The first explosion in space happened on 29 June 1961: the Thor-Able upper stage exploded two hours after the separation from the Transit 4A navigation satellite; the number of fragments tracked from the ground was about 294. The first known accidental collision in space occurred on 24 July 1996: the 50 kg microsatellite Cerise was hit by the explosion of the rocket Ariane. Dramatic breakup events occurred in 2007 with the collision of the satellite Fengyun–1C with an anti-satellite missile; the collision generated more than 3 000 fragments. In 2009, the collision of the Cosmos 2251 and the Iridium 33 spacecraft generated more than 1 700 fragments. The proliferation of this population of space debris is a source of hazard for space assets and human operated missions. An analysis of the collision risk against space debris is described in [66, 82, 83].

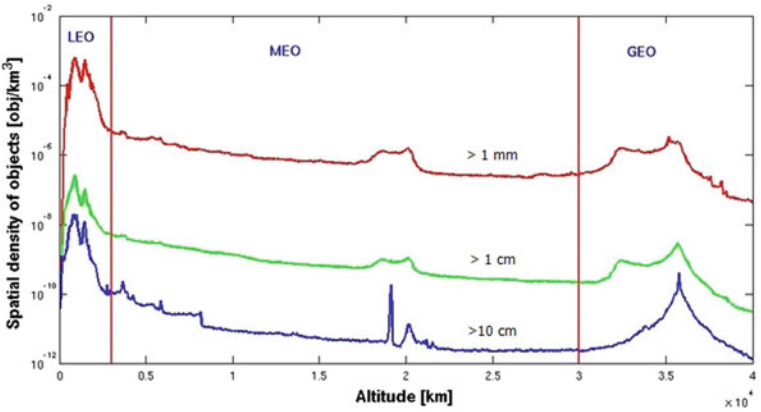


Fig. 1 Spatial density of objects as a function of the altitude for three different size thresholds: objects with diameter larger than 1 mm (red line), 1 cm (green line) and 10 cm (blue line) (from http://www.scholarpedia.org/article/Space debris)

The scientific community is aware that the so-called *Kessler syndrome* [64], according to which the density of objects is high enough that collisions between objects lead to a cascade effect, has already begun in LEO. In this region, the growth of space debris cannot be stopped, but only controlled. The Inter-Agency Space Debris Coordination Committee (IADC), founded in 1993, is an international forum of governmental bodies for the coordination of activities related to the issues of man-made and natural debris in space with the following aims: exchange information on space debris research activities between member space agencies, facilitate opportunities for co-operation in space debris research, review the progress of ongoing co-operative activities, identify debris mitigation options. The latter point calls in action the capability to understand the orbital evolution of space debris to devise maintenance, control strategies and mitigation. In particular, end-of-life disposal strategies consist in the transfer of still-operational spacecraft into graveyard orbits or rather could be obtained by provoking an eccentricity growth to cause an atmospheric re-entry.

With this in mind, it will be important to pursue the following aims:

– seek *stability regions*, to minimize the eccentricity growth and future interactions with operational spacecraft in that region;
– take advantage of the natural dynamics, to aim at re-enter into the atmosphere or to move into a graveyard orbit at the end of the operational life.

A successful strategy will require to understand the resonances, to locate regular dynamics, to study the mechanisms for the onset of chaos (see [11]). To this end, we need to distinguish between the different regions. Precisely, in MEO and GEO we have a conservative dynamics mainly affected by the monopole term of the Earth's attraction, Earth's oblateness, the attraction of the Moon, the influence of the Sun and the effect of Solar radiation pressure (hereafter SRP); in LEO we need to take

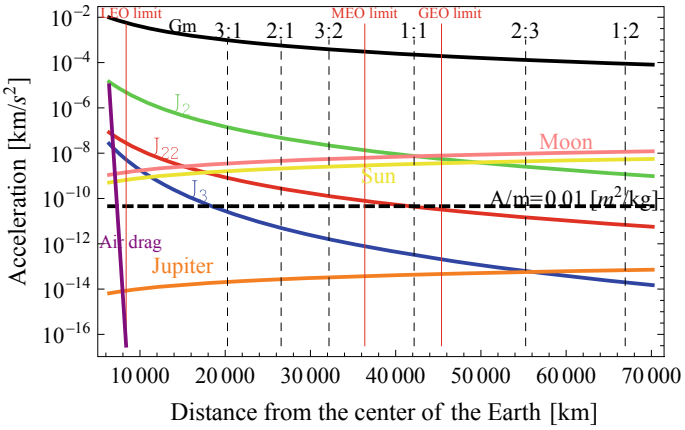


Fig. 2 Orders of magnitude of various perturbations of a debris orbit. The vertical dashed lines correspond to main tesseral resonances in MEO, GEO and outside the geostationary ring

into account the dissipative effects provoked by the atmospheric drag. The orders of magnitude of the different perturbations on a debris orbit are shown in Fig. 2.

We remark that recent results concerning the computation of proper elements have been obtained in [21] (see also [27, 77]), while perturbative methods and Nekhoroshev-like techniques have been implemented in [12, 29]. We also mention that the dynamics of space debris can be connected to the theory of normally hyperbolic invariant manifolds (NHIMs); a short suggestion is given in [18] and numerical results are available in [43].

A key role is played by resonances, that we classify as:

– tesseral and gravitational resonances, whenever there is a commensurability between the orbital period of the debris and the period of Earth's rotation (see [7–9, 13–15, 42, 70, 85, 88, 90, 91]);
– lunisolar resonances, which involve commensurabilities between the slow angles, i.e. the argument of perigee, the longitude of the ascending node or the mean anomaly of Moon and Sun (see [2, 8, 16, 18, 28, 54, 55, 76, 78]).

Another key role is played by the solar radiation pressure, which influences the orbits of high area-to-mass ratio objects (HAMR) (see [39, 89, 90]). We refer to [1, 10, 25, 30, 73, 79] for studies of the solar radiation pressure effects on the long–term evolution of space debris, to [34, 38, 53] for a description of the Earth's shadowing influence on circumterrestrial orbits and to [69] for secondary resonances involving the Sun's longitude.

In LEO one needs to include the dissipation due to the atmospheric drag (see [6, 40, 74]) and higher order harmonics (see [49]). The combined effect of air drag and resonances (tesseral or lunisolar) has been recently addressed in [3, 17, 86].

To stabilize the space debris population growth, IADC (see [57]) adopted a set of guidelines which are based on the following general principles:

– preventing on-orbit break-ups;
– removing spacecraft and orbital stages that have reached the end of their mission operations from densely populated orbit regions;
– limiting the objects released during normal operations.

Two protected regions about the Earth have been identified:

– Region A, the Low Earth Orbit (LEO) protected region,
– Region B, the Geosynchronous Protected Region, a segment of a spherical shell that includes the geosynchronous orbit (GEO) (see [57]).

For these regions, the mitigation measures proposed by IADC are able to reduce strongly the growth of the number of objects. In LEO, the orbital lifetime could vary from months to centuries. For an object with a large ballistic coefficient, say $B = 200\,\mathrm{cm}^2/\mathrm{kg}$, the orbital lifetime is 1 month at 300 km, 1 year at 400 km, while it amounts to decades at 600–700 km and centuries at altitudes larger than 900 km (see [58]). As for mitigation strategies in LEO, the space systems interfering with the LEO region should be de-orbited, or manoeuvred into an orbit with a reduced lifetime. The post-mission orbit lifetime is officially stated to be no longer than 25 years.

Another strategy for ensuring the LEO's resilience involves the implementation of debris active removal concepts (see [56]). Nowadays, comprehensive surveys, pursuing the research and development of technologies and techniques with the potential of removing either small or large debris, are strategic actions for the industry and space agencies. However, such concepts are not likely to be operational very soon. In the case of Region B, spacecraft at their end-of-life should be moved to a disposal orbit with the perigee higher than the geostationary altitude by an amount ΔH (km) given by

$$\Delta H = 235 + C_r\, 1000\, A/m\,,$$

where A/m is the area–to–mass ratio, while C_r is radiation pressure coefficient, typically between 1 and 2, which specifies the amount of solar radiation transmitted, absorbed and reflected by the spacecraft. The GEO and near-GEO regimes, including the orbital debris flux, the current operations, the end–of–life disposal of satellites, and transfer orbits, are described in [4, 36, 44, 60, 61, 67, 80, 87]. At present, there are no internationally agreed mitigation guidelines in MEO.

In the light of the panorama presented before, it is extremely important to address some questions on the dynamics of the space debris: where are the regular regions where the debris can be safely located? what is the effect of chaotic diffusion in the Earth's environment? which is the role of the different resonances that can occur due to Earth, Sun and Moon? In the following sections we will try to give some results by introducing different models and several kinds of resonances, providing some details mainly based on our past works on these subjects. In particular, in Sect. 2 we describe the equations of motion using the Cartesian and Hamiltonian approaches. In Sect. 3 we describe the resonances in MEO and GEO, while in Sect. 4 we give some details on the dynamics in LEO. Secular resonances are the content of Sect. 5. Conclusions and perspectives are presented in Sect. 6.

2 Equations of Motion and Resonances

In the following sections we will introduce the formulation of the problem using the Cartesian formalism (Sect. 2.1) and the Hamiltonian formalism (Sect. 2.2). The main term of the expansion of the geopotential, called the J_2 term, is introduced in Sect. 2.3. Once we have written the different contributions of the Hamiltonian, we analyze their expansion, which leads to the classification of the terms according to whether they contain fast or slow angles (Sect. 2.4). As a consequence, we will introduce different definitions of resonances. In Sect. 2.5 we introduce the effect of the atmospheric drag, which is effective at low altitudes.

2.1 Cartesian Equations of Motion

We consider a quasi-inertial reference frame with origin in the barycenter of the Earth and axes parallel to an inertial frame. Later, it will be convenient to introduce also a synodic reference frame with origin in the barycenter of our planet and rotating with the angular velocity of the Earth.

We consider the effects of the geopotential, the gravitational attractions of Sun and Moon, and the effect of solar radiation pressure. In the quasi-inertial frame, the equations of motion of a small object whose position is denoted by $\mathbf{r}$ is given by

$$\begin{aligned}\ddot{\mathbf{r}} = {} & R_3(-\theta)\,\nabla V(\mathbf{r}) - \mathcal{G}m_S\Big(\frac{\mathbf{r}-\mathbf{r}_S}{|\mathbf{r}-\mathbf{r}_S|^3}+\frac{\mathbf{r}_S}{|\mathbf{r}_S|^3}\Big)\\ & - \mathcal{G}m_M\Big(\frac{\mathbf{r}-\mathbf{r}_M}{|\mathbf{r}-\mathbf{r}_M|^3}+\frac{\mathbf{r}_M}{|\mathbf{r}_M|^3}\Big) + C_r\,P_r\,a_S^2\,(\frac{A}{m})\,\frac{\mathbf{r}-\mathbf{r}_S}{|\mathbf{r}-\mathbf{r}_S|^3} + \mathbf{a}_{ng}\,, \end{aligned} \tag{1}$$

where $\mathcal{G}$ is the gravitational constant, θ is the sidereal time, R_3 denotes a rotation about the vertical axis of angle $-\theta$, V is the geopotential, $\mathbf{r}_S$ and $\mathbf{r}_M$ are the position vectors of Sun and Moon; the last but one term describes the solar radiation pressure, which is due to the absorption/reflection of photons by the surface of the body, and it depends on the area-to-mass ratio A/m, the reflectivity coefficient C_r (fixed to 1 in simulations), the radiation pressure $P_r = 4.56 \cdot 10^{-6}\ [N/m^2]$ at $a_S = 1\ AU$; the last term $\mathbf{a}_{ng}$ stands for the acceleration induced by other non-gravitational effects.

In terms of spherical harmonics, the Earth's gravity potential can be expanded as [62, 63]

$$\begin{aligned} V(r,\phi,\lambda) &= -\sum_{n=0}^{\infty}\sum_{m=0}^{n} V_{nm} \\ &= \frac{\mathcal{G}M_E}{r}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\Big(\frac{R_E}{r}\Big)^n P_{nm}(\sin\phi)\,J_{nm}\cos(m\lambda-\lambda_{nm})\,, \end{aligned}$$

Table 1 The coefficients J_{nm} (in units of 10^{-6}) and the constants λ_{nm} up to degree and order 5; values computed from [33] (see also [22])

n	m	J_{nm}	λ_{nm}
2	0	1082.6261	0
2	1	0.001807	$-81.^{\circ}5116$
2	2	1.81559	$75.^{\circ}0715$
3	0	−2.53241	0
3	1	2.20947	$186.^{\circ}9692$
3	2	0.37445	$72.^{\circ}8111$
4	0	−1.619331	0
4	1	0.67864	$41.^{\circ}4529$
4	2	0.16759	$121.^{\circ}0589$
4	3	0.060421	$56.^{\circ}1784$

where (r, ϕ, λ) are the spherical coordinates in the Earth-centered synodic frame, M_E is the mass of the Earth, $r = |\mathbf{r}|$, P_{nm} are the Legendre polynomials, J_{nm} are the harmonic coefficients, λ_{nm} are constants.

It is important to stress that $J_2 = J_{20}$ is the largest coefficient in the expansion, since it is about 500 times larger than any other coefficient J_{nm}. Values of the coefficients J_{nm} and of the constants λ_{nm} are given in Table 1.

The sizes of the different contributions on an object orbiting around the Earth are given in Fig. 2, which compares different effects in terms of the distance from the center of the Earth. The Keplerian part is dominant at all distances. In LEO, the atmospheric drag is very important. The effects of J_2, Sun and Moon are all extremely relevant, but it is worth noticing that above a certain distance (around 40 000 km) the contribution of the Moon (and later of the Sun) becomes more important than the J_2 effect.

2.2 *Hamiltonian Equations of Motion*

To introduce the equations of motion using the Hamiltonian formalism, it is convenient to use the action–angle Delaunay variables denoted as $(L, G, H, M, \omega, \Omega)$. The actions are related to the orbital elements (a, e, i), standing for semimajor axis, eccentricity, inclination, by the following formulae:

$$L = \sqrt{\mu_E a}\,, \qquad G = L\sqrt{1 - e^2}\,, \qquad H = G\cos i\,.$$

The angle variables (M, ω, Ω) are, respectively, the mean anomaly, the argument of perigee and the longitude of the ascending node, while $\mu_E = \mathcal{G}\, M_E$. The equations of motion are given by

$$\dot{M} = \frac{\partial \mathcal{H}}{\partial L}\,, \qquad \dot{\omega} = \frac{\partial \mathcal{H}}{\partial G}\,, \qquad \dot{\Omega} = \frac{\partial \mathcal{H}}{\partial H}$$
$$\dot{L} = -\frac{\partial \mathcal{H}}{\partial M}\,, \qquad \dot{G} = -\frac{\partial \mathcal{H}}{\partial \omega}\,, \qquad \dot{H} = -\frac{\partial \mathcal{H}}{\partial \Omega}\,, \tag{2}$$

where the Hamiltonian can be written as the sum of different contributions:

$$\mathcal{H} = -\frac{\mu_E^2}{2L^2} + \mathcal{H}_{Earth}(a, e, i, M, \omega, \Omega, \theta) + \mathcal{H}_{Moon}(a, e, i, M, \omega, \Omega, \Lambda_M) + \mathcal{H}_{Sun}(a, e, i, M, \omega, \Omega, \Lambda_S) + \mathcal{H}_{SRP}(a, e, i, M, \omega, \Omega, \Lambda_S)\,, \tag{3}$$

with θ the sidereal time, $\Lambda_M = (a_M, e_M, i_M, M_M, \omega_M, \Omega_M)$ denotes the orbital elements of the Moon with respect to the ecliptic, $\Lambda_S = (a_S, e_S, i_S, M_S, \omega_S, \Omega_S)$ denotes the orbital elements of the Sun with respect to the celestial equator. We stress that the elements of the Moon are referred to ecliptic plane, since for this choice the inclination i_M of the Moon is nearly constant, while the rate of variation of the argument of perihelion ω_M of the Moon and that of the lunar longitude of the ascending node Ω_M are nearly linear, which is not the case of the Moon's elements with respect to the celestial equator as they vary nonlinearly in time. The Hamiltonian (3) is the sum of the Keplerian part, the effect of the geopotential $\mathcal{H}_{Earth}$, the contributions of Moon $\mathcal{H}_{Moon}$ and Sun $\mathcal{H}_{Sun}$, the effect of the solar radiation pressure $\mathcal{H}_{SRP}$.

The term $\mathcal{H}_{Earth}$ can be expanded in Fourier series as

$$\mathcal{H}_{Earth} = \frac{\mu_E}{a} \sum_{n=2}^{\infty} \sum_{m=0}^{n} \left(\frac{R_E}{a}\right)^n J_{nm} \sum_{p=0}^{n} F_{nmp}(i) \sum_{q=-\infty}^{\infty} G_{npq}(e)\, cs_{nm}\Big(\Psi_{nmpq}(M, \omega, \Omega, \theta)\Big)$$

where R_E is the radius of the Earth, J_{nm} are the harmonic coefficients, F_{nmp}, G_{npq} are the inclination and eccentricity functions [22, 63], cs_{nm} is the cosine function if $n - m$ is even and the sine function if $n - m$ is odd, and

$$\Psi_{nmpq}(M, \omega, \Omega, \theta) = (n - 2p)\omega + (n - 2p + q)M + m(\Omega - \theta) - m\lambda_{nm}$$

for some constants λ_{nm} as in Table 1.

The secular part of the geopotential, limited to the contributions of first order in J_2, J_3, J_4, can be written as

$$
\begin{aligned}
R_{earth}^{sec} \cong & \frac{\mu_E R_E^2 J_2}{a^3}\Big(\frac{3}{4}\sin^2 i - \frac{1}{2}\Big)(1-e^2)^{-3/2} \\
& + \frac{2\mu_E R_E^3 J_3}{a^4}\Big(\frac{15}{16}\sin^3 i - \frac{3}{4}\sin i\Big)e(1-e^2)^{-5/2}\sin\omega \\
& + \frac{\mu_E R_E^4 J_4}{a^5}\Big[\Big(-\frac{35}{32}\sin^4 i + \frac{15}{16}\sin^2 i\Big)\frac{3e^2}{2}(1-e^2)^{-7/2}\cos(2\omega) \\
& + \Big(\frac{105}{64}\sin^4 i - \frac{15}{8}\sin^2 i + \frac{3}{8}\Big)\Big(1+\frac{3e^2}{2}\Big)(1-e^2)^{-7/2}\Big]. \qquad (4)
\end{aligned}
$$

The disturbing functions due to the Sun, Moon and solar radiation pressure can be expanded in Fourier series as [20, 63, 68]:

$$
\mathcal{H}_b = \sum \mathcal{A}^{Sun/Moon/SRP}_{k_1k_2k_3k_4k_5k_6}(a,e,i,a_b,e_b,i_b)\cos(k_1M + k_2M_b + k_3\omega + k_4\omega_b + k_5\Omega + k_6\Omega_b),
$$

where k_j are integers and b stands for 'Sun', 'Moon' or 'SRP', and with the additional convention that whenever b is 'SRP' in the left hand side, then b is 'Sun' in the right hand side.

2.3 *Effects of J_2*

In this section we focus on the effect of J_2 [62, 63, 65]. We approximate $\mathcal{H}_{Earth}$ with the J_2 term only, again averaged over the fast variables, say

$$
\mathcal{H}_{Earth} \simeq \frac{R_E^2 J_2 \mu_E^4}{4}\,\frac{1}{L^3G^3}\left(1 - 3\frac{H^2}{G^2}\right),
$$

and then we consider the Hamiltonian given by the Keplerian part and the J_2 effect:

$$
\mathcal{H}_{Kepler+J_2} = -\frac{\mu_E^2}{2\,L^2} + \frac{R_E^2 J_2 \mu_E^4}{4}\,\frac{1}{L^3G^3}\left(1 - 3\frac{H^2}{G^2}\right). \qquad (5)
$$

Since $\mathcal{H}_{Kepler+J_2}$ depends only on the actions, then the quantities L, G and H (or equivalently a, e and i) are constant, while the Delaunay angles M, ω and Ω vary linearly in time with rates

$$
\begin{aligned}
\dot{M} &\simeq 6135.7\Big(\frac{R_E}{a}\Big)^{3/2} - 4.98\Big(\frac{R_E}{a}\Big)^{7/2}(1-e^2)^{-3/2}(1-3\cos^2 i)\ {}^o/day \\
\dot{\omega} &\simeq 4.98\Big(\frac{R_E}{a}\Big)^{7/2}(1-e^2)^{-2}(5\cos^2 i - 1)\ {}^o/day \\
\dot{\Omega} &\simeq -9.97\Big(\frac{R_E}{a}\Big)^{7/2}(1-e^2)^{-2}\cos i\ {}^o/day\,.
\end{aligned}
$$

As a consequence of the above formulae, we may summarize as follows the main effects of J_2: a slow change of the rate of the mean anomaly, a precession of the perigee and a secular regression of the orbital node. We also notice that, when we use the Hamiltonian (5), we obtain that

$$\dot{\omega} = \frac{\partial \mathcal{H}_{Kepler+J_2}}{\partial G} = -\frac{3}{4}\frac{R_E^2 J_2 \mu_E^4}{L^3 G^4}\left(1 - 5\frac{H^2}{G^2}\right),$$

so that $\dot{\omega} = 0$ for $1 - 5\frac{H^2}{G^2} = 0$. Since $H = G\cos i$, we obtain that $\dot{\omega} = 0$ when the inclination is equal to the critical values

$$i_c = 63.43^o\,, \qquad i_c = 116.56^o\,.$$

2.4 Classification of the Arguments in the Disturbing Functions

With reference to the Hamiltonian (3), the Fourier expansions of $\mathcal{H}_{Earth}$, $\mathcal{H}_{Sun}$ and $\mathcal{H}_{Moon}$ contain infinite number of terms of the following form (see [20, 46, 54, 68] for explicit formulas of these terms):

$$\begin{aligned}
&\mathcal{A}^{Earth}_{k_1k_2k_3k_4}(a,e,i)\cos(k_1M + k_2\theta + k_3\omega + k_4\Omega)\\
&\mathcal{A}^{Sun/SRP}_{k_1k_2k_3k_4k_5k_6}(a,e,i,a_S,e_S,i_S)\cos(k_1M + k_2M_S + k_3\omega + k_4\omega_S + k_5\Omega + k_6\Omega_S)\\
&\mathcal{A}^{Moon}_{k_1k_2k_3k_4k_5k_6}(a,e,i,a_M,e_M,i_M)\cos(k_1M + k_2M_M + k_3\omega + k_4\omega_M + k_5\Omega + k_6\Omega_M)\,,
\end{aligned} \tag{6}$$

where $k_1, k_2, k_3, k_4, k_5, k_6$ are integers. The angles involved in the expansions in (6) can be classified as follows:

- fast angles: M and θ, since $\dot{M} > 360^\circ/day$ and $\dot{\theta} = 360^\circ/day$;
- semi–fast angles: M_S and M_M, since $\dot{M}_S \simeq 1^\circ/day$ and $\dot{M}_M \simeq 13.06^\circ/day$;
- slow angles: ω, Ω, ω_M, Ω_M, since $\dot{\omega}$ and $\dot{\Omega}$ can be approximated as described in Sect. 2.3, ω_S and Ω_S are constant, $\dot{\omega}_M \simeq 0.164^\circ/day$ and $\dot{\Omega}_M \simeq -0.053^\circ/day$.

It is convenient to adopt the following classification of the terms of the expansions:

(1) short periodic terms: the arguments involve the fast angles (M or θ), and $\dot{\theta}$ and $\dot{M}$ are not commensurable;
(2) resonant terms: the arguments involve the fast angles (M or θ), and there is a commensurability between $\dot{\theta}$, $\dot{M}$, as well as the much smaller frequencies $\dot{\omega}$, $\dot{\Omega}$;
(3) semi–secular terms: the cosine arguments are independent of the fast angles θ and M, but depend on either M_S or M_M;

(4) secular terms: the cosine arguments are independent of the sidereal time θ and the mean anomalies M, M_S and M_M.

As a consequence, we have the following types of resonances:

(i) *tesseral resonances,* whenever the orbital period of the debris, the rotation of the Earth, the periods of the argument of perigee and that of the longitude of the ascending node satisfy a commensurability condition of the form $\ell\dot{M} - j\dot{\theta} + j\dot{\Omega} + \ell\dot{\Omega} = 0$, $\ell, j \in \mathbb{N}$. As effect of tesseral resonances, the semi–major axis varies on a time scale of the order of hundred of days;

(ii) *lunisolar resonances*, involving a third–body (Sun-Moon) perturber and further split as
– *secular resonances,* if $k_1\dot{\omega} + k_2\dot{\Omega} + k_3\dot{\omega}_b + k_4\dot{\Omega}_b = 0$ for $(k_1, k_2, k_3, k_4) \in \mathbb{Z}^4$, $b = S, M$; secular resonances provoke long-term variations of eccentricity and inclination on time scales of the order of tens (or hundreds) of years;
– *semi–secular resonances,* which involve the mean anomaly of the Sun or of the Moon; these resonances mostly take place in the LEO region;
– *mean motion resonances* between the mean motion of the debris and Sun/Moon, which never occur in LEO, MEO and GEO.

2.5 Dissipative Effects: The Atmospheric Drag

At low altitudes, say less than 2 000 km, we need to modify Hamilton's equations (2) to take into account the atmospheric drag as

$$\dot{M} = \frac{\partial \mathcal{H}}{\partial L}\,, \qquad \dot{\omega} = \frac{\partial \mathcal{H}}{\partial G}\,, \qquad \dot{\Omega} = \frac{\partial \mathcal{H}}{\partial H}$$
$$\dot{L} = -\frac{\partial \mathcal{H}}{\partial M} + F_L\,, \qquad \dot{G} = -\frac{\partial \mathcal{H}}{\partial \omega} + F_G\,, \qquad \dot{H} = -\frac{\partial \mathcal{H}}{\partial \Omega} + F_H\,,$$

where F_L, F_G, F_H are the components of the contribution due to the atmospheric drag. The effects of the atmospheric drag on the orbital elements are described by the averaged equations as formulated in [71] (see also [22]):

$$\frac{da}{dt} = \frac{1}{2\pi}\int_0^{2\pi} B\,\rho\,\mathcal{F}^{(a)}(a, e, i, M)\,dM$$
$$\frac{de}{dt} = \frac{1}{2\pi}\int_0^{2\pi} B\,\rho\,\mathcal{F}^{(e)}(a, e, i, M)\,dM$$
$$\frac{di}{dt} = 0\,,$$

where B is the ballistic coefficient and $\rho(h)$ is the atmospheric density depending on the altitude h and for which there are density models as in [50, 51, 59]. We are assuming a non-rotating atmosphere, which leads to $\frac{di}{dt} = 0$.

For the local density we use the barometric formula given by

$$\rho(h) = \rho_0 \, \exp\left(-\frac{h - h_0}{H_0}\right),$$

where ρ_0 is the (minimum, mean or maximum) density, estimated for (minimum, mean or maximum) solar activity at the reference altitude h_0, while H_0 is the scaling height at h_0 (see [50]). Using the expressions

$$\dot{L} = \frac{\partial L}{\partial a}\dot{a}\,, \quad \dot{G} = \dot{L}\sqrt{1-e^2} + \frac{\partial G}{\partial e}\dot{e}\,, \quad \dot{H} = \dot{G}\cos i - G \sin i \, \frac{di}{dt}\,,$$

we can derive an explicit formula for F_L, F_G and F_H, which characterize the air drag in the dynamical equations (7).

3 Tesseral Resonances in GEO and MEO Regions

We remind that tesseral resonances occur when a relation of the form $\ell\dot{M} - j\dot{\theta} + j\dot{\Omega} + (\ell - q)\dot{\Omega} = 0$, $\ell, j \in \mathbb{N}, q \in \mathbb{N}$, is satisfied (see [19]).

When $J_2 = 0$, then $\dot{\Omega} = \dot{\Omega} = 0$ and therefore the tesseral resonance reduces to a commensurability between the orbital period of the debris and the sidereal rotation of the Earth: $\ell\dot{M} - j\dot{\theta} = 0$ for $\ell, j \in \mathbb{N}$. In this case, we will speak of $j : \ell$ *gravitational resonances*.

By using Kepler's third law, it follows that a $j : \ell$ gravitational resonance corresponds to a semimajor axis equal to $a_{j:\ell} = (j/\ell)^{-2/3} \, a_{geo}$, where $a_{geo} = 42\,164.1696$ km is the semimajor axis of the geosynchronous orbit. We remark that the value of a_{geo} is the semimajor axis of a Keplerian orbit around a spherical object with the mass of the Earth and with orbital period equal to the Earth's rotational period. Table 2 gives the semimajor axis of the resonances, for several values of ℓ and j.

For a specific resonance, we can approximate $\mathcal{H}_{Earth}$ as the sum of the secular and resonant parts, which are obtained averaging over the short-period terms and retaining the resonant contributions of the angles; the variation of the orbital elements is mainly

Table 2 Value of the semimajor axis for different resonances

$j : \ell$	a in km	$j : \ell$	a in km
1:1	42164.2	4:3	34805.8
2:1	26561.8	5:1	14419.9
3:1	20270.4	5:2	22890.2
3:2	32177.3	5:3	29994.7
4:1	16732.9	5:4	36336

governed by the secular and resonant parts. This leads to consider the following Hamiltonian:

$$\mathcal{H}_{Earth} \cong \mathcal{H}^{sec}_{Earth}(L, G, H, \omega) + \mathcal{H}^{res\ j:\ell}_{Earth}(L, G, H, \ell M - j\theta, \omega, \Omega)\ , \qquad (7)$$

and we can provide explicit analytical expressions for $\mathcal{H}^{sec}_{Earth}$, $\mathcal{H}^{res\ j:\ell}_{Earth}$.

The Hamiltonian (7) is simple enough to allow for explicit computations, though still retaining the essential features of the dynamics. In [13–19], we have used it for the followings tasks:

- expand the Hamiltonian (7) around the resonant value L_{res} and retain only the largest term; after some canonical transformations, we get a pendulum-like Hamiltonian from which one can give an estimate of the amplitude around the resonance;
- analyze the dominant terms, namely the terms prevailing in specific regions of the parameter space;
- compute the location of the equilibria, the splitting or overlapping of the resonant islands, the occurrence of bifurcations.

In this Section we will review some results associated to the 1:1 and 2:1 tesseral resonances (see, respectively, Sects. 3.1 and 3.2), while we will discuss in Sect. 4 tesseral resonances occurring in the LEO region.

3.1 The 1:1 Resonance

For the 1:1 resonance, the expression of the resonant Hamiltonian is given by

$$\begin{aligned}
R^{res\,1:1}_{Earth} \cong\ & \frac{\mu_E R_E^2 J_{22}}{a^3}\left\{\frac{3}{4}(1+\cos i)^2(1-\frac{5}{2}e^2)\cos[2(M-\theta+\omega+\Omega-\lambda_{22})]\right. \\
& \left. + \frac{27}{8}e^2\sin^2 i\cos[2(M-\theta+\Omega-\lambda_{22})]\right\} \\
& + \frac{\mu_E R_E^2 J_{21}}{a^3}\left\{\frac{3}{4}\sin i(1+\cos i)(-\frac{e}{2})\sin(M-\theta+2\omega+\Omega-\lambda_{21})\right. \\
& \left. + \frac{3}{2}e(-\frac{3}{2}\sin i\cos i)\sin[M-\theta+\Omega-\lambda_{21}]\right\} \\
& + \frac{\mu_E R_E^3 J_{31}}{a^4}\left\{-\frac{15}{16}\sin^2 i(1+\cos i)\frac{e^2}{8}\cos(M-\theta+3\omega+\Omega-\lambda_{31})\right. \\
& + \left(\frac{15}{16}\sin^2 i(1+3\cos i)-\frac{3}{4}(1+\cos i)\right)(1+2e^2)\cos(M-\theta+\omega+\Omega-\lambda_{31}) \\
& \left. + \left(\frac{15}{16}\sin^2 i(1-3\cos i)-\frac{3}{4}(1-\cos i)\right)\frac{11e^2}{8}\cos(M-\theta-\omega+\Omega-\lambda_{31})\right\} \\
& + \frac{\mu_E R_E^3 J_{32}}{a^4}\left\{-\frac{15}{8}\sin i(1+\cos i)^2 e\sin(2M-2\theta+3\omega+2\Omega-2\lambda_{32})\right. \\
& \left. + +\frac{45}{8}\sin i(1-2\cos i-3\cos^2 i)e\sin(2M-2\theta+\omega+2\Omega-2\lambda_{32})\right\}
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{\mu_E R_E^3 J_{33}}{a^4}\Big\{\frac{15}{8}(1+\cos i)^3(1-6e^2)\cos[3(M-\theta+\omega+\Omega-\lambda_{33})] \\
&+ \frac{45}{8}\sin^2 i(1+\cos i)\frac{53e^2}{8}\cos(3M-3\theta+\omega+3\Omega-3\lambda_{33})\Big\} \\
&+ \frac{\mu_E R_E^4 J_{41}}{a^5}\Big\{\Big(\frac{35}{16}\sin^3 i(1+2\cos i)-\frac{15}{8}(1+\cos i)\sin i\Big)\frac{e}{2}\sin(M-\theta+2\omega+\Omega-\lambda_{41}) \\
&+ \cos i\Big(\frac{15}{4}\sin i-\frac{105}{16}\sin^3 i\Big)\frac{5e}{2}\sin(M-\theta+\Omega-\lambda_{41})\Big\} \\
&+ \frac{\mu_E R_E^4 J_{42}}{a^5}\Big\{-\frac{105}{32}\sin^2 i(1+\cos i)^2\,\frac{e^2}{2}\cos[2(M-\theta+2\omega+\Omega-\lambda_{42})] \\
&+ \Big(\frac{105}{8}\sin^2 i\cos i(1+\cos i)-\frac{15}{8}(1+\cos i)^2\Big)(1+e^2)\cos[2(M-\theta+\omega+\Omega-\lambda_{42})] \\
&+ \Big(\frac{105}{16}\sin^2 i(1-3\cos^2 i)-\frac{15}{4}\sin^2 i\Big)5e^2\cos[2(M-\theta+\Omega-\lambda_{42})]\Big\} \\
&+ \frac{\mu_E R_E^4 J_{43}}{a^5}\Big\{\frac{105}{16}\sin i(1+\cos i)^3\Big(-\frac{3e}{2}\Big)\sin(3M-3\theta+4\omega+3\Omega-3\lambda_{43}) \\
&+ \frac{105}{8}\sin i(1-3\cos^2 i-2\cos^3 i)\Big(\frac{9e}{2}\Big)\sin(3M-3\theta+2\omega+3\Omega-3\lambda_{43})\Big\} \\
&+ \frac{\mu_E R_E^4 J_{44}}{a^5}\Big\{\frac{105}{16}(1+\cos i)^4\Big(1-11e^2\Big)\cos[4(M-\theta+\omega+\Omega-\lambda_{44})] \\
&+ \frac{105}{4}\sin^2 i(1+\cos i)^2\Big(\frac{53e^2}{4}\Big)\cos(4M-4\theta+2\omega+4\Omega-4\lambda_{44})\Big\}\,.
\end{aligned}
$$

We consider a Hamiltonian function composed by the Keplerian part, the secular contribution and the resonant part. In Fig. 3 we analyze several models obtained taking approximations at different orders of the 1:1 resonant Hamiltonian, precisely at degrees and orders $n = m = 2$, $n = m = 3$ and $n = m = 4$, respectively. We plot the semimajor axis a versus the stroboscopic mean node λ, which is defined for the 1:1 resonance as $\lambda = M - \theta + \omega + \Omega$; the color scale gives the FLI computed as in the Appendix (see [37], compare also with [41]). The figure shows some differences between the plots at orders 2 and 3, while those at orders 3 and 4 are quite similar, thus indicating that for these values of the parameters the order 3 is already a good approximation of the expansion. Finally, we validate the Hamiltonian model by comparing it with the results obtained integrating the Cartesian equations of motion; the bottom–right panel of Fig. 3 shows the results obtained by integrating the Cartesian equations of motion, considering the expansion of the Earth' geopotential up to degree and order $n = m = 3$ and the influence of the Moon, Sun and solar radiation pressure with $A/m = 0.1\,[\mathrm{m}^2/\mathrm{kg}]$.

The cartographic study based on the FLI (or any other chaos indicator) can be used to get much information on the models and the dynamics. For example, the top panels of Fig. 4 are obtained using the same model, but taking different inclinations, thus showing how the amplitude of libration around the 1:1 resonance varies as a function of the orbital inclination. The lower panels of Fig. 4 are obtained by integrating the Cartesian equations of motion, using only an expansion to third order of the geopotential and adding also the effects of Sun, Moon and solar radiation pressure.

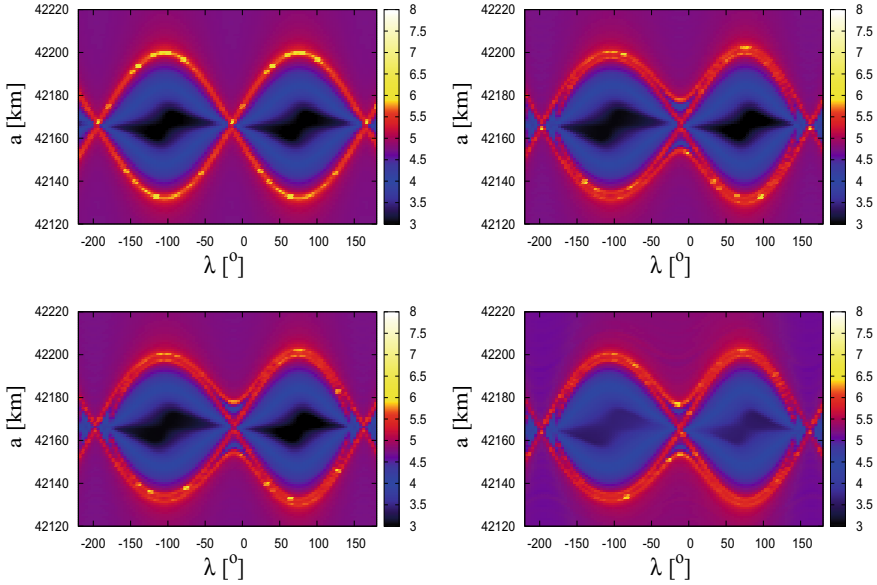


Fig. 3 FLI for the 1:1 resonance for $e = 0.01$, $i = 5^o$, $\omega = \Omega = 0^o$: Hamilton's equations with $\mathcal{H}_{Earth}$ approximated to degree and order $n = m = 2$ (top left), $n = m = 3$ (top right), and $n = m = 4$ (bottom left) and Cartesian equations for $n = m = 3$ and adding the influence of the Moon, Sun and SRP with $A/m = 0.1[\mathrm{m}^2/\mathrm{kg}]$ (bottom right)

3.2 The 2:1 Resonance

For the 2:1 resonance, the expression of the resonant Hamiltonian is given by[1]

$$
\begin{aligned}
R^{res2:1}_{Earth} \cong\ & \frac{\mu_E R_E^2 J_{22}}{a^3}\Big\{\frac{3}{4}(1+\cos i)^2\Big(-\frac{e}{2}+\frac{e^3}{16}\Big)\cos(M-2\theta+2\omega+2\Omega-2\lambda_{22}) \\
& + \frac{3}{2}\sin^2 i\Big(\frac{3}{2}e+\frac{27}{16}e^3\Big)\cos(M-2\theta+2\Omega-2\lambda_{22}) \\
& + \frac{1}{64}(1-\cos i)^2 e^3\cos(M-2\theta-2\omega+2\Omega-2\lambda_{22})\Big\} \\
& + \frac{\mu_E R_E^3 J_{32}}{a^4}\Big\{\frac{15}{8}\sin i(1+\cos i)^2\Big(\frac{e^2}{8}+\frac{e^4}{48}\Big)\sin(M-2\theta+3\omega+2\Omega-2\lambda_{32}) \\
& + \frac{15}{8}\sin i(1-2\cos i-3\cos^2 i)\Big(1+2e^2+\frac{239e^4}{64}\Big)\sin(M-2\theta+\omega+2\Omega-2\lambda_{32}) \\
& - \frac{15}{8}\sin i(1+2\cos i-3\cos^2 i)\Big(\frac{11e^2}{8}+\frac{49e^4}{16}\Big)\sin(M-2\theta-\omega+2\Omega-2\lambda_{32}) \\
& - \frac{5}{1024}\sin i(1-\cos i)^2 e^4\sin(M-2\theta-3\omega+2\Omega-2\lambda_{32})\Big\}
\end{aligned}
$$

[1] Notice that three coefficients in [13] were not correct and we take the opportunity to provide here the correct expansion.

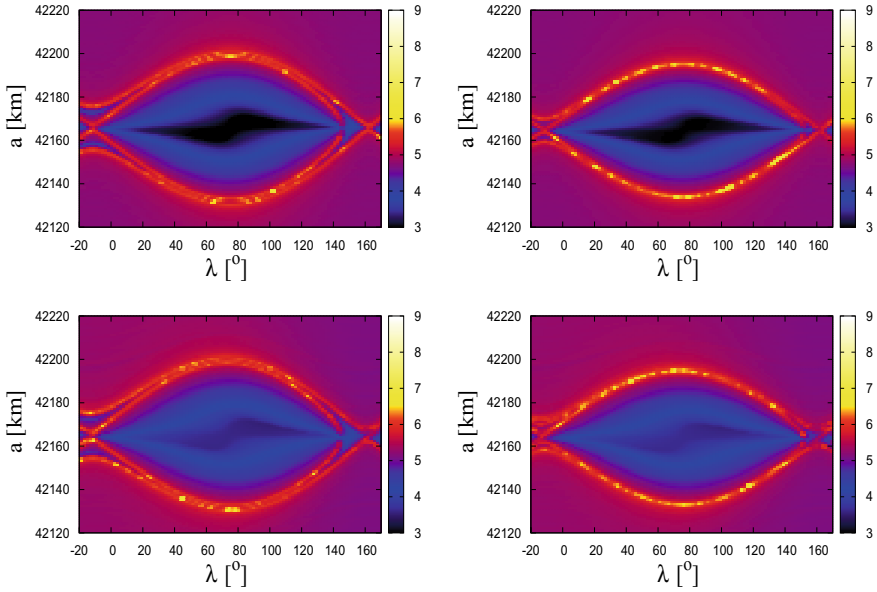


Fig. 4 1:1 resonance for $e = 0.1$, $\omega = \Omega = 0^o$ and inclination $i = 20^o$ (left panels), respectively $i = 40^o$ (right panels). Top panels: Hamilton's equations for $n = m = 4$. Bottom panels: Cartesian equations for $n = m = 3$, adding the influence of the Moon, Sun and SRP with $A/m = 0.1[\mathrm{m}^2/\mathrm{kg}]$

$$
\begin{aligned}
&+ \frac{\mu_E R_E^4 J_{42}}{a^5}\Big\{\frac{35}{512}\sin^2 i(1+\cos i)^2 e^3\cos(M-2\theta+4\omega+2\Omega-2\lambda_{42})\\
&+ \Big(\frac{105}{8}\sin^2 i\cos i(1+\cos i)-\frac{15}{8}(1+\cos i)^2\Big)\Big(\frac{e}{2}+\frac{33e^3}{16}\Big)\cos(M-2\theta+2\omega+2\Omega-2\lambda_{42})\\
&+ \Big(\frac{105}{16}\sin^2 i(1-3\cos^2 i)-\frac{15}{4}\sin^2 i\Big)\Big(\frac{5e}{2}+\frac{135e^3}{16}\Big)\cos(M-2\theta+2\Omega-2\lambda_{42})\\
&- \Big(\frac{105}{8}\sin^2 i\cos i(1-\cos i)+\frac{15}{8}(1-\cos i)^2\Big)\frac{49e^3}{48}\cos(M-2\theta-2\omega+2\Omega-2\lambda_{42})\Big\}\\
&+ \frac{\mu_E R_E^4 J_{44}}{a^5}\Big\{\frac{105}{16}(1+\cos i)^4\Big(\frac{e^2}{2}-\frac{e^4}{3}\Big)\cos[2(M-2\theta+2\omega+2\Omega-2\lambda_{44})]\\
&+ \frac{105}{4}\sin^2 i(1+\cos i)^2\Big(1+e^2+\frac{65e^4}{16}\Big)\cos[2(M-2\theta+\omega+2\Omega-2\lambda_{44})]\\
&+ \frac{315}{8}\sin^4 i\Big(5e^2+\frac{155e^4}{12}\Big)\cos[2(M-2\theta+2\Omega-2\lambda_{44})]\\
&+ \frac{105}{4}\sin^2 i(1-\cos i)^2\frac{67e^4}{48}\cos[2(M-2\theta-\omega+2\Omega-2\lambda_{44})]\Big\}\,.
\end{aligned}
$$

Denoting by σ the angle $\sigma = M + \omega - 2(\theta - \Omega)$, let us approximate the 2:1 resonant Hamiltonian as the sum of three terms, t_1, t_2, t_3, defined as

$$
t_1 = \frac{\mu_E R_E^2 J_{22}}{a^3}\Big\{\frac{3}{4}(1+\cos i)^2\Big(-\frac{e}{2}\Big)\Big\}\cos(\sigma+\omega-2\lambda_{22})
$$

$$t_2 = \frac{\mu_E R_E^2 J_{22}}{a^3}\left\{\frac{3}{2}\sin^2 i\left(\frac{3}{2}e\right)\right\}\cos(\sigma-\omega-2\lambda_{22})$$
$$t_3 = \frac{\mu_E R_E^3 J_{32}}{a^4}\left\{\frac{15}{8}\sin i(1-2\cos i-3\cos^2 i)\left(1+2e^2\right)\right\}\sin(\sigma-2\lambda_{32})\ . \quad (8)$$

These terms are dominant in various regions of the phase space (see [13]). Hence, we consider the Hamiltonian

$$\mathcal{K}(L,G,H,\sigma,\omega,\Omega) = -\frac{\mu_E^2}{2L^2} - 2\dot{\theta}L + R_{earth}^{sec} + t_1 + t_2 + t_3\ , \quad (9)$$

where the secular part is given by

$$R_{earth}^{sec} \cong \frac{\mu_E R_E^2 J_2}{a^3}\left(\frac{3}{4}\sin^2 i - \frac{1}{2}\right)(1-e^2)^{-3/2}\ .$$

We notice that the three terms in (9) contain different combinations of the angles, precisely $\sigma, \sigma+\omega, \sigma-\omega$. When integrating Hamilton's equations associated to (9) with $t_2 = t_3 = 0$, we obtain a pendulum like structure as in the upper-left panel of Fig. 5; the same happens in the upper-right and lower-left panels of Fig. 5, which are obtained, respectively, setting $t_1 = t_3 = 0$ and $t_1 = t_2 = 0$. Indeed, the angle Ω is cyclic and it does not influence the location of the equilibria. On the contrary, ω plays an important role, since there are comparable terms for $\sigma \pm \omega - 2\lambda_{22}$ and $\sigma - 2\lambda_{32}$, thus showing that the location of the equilibria and the pattern of the resonances are strongly affected by ω. When considering all terms, there is a superposition of the resonances associated to the different terms, which generates chaos as in the lower-right panel of Fig. 5.

The overlap of resonances as a mechanism for the onset of chaos was already discussed in [24]; in a nutshell, with reference to Fig. 6, the left plot shows two islands around two different equilibria, such that the distance between the equilibria (dashed line) is greater than the sum of the semi-amplitudes of the resonant islands (dotted lines); in the right panel, instead, the distance between the equilibria is smaller than the sum of the semi-amplitudes, which is the situation in which chaos can appear.

An interesting phenomenon associated to the 2:1 resonance is that of *transcritical bifurcations*. We notice that the term t_3 in (8) contains the following function

$$f : [0^o, 90^o] \longrightarrow \mathbb{R}\ , \qquad f(i) = -\sin i(1-2\cos i-3\cos^2 i)\ , \quad (10)$$

which changes sign at $i = 70.53^o$. Precisely, for small eccentricities and for $i < 70.53^o$ the stable point is at about $\sigma \simeq 55^o$ and the unstable equilibrium is located at $\sigma \simeq 235^o$. For $i > 70.53^o$ the situation is opposite: the unstable point is at $\sigma \simeq 55^o$, while the stable equilibrium is located at $\sigma \simeq 235^o$. This shows that at $i = 70.53^o$ we have the occurrence of a transcritical bifurcation (Fig. 7).

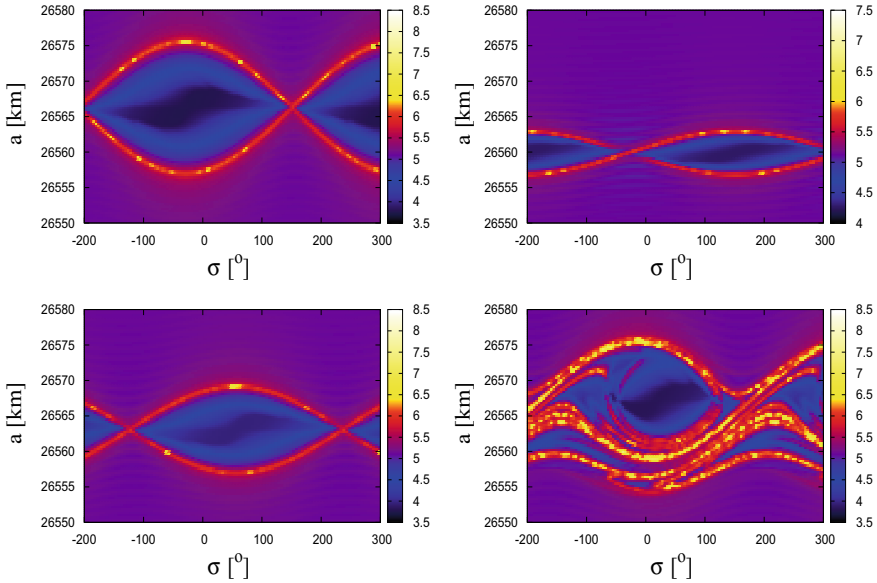


Fig. 5 FLI plots associated to (9) for $e = 0.15$, $i = 15^o$, $\omega = \Omega = 0$ with $t_2 = t_3 = 0$ (upper-left panel), $t_1 = t_3 = 0$ (upper-right panel), $t_1 = t_2 = 0$ (lower-left panel), all terms considered (lower-right panel)

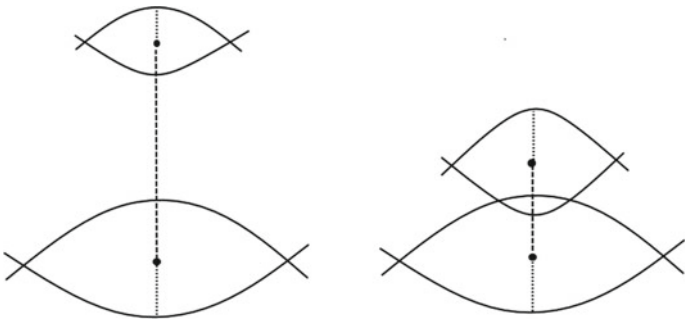


Fig. 6 Non-overlapping resonances (left panel), overlapping resonances (right panel)

4 Tesseral Resonances in the LEO Region

In this Section we concentrate on tesseral resonances occurring in the LEO region. Following [17], we consider the resonances of order $11 : 1$, $12 : 1$, $13 : 1$, $14 : 1$, which occur between about 880 and 2 146 km, as shown in Table 3, which provides also the altitudes at perigee and apogee for an object with eccentricity equal to 0.02. The $15 : 1$ resonance has been investigated in [35]. The values from Table 3

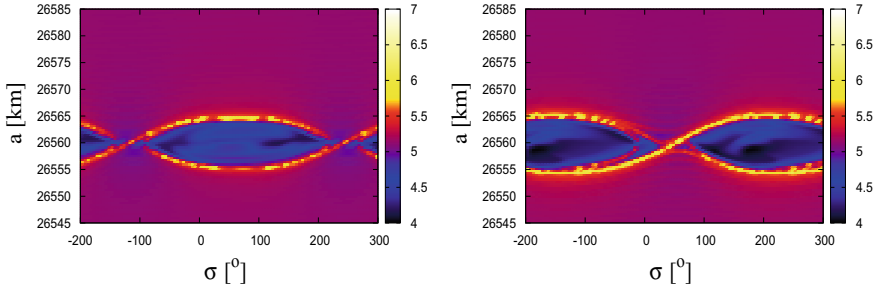


Fig. 7 A transcritical bifurcation for the 2:1 resonance for $e = 0.01$, $\omega = 0$, $\Omega = 0$: $i = 60^o$ (left panel), $i = 85^o$ (right panel)

are computed by using Kepler's third law. However, as discussed in [17], the exact altitude of each resonance depends also on the value of the inclination.

For each resonance, we approximate the Hamiltonian $\mathcal{H}_{Earth}$ as the sum of the secular and resonant parts:

$$\mathcal{H}_{Earth} \cong \mathcal{H}^{sec}_{Earth} + \mathcal{H}^{res\, m:1}_{Earth} \ .$$

We denote the resonant angle by $\sigma_{m1} = M - m\theta + \omega + m\Omega$. In [17] we introduced three different models, with increasing difficulty, apt to study the dynamics in LEO. Their definition is the following:

$\diamond$ Model 1 (or the toy model): a 1 d.o.f. problem which considers the effects of the Keplerian part, a J_2 approximation for the secular part, the resonant part containing just the five most important terms, the dissipative part characterized by the function F_L; the effect of the solar cycle is disregarded. The toy model is described by the following equations:

$$\begin{aligned} \dot{\sigma}_{m1} &= h^{(m)}_{,L}(L, G, H) + \varepsilon \mathcal{A}^{(m)}_{,L}(L, G, H)\ \cos(\sigma_{m1} - \varphi^{(m)}) \ , \\ \dot{L} &= \varepsilon \mathcal{A}^{(m)}(L, G, H) \sin(\sigma_{m1} - \varphi^{(m)}) - \eta D^{(m)}_{L}(L, G, H) \ , \end{aligned}$$

Table 3 The semimajor axis and the altitude corresponding to some resonances of order $m : 1$, as well as the perigee and apogee altitudes of a resonant elliptic orbit with $e = 0.02$. The altitudes are computed by considering the reference value $R_E = 6378.14$ [km] for the Earth's radius

$m : 1$	a (km)	Altitude (km)	Perigee altitude for $e = 0.02$ (km)	Apogee altitude for $e = 0.02$ (km)
11:1	8524.75	2146.61	1976.25	2317.25
12:1	8044.32	1666.18	1505.43	1827.21
13:1	7626.31	1248.17	1095.78	1400.84
14:1	7258.69	880.55	735.52	1025.86

where the actions G and H are constants, ε is a small parameter (of the order of 10^{-9}), artificially introduced so that the functions $h^{(m)}$, $\mathcal{A}^{(m)}$ and $D_{\scriptscriptstyle L}^{(m)}$ have comparable magnitudes, and $\eta = B\rho$ is the dissipative parameter. We refer to [15] for the explicit form of the functions $h^{(m)}$, $\mathcal{A}^{(m)}$, $D_L^{(m)}$ and the derivative $h_{,L}^{(m)}$, $\mathcal{A}_{,L}^{(m)}$ of the functions $h^{(m)}$, $\mathcal{A}^{(m)}$ with respect to L.

$\diamond$ Model 2 (or the dissipative model of LEO resonances (DMLR)): a 3 d.o.f. problem described in terms of the Keplerian part, the disturbing function due to the Earth $\mathcal{H}_{Earth}$, the dissipative functions F_L, F_G, F_H; the effect of the solar cycle is disregarded.

$\diamond$ Model 3 (or the full model): including the Keplerian part, the disturbing function due to the Earth $\mathcal{H}_{Earth}$, the dissipative functions F_L, F_G, F_H, the solar cycle, and the influence of the Moon and Sun.

The following result, borrowed from [17], provides the existence of equilibrium points in Model 1.

Theorem 1 *Let us take eccentricity and inclination in the intervals* $e \in [0, 0.02]$ *and* $i \in [0^o, 120^o]$. *Let* $(\sigma_{m1}^{(0)}, L_0)$ *be an equilibrium point for the conservative version of Model 1. Assume that the parameters* η, ε *satisfy the following inequalities:*

$$\left| \frac{\eta D_{\scriptscriptstyle L}^{(m)}(L_0, G_0, H_0)}{\varepsilon \mathcal{A}^{(m)}(L_0, G_0, H_0)} \right| \leq 1 - \delta \,, \qquad \gamma_1 \varepsilon + \gamma_2 \eta + \gamma_3 \varepsilon^2 < \delta$$

for some positive real constants δ, γ_1, γ_2 *and* γ_3.

Then, the dissipative version of Model 1 admits equilibrium points $(\sigma_{m1}^{(1)}, L_1)$ *which, at first order in* η, *are defined by the following expressions:*

$$\sigma_{m1}^{(1)} = \sigma_{m1}^{(0)} + \frac{D_{\scriptscriptstyle L}^{(m)}(L_0, G_0, H_0)}{\varepsilon \mathcal{A}^{(m)}(L_0, G_0, H_0) \cos(\sigma_{m1}^0 - \varphi^{(m)})} \, \eta \,, \qquad L_1 = L_0 \,.$$

As shown in [17], if $(\sigma_{m1}^{(0)}, L_0)$ is a center (saddle) for the conservative toy model, then $(\sigma_{m1}^{(1)}, L_1)$ is an unstable spiral (saddle) for the dissipative toy model. Beside, in [17] a thorough analysis of the location of the equilibrium points is performed, together with a study of the position of the equilibria as a function of η. Moreover, [17] provides numerical evidence that the results for Model 1 are valid also for Models 2 and 3.

Let us now consider Model 3. Figure 8 shows the typical phenomena of temporary capture into resonance and passage through resonance. The behavior in the neighborhood of a resonance has a strongly stochastic feature, since a small change in the initial conditions, or a small perturbation, leads to a drastically different scenario. The curves depicted in Fig. 8 describe the evolution of three orbits in the vicinity of the 14:1 resonance, derived for the same ballistic coefficient and initial conditions, but the initial value of the critical angle $\sigma_{14\,1}$. The capture time depends on various factors such as the value of the ballistic coefficient, the density of the atmosphere, the initial conditions.

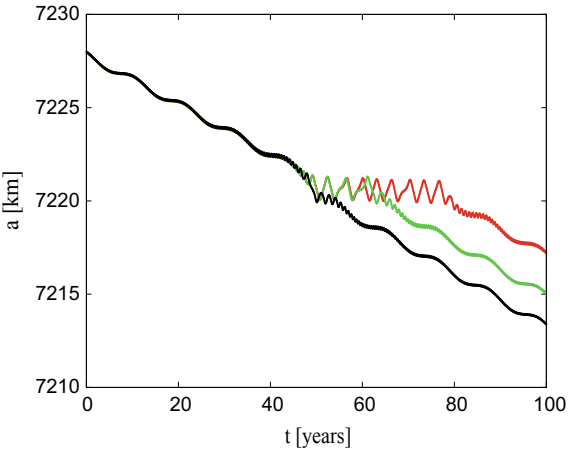


Fig. 8 The 14:1 resonance: temporary capture into resonance (green and red curves) and passage through the resonance (black line), under the Model 3. The ballistic coefficient and initial conditions for the three orbits are $B = 50\,[\mathrm{cm}^2/\mathrm{kg}]$, $e = 0.005$, $i = 65°$, $\omega = \Omega = 0°$ and respectively $\sigma_{14\,1} = 35°$ (red), $\sigma_{14\,1} = 40°$ (green), $\sigma_{14\,1} = 100°$ (black)

5 Secular Resonances

With the aim of studying the dynamics in a MEO region which does not belong to the libration regions of tesseral resonances, then one can reduce to consider the following Hamiltonian:

$$\mathcal{H}^{sec} = \mathcal{H}^{sec}_{Earth} + \mathcal{H}^{sec}_{Sun} + \mathcal{H}^{sec}_{Moon}\,,$$

where

$$\mathcal{H}^{sec}_{Earth} = \frac{R_E^2 J_2 \mu_E^4}{4}\,\frac{1}{L^3 G^3}\left(1 - 3\frac{H^2}{G^2}\right),$$

that is, we consider only the most important contribution corresponding to the J_2 gravity coefficient of the Earth's perturbation.

The terms $\mathcal{H}^{sec}_{Sun}$ and $\mathcal{H}^{sec}_{Moon}$ are defined by averaging $\mathcal{H}_{Sun}$ and $\mathcal{H}_{Moon}$ over both the mean anomaly M of the debris and the mean anomaly M_b of the perturbing body, as well as by truncating the series expansions up to the second order in the ratio a/a_b with $b = S, M$; we will refer to this truncation as the *quadrupolar approximation*. Notice that we neglected the Keplerian part, since the mean anomaly M is an ignorable variable and therefore L is constant. Moreover, $\mathcal{H}^{sec}$ depends on time only through Ω_M, which varies linearly with rate $\dot{\Omega}_M \simeq -0.053°/day$. In conclusion, the Hamiltonian $\mathcal{H}^{sec}$ is *a two degrees of freedom non-autonomous Hamiltonian*, depending on the parameter L.

Figure 9 shows a validation of the Hamiltonian approach by comparing two different orbits with the results obtained integrating the Cartesian equations of motion. There is a very good overlap for a very long time interval. It is worth noticing that the integration of Hamilton's equations takes a few seconds, while that of the Cartesian equations takes tens of minutes.

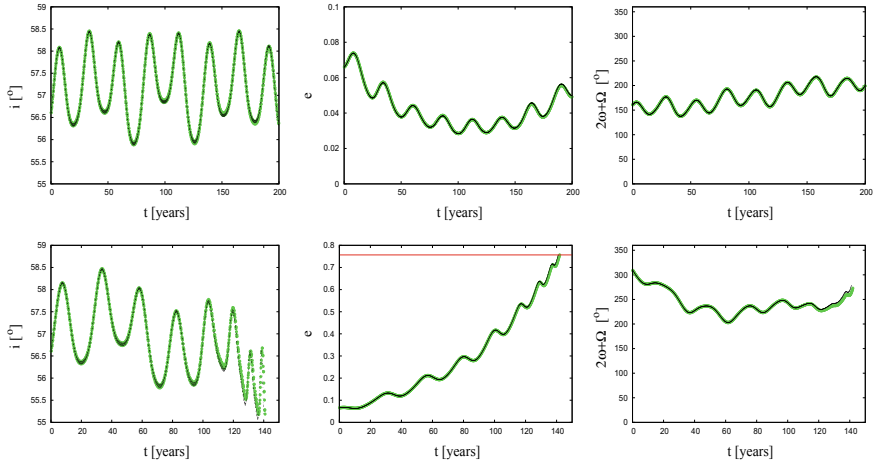


Fig. 9 Integration of the two orbits for which $a = 26\,700$ [km], $e(0) = 0.066$, $i(0) = 56.6^o$, $\Omega(0) = 100^o$, and $\omega(0) = 30^o$ (top panels), respectively, $\omega(0) = 105^o$ (bottom plots). Green color—Hamiltonian formulation, black color—Cartesian model. Red line—the eccentricity value leading to re-entry

5.1 Types of Secular Resonances

The secular parts $\mathcal{H}^{sec}_{Sun}$ and $\mathcal{H}^{sec}_{Moon}$ have the form [20, 28, 62, 68]

$$\mathcal{H}^{sec}_{Sun} = \sum_{m=0}^{2}\sum_{p=0}^{2} \mathcal{A}^{Sun}_{mp}(a, a_S, e, e_S, i, i_S)\cos\Big((2-2p)\omega + m(\Omega - \Omega_S)\Big)\,,$$

and

$$\mathcal{H}^{sec}_{Moon} = \sum_{m=0}^{2}\sum_{s=0}^{2}\sum_{p=0}^{2} \mathcal{A}^{Moon}_{msp}(a, a_M, e, e_M, i, i_M)\cos\Big((2-2p)\omega + m\Omega \pm s\Omega_M\Big)$$

for suitable coefficients $\mathcal{A}^{Sun}_{mp}$ and $\mathcal{A}^{Moon}_{msp}$. In the MEO region two types of secular resonances are possible:

(i) the ones that do not involve the rate $\dot\Omega_M$, namely

$$(2-2p)\dot\Omega + m\dot\Omega = 0\,,$$

called also *resonances depending only on inclination;*

(ii) the resonances that satisfy a relation of the following form:

$$(2-2p)\dot\Omega + m\dot\Omega + \kappa\dot\Omega_M = 0, \qquad \kappa \neq 0\,.$$

The first set of resonances does not depend on the semi-major axis and eccentricity, because using

$$\dot{\omega} \simeq 4.98\Big(\frac{R_E}{a}\Big)^{7/2}(1-e^2)^{-2}(5\cos^2 i - 1)\ ^o/day\ ,$$
$$\dot{\Omega} \simeq -9.97\Big(\frac{R_E}{a}\Big)^{7/2}(1-e^2)^{-2}\cos i\ ^o/day\ ,$$

then the condition $(2-2p)\dot{\Omega} + m\dot{\Omega} = 0$ gives a relation that depends only on the inclination. Such resonances have the following locations:

$$\begin{aligned}
\dot{\Omega} = 0 \quad & \text{at} \quad 90^\circ\ , \\
\dot{\omega} + \dot{\Omega} = 0 \quad & \text{at} \quad 46.4^\circ \quad \text{or} \quad 106.9^\circ\ , \\
-\dot{\omega} + \dot{\Omega} = 0 \quad & \text{at} \quad 73.2^\circ \quad \text{or} \quad 133.6^\circ\ , \\
-2\dot{\omega} + \dot{\Omega} = 0 \quad & \text{at} \quad 69.0^\circ \quad \text{or} \quad 123.9^\circ\ , \\
2\dot{\omega} + \dot{\Omega} = 0 \quad & \text{at} \quad 56.1^\circ \quad \text{or} \quad 111.0^\circ\ .
\end{aligned}$$

The resonances involving the lunar ascending node Ω_M are responsible for the existence of a web–like structure of resonances in the phase space [28, 32, 76]. Figure 10 shows the web structure of the resonances in the plane (G, H). The units of length and time are normalized so that the geostationary distance is unity (it amounts to 42 164.1696 km) and that the period of the Earth's rotation is equal to 2π.

5.2 *Effects of Secular Resonances*

In contrast with tesseral resonances, whose main effect is a variation of the semi-major axis on a time scale of the order of hundreds of days, secular resonances influence the evolution of eccentricity and inclination on time scales of the order of tens (or hundreds) of years. Although lunar and solar perturbations are investigated since the beginning of space age [26, 54, 55, 72], recent qualitative and quantitative studies [2, 16, 18–20, 23, 28, 31, 45, 48, 75, 81, 84, 92] show that the dynamics of secular resonances is very complex. Several mathematical models [16, 18, 20] introduced to get better insights into the dynamics of secular resonances reveal that, in many cases, a specific resonance cannot be modeled by a pendulum–type system, but one should use extended fundamental models, as described in [8, 52]. Besides, the web structure of lunisolar secular resonances, revealed by the two degrees of freedom non–autonomous Hamiltonian $\mathcal{H}^{sec}$ and depicted in Fig. 10, is responsible for intricate interactions that lead to a plethora of dynamical phenomena. Depending on the region explored, one can find bifurcation of equilibria [16, 18]; libration regions leading to large as well as to small excursions in eccentricity; overlapping of resonances and the onset of chaos [32, 76]; transport in the phase space [28]. The

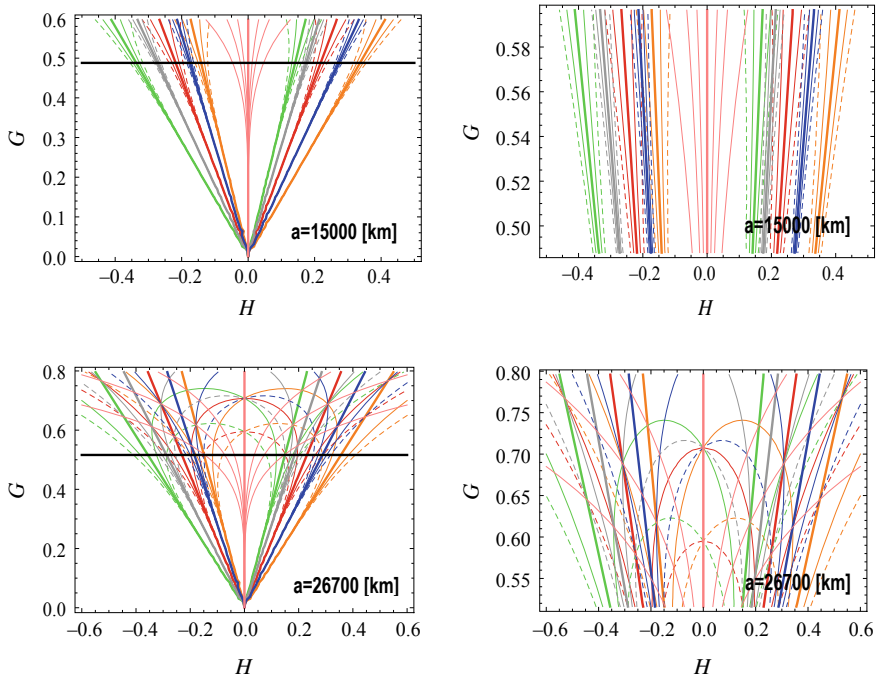


Fig. 10 The web structure of lunisolar resonances at $a = 15000$ [km] (top plots) and respectively $a = 26700$ [km] (bottom panels): $\dot{\omega} + \dot{\Omega} = 0$ (orange color, $i = 46.4°$, $i = 106.9°$); $2\dot{\omega} + \dot{\Omega} = 0$ (blue color, $i = 56.1°$, $i = 111°$); $\dot{\omega} = 0$ (red color, $i = 63.4°$, $i = 116.6°$); $2\dot{\omega} - \dot{\Omega} = 0$ (grey color, $i = 69.0°$); $\dot{\omega} - \dot{\Omega} = 0$ (green color, $i = 73.2°$, $i = 133.6°$); $\dot{\Omega} = 0$ (pink color, $i = 90°$). The black horizontal line in the left panels represents the Delaunay action G leading to re-entry. Right panels provide a zoom-in of the non-colliding-orbits regions

practical implications of these phenomena are closely connected with the design of disposal orbits of non-operative satellites and the population of space debris. For instance, the long–term growth in eccentricity, observed for disposal orbits of various satellites, such as GPS, GLONASS, and GALILEO (see [23]) may be viewed as an effect of the lunisolar resonances and their complex interactions.

As a final remark of this Section, we mention that the two orbits shown in Fig. 9 are both located inside the libration regions of the resonance $2\dot{\omega} + \dot{\Omega} = 0$. Indeed, the right panels of Fig. 9 show that the critical angle $2\omega + \Omega$ librates. The orbit whose evolution is depicted in the top panels of Fig. 9 is located inside a libration region leading to small excursions in eccentricity, while the other one is a colliding orbit.

6 Conclusions and Perspectives

The powerful tool offered by the Hamiltonian formulation of the equations of motion of space debris allows us to get important information about the dynamics of the millions of objects that populate the sky around our planet. In particular, we have been able to make an accurate study of tesseral, gravitational and lunisolar secular resonances, using both numerical and analytical tools. We have noticed that tesseral resonances induce variations of the semi-major axis, while lunisolar secular resonances provoke variations of the eccentricity and inclination. Besides, lunisolar secular resonances provoke transcritical bifurcations and overlapping of resonances that give rise to chaotic motions.

The numerical and analytical studies of the LEO tesseral resonances highlighted an orbital decay caused by the air drag, which can be (temporarily) balanced by resonant effects, although such phenomenon depends on a large class of values of dynamic parameters (ballistic coefficient, inclination, eccentricity, etc.). Indeed, in LEO we have observed several different behaviors, among which temporary capture or passage through a resonance, trapped motions, escape motions, shift of equilibria along some axes, bifurcations.

The study of the dynamics of space debris can be continued in several directions, among which a more detailed analysis of the LEO resonance to discriminate between the possibility of parking satellites in the close vicinity of the equilibrium points, or, on the contrary, to avoid such zones due to a possible accumulation of a large number of space debris in these regions. We believe and hope that such tools might be used to design suitable disposal orbits that could help to solve the problem of the sustainability of the Earth's environment.

Acknowledgements A.C. acknowledges the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C18000100006, MIUR-PRIN 20178CJA2B "New Frontiers of Celestial Mechanics: theory and Applications". A.C. and C.G. acknowledges EU H2020 MSCA ETN Stardust-Reloaded Grant Agreement 813644.

Appendix: FLI

In this Appendix, we briefly review the definition of *Fast Lyapunov Indicators* (hereafter FLI) for which we refer the reader to [37] (see also [47]). Details on Lyapunov exponents are provided in [5].

Given a vector field $\mathbf{F} = \mathbf{F}(\mathbf{X}, t)$ for $\mathbf{X} \in \mathbb{R}^n$, $t \in \mathbb{R}$, we consider the equations of motion and the associated variational equations defined as

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X}, t) \ , \qquad \frac{d\mathbf{V}}{dt} = \frac{\partial \mathbf{F}}{\partial \mathbf{X}}(\mathbf{X}, t)\mathbf{V} \ .$$

For an initial condition $\mathbf{X}(0)$, the FLI at a given time $t = T$ is defined as

$$\mathrm{FLI}(\mathbf{X}(0), \mathbf{V}(0), T) \equiv \sup_{0<t\leq T} \log \|\mathbf{V}(t)\| \, .$$

We remark that the FLI provides much information on the dynamics, among which the regular or chaotic character of the dynamics, the location of the equilibrium points, the role of higher degree harmonic terms.

References

1. Alessi, E.M., Colombo, C., Rossi, A.: Phase space description of the dynamics due to the coupled effect of the planetary oblateness and the solar radiation pressure perturbations. Celest. Mech. Dyn. Astr. **131**, 43 (2019)
2. Alessi, E.M., Deleflie, F., Rosengren, A.J., Rossi, A., Valsecchi, G.B., Daquin, J., Merz, K.: A numerical investigation on the eccentricity growth of GNSS disposal orbits. Celest. Mech. Dyn. Astr. **125**(1), 71–90 (2016)
3. Alessi, E.M., Schettino, G., Rossi, A., Valsecchi, G.B.: Natural highways for end-of-life solutions in the LEO region. Celest. Mech. Dyn. Astron. **130**, n. 34 (2018)
4. Anderson, P.V., Schaub, H.: Local orbital debris flux study in the geostationary ring. Adv. Space Res. **51**, 2195–2206 (2013)
5. Arnold, L., Wihstutz, V.: Lyapunov exponents, a Survey. Lecture Notes in Mathematics, vol. 1186, pp. 1–26. Springer (1986)
6. Bezdek, A., Vokrouhlický, D.: Semianalytic theory of motion for close-Earth spherical satellites including drag and gravitational perturbation. Planetary and Space Science **52**(14), 1233–1249 (2004)
7. Belyanin, S., Gurfil, P.: Semianalytical study of geosynchronous orbits about a precessing oblate Earth under lunisolar gravitation and tesseral resonance. J. Astr. Sci. **57**, 513–543 (2010)
8. Breiter, S.: Lunisolar resonances revisited. Celest. Mech. Dyn. Astr. **81**, 81–91 (2001)
9. Breiter, S., Wytrzyszczak, I., Melendo, B.: Long–term predictability of orbits around the geosynchronous altitude. Adv. Space Res. **35**, 1313–1317 (2005)
10. Casanova, D., Petit, A., Lemaitre, A.: Long-term evolution of space debris under the J_2 effect, the solar radiation pressure and the solar and lunar perturbations. Celest. Mech. Dyn. Astr. **123**, 223–238 (2015)
11. Celletti, A.: Stability and Chaos in Celestial Mechanics. Springer, Published in Association with Praxis Publishing Ltd., Berlin (Chichester, ISBN: 978-3-540-85145-5) (2010)
12. Celletti, A., De Blasi, I., Efthymiopoulos, C.: Nekhoroshev sstimates for satellites' orbital stability. Submitted to Celest. Mech. Dyn. Astr. (2022)
13. Celletti, A., Galeş, C.: On the dynamics of space debris: 1:1 and 2:1 resonances. J. Nonlinear Sci. **24**(6), 1231–1262 (2014)
14. Celletti, A., Galeş, C.: Dynamical investigation of minor resonances for space debris. Celest. Mech. Dyn. Astr. **123**(2), 203–222 (2015)
15. Celletti, A., Galeş, C.: A study of the main resonances outside the geostationary ring. Adv. Space Res. **56**, 388–405 (2015)
16. Celletti, A., Galeş, C.: A study of the lunisolar secular resonance $2\dot{\omega} + \dot{\Omega} = 0$. Front. Astron. Space Sci.-Fund. Astr. (2016). http://dx.doi.org/10.3389/fspas.2016.00011
17. Celletti, A., Galeş, C.: Dynamics of resonances and equilibria of Low Earth Objects. SIAM J. Appl. Dyn. Syst. **17**, 203–235 (2018)
18. Celletti, A., Galeş, C., Pucacco, G.: Bifurcation of lunisolar secular resonances for space debris orbits. SIAM J. Appl. Dyn. Syst. **15**, 1352–1383 (2016)
19. Celletti, A., Galeş, C., Lhotka, C.: Resonances in the Earth's space environment. Comm. Nonlin. Sc. Num. Sim. **84**, 105185 (2020)

20. Celletti, A., Galeş, C., Pucacco, G., Rosengren, A.: Analytical development of the lunisolar disturbing function and the critical inclination secular resonance. Celest. Mech. Dyn. Astron. **127**(3), 259–283 (2017)
21. Celletti, A., Pucacco, G., Vartolomei, T.: Proper elements for space debris. Celest. Mech. Dyn. Astr. **134**, 11 (2022)
22. Chao, C.C.: Applied Orbit Perturbation and Maintenance. Aerospace Press Series, AIAA, Reston, Virgina (2005)
23. Chao, C.C., Gick, R.A.: Long-term evolution of navigation satellite orbits: GPS/GLONASS/GALILEO. Adv. Space Res. **34**, 1221–1226 (2004)
24. Chirikov, B.V.: A universal instability of many-dimensional oscillator systems. Phys. Rep. **52**, 263–379 (1979)
25. Colombo, C., Lücking, C., McInnes, C.-R.: Orbital dynamics of high area-to-mass ratio spacecraft with J_2 and solar radiation pressure for novel Earth observation and communication services. Acta Astronautica **81**, 137–150 (2012)
26. Cook, G.E.: Luni-solar perturbations of the orbit of an Earth satellite. Geophys. J. **6**, 271–291 (1962)
27. Cook, G.E.: Perturbations of near-circular orbits by the Earth's gravitational potential. Planet. Space Sci. **14**, 433–444 (1966)
28. Daquin, J., Rosengren, A.J., Alessi, E.M., Deleflie, F., Valsecchi, G.B., Rossi, A.: The dynamical structure of the MEO region: long-term stability, chaos, and transport. Celest. Mech. Dyn. Astr. **124**, 335–366 (2016)
29. De Blasi, I., Celletti, A., Efthymiopoulos, C.: Semi-analytical estimates for the orbital stability of Earth's satellite. J. Nonlinear Sci. **31**, 93 (2021)
30. Deienno, R., Merguizo Sanchez, D., Bertachini de Almeida Prado, A.F., Smirnov, G.: Satellite de-orbiting via controlled solar radiation pressure. Celest. Mech. Dyn. Astr. **126**(4), 433–459 (2016)
31. Deleflie, F., Rossi, A., Portmann, C., Metris, G., Barlier, F.: Semi-analytical investigations of the long term evolution of the eccentricity of Galileo and GPS-like orbits. Adv. Space Res. **47**, 811–821 (2011)
32. Ely, T.A., Howell, K.C.: Dynamics of artificial satellite orbits with tesseral resonances including the effects of luni-solar perturbations. Dyn. Stab. Syst. **12**(4), 243–269 (1997)
33. Earth Gravitational Model (2008). http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008/
34. Ferraz Mello, S.: Analytical study of the Earth's shadowing effects on satellite orbits. Celest. Mech. Dyn. Astron. **5**, 80–101 (1972)
35. Formiga, J.K.S., Vilhena de Moraes, R.: 15:1 Resonance effects on the orbital motion of artificial satellites. J. Aerosp. Techn Man. **3**(3), 251–258 (2011)
36. Friesen, L.J., Kessler, D.J., Zook, H.A.: Reduced debris hazard resulting from a stable inclined geo-synchronous orbit. Adv. Space Res. **13**(8), 231–241 (1993)
37. Froeschlé, C., Lega, E., Gonczi, R.: Fast Lyapunov indicators. Application to asteroidal motion. Celest. Mech. Dyn. Astr. **67**(1), 41–62 (1997)
38. Früh, C., Kelecy, T.M., Jah, M.K.: Coupled orbit-attitude dynamics of high area-to-mass ratio (HAMR) objects: influence of solar radiation pressure, Earth's shadow and the visibility in light curves. Celest. Mech. Dyn. Astr. **117**, 385–404 (2013)
39. Gachet, F., Celletti, A., Pucacco, G., Efthymiopoulos, C.: Geostationary secular dynamics revisited: application to high area-to-mass ratio objects. Celest. Mech. Dyn. Astr. **128**(2–3), 149–181 (2017)
40. Gaias, G., Ardaens, J.-S., Montenbruck, O.: Model of J_2 perturbed satellite relative motion with time-varying differential drag. Celest. Mech. Dyn. Astr. **123**(4), 411–433 (2015)
41. Galeş, C.: A cartographic study of the phase space of the restricted three body problem. Application to the Sun-Jupiter-Asteroid system. Commun. Nonlinear Sc. Num. Sim. **17**, 4721–4730 (2012)
42. Gedeon, G.: Tesseral resonance effects on satellite orbits. Cel. Mech. **1**(2), 167–189 (1969)

43. Gkolias, I., Daquin, J., Skoulidou, D.K., Tsiganis, K., Efthymiopoulos, C.: Chaotic transport of navigation satellites. Chaos **29**, 101106 (2019)
44. Gkolias, I., Colombo, C.: Towards a sustainable exploitation of the geosynchronous orbital region. Celest. Mech. Dyn. Astr. **131**(19) (2019)
45. Gkolias, I., Daquin, J., Gachet, F., Rosengren, A.J.: From order to chaos in Earth satellite orbits. Astron. J. **152**(5), 119 (2016)
46. Giacaglia, G.E.O.: A note on Hansen's coefficients in satellite theory. Celest. Mech. **14**, 515–523 (1976)
47. Guzzo, M., Lega, E.: Geometric chaos indicators and computations of the spherical hypertube manifolds of the spatial circular restricted three-body problem. Physica D **373**, 35–58 (2018)
48. Gondelach, D.J., Armellin, R., Wittig, A.: On the predictability and robustness of Galileo disposal orbits. Celest. Mech. Dyn. Astr. **131**, 60 (2019)
49. Hautesserres, D., Lara, M.: Intermediary LEO propagation including higher order zonal harmonics. Celest. Mech. Dyn. Astr. **127**, 505 (2017)
50. Hedin, A.E.: MSIS-86 thermospheric model. J. Geophys. Res. **92**, 4649–4662 (1986)
51. Hedin, A.E.: Extension of the MSIS thermosphere model into the middle and lower atmosphere. J. Geophys. Res. **96**, 1159–1172 (1991)
52. Henrard, J., Lemaitre, A.: A second fundamental model for resonance. Celest. Mech. **30**(2), 197–218 (1983)
53. Hubaux, Ch., Lemaitre, A.: The impact of Earth's shadow on the long-term evolution of space debris. Celest. Mech. Dyn. Astr. **116**, 79–95 (2013)
54. Hughes, S.: Earth satellite orbits with resonant lunisolar perturbations. I. Resonances dependent only on inclination. Proc. R. Soc. Lond. A **372**, 243–264 (1980)
55. Hughes, S.: Earth satellite orbits with resonant lunisolar pertubations. II. Some resonances dependent on the semi-major axis, eccentricity and inclination. Proc. R. Soc. Lond. A **375**, 379–396 (1981)
56. IADC-11-04, Apr. 2013, Space Debris IADC Assessment Report for 2010. https://www.iadc-home.org/
57. IADC-02-01, Revision 2, Mar. 2020, IADC Space Debris Mitigation Guidelines. https://orbitaldebris.jsc.nasa.gov/library/iadc-space-debris-guidelines-revision-2.pdf
58. ISO 27852:2016, Space systems—Estimation of orbit lifetime. https://www.iso.org/standard/68572.html
59. Jacchia, L.G.: Revised static models of the thermosphere and exosphere with empirical temperature profiles. Smithsonian Astrophysical Observatory, Science Report No. 332, Cambridge, MA (1971)
60. Jehn, R., Agapov, V., Hernadez, C.: End-of-Life disposal of Geostationary satellites. In: Proceedings of the Fourth European Conference on Space Debris, ESA SP587, **373J** (2005)
61. Johnson, N.L.: A new look at the GEO and near-GEO regimes: operations, disposals, and debris. Acta Astronautica **80**, 82–88 (2012)
62. Kaula, W.M.: Development of the lunar and solar disturbing functions for a close satellite. Astron. J. **67**, 300–303 (1962)
63. Kaula, W.M.: Theory of Satellite Geodesy. Blaisdell Publ. Co. (1966)
64. Kessler, D.J., Cour-Palais, B.G.: Collision frequency of artificial satellites: the creation of a debris belt. J. Geophys. Res. **83**(A6), 2637–2646 (1978)
65. King-Hele, D.G.: The effect of the Earth's oblateness on the orbit of a near satellite. Proc. R. Soc. Lond. A **247**, 49–72 (1958)
66. Klinkrad, H.: Space Debris: Models and Risk Analysis. Springer-Praxis, Berlin-Heidelberg (2006)
67. Krisko, P.H., Hall, D.T.: Geosynchronous region orbital debris modeling with GEO, EVOLVE 2.0. Adv. Space Res. **34**, 1166–1170 (2004)
68. Lane, M.T.: On analytic modeling of lunar perturbations of artificial satellites of the Earth. Celest. Mech. Dynam. Astr. **46**(4), 287–305 (1989)
69. Lemaitre, A., Delsate, N., Valk, S.: A web of secondary resonances for large A/m geostationary debris. Celest. Mech. Dyn. Astr. **104**, 383–402 (2009)

70. Lhotka, C., Celletti, A., Galeş, C.: Poynting-Robertson drag and solar wind in the space debris problem. Mon. Not. Roy. Ast. Soc. **460**, 802–815 (2016)
71. Liu, J.J.F., Alford, R.L.: Semianalytic theory for a close-Earth artificial satellite. J. Guid. Control **3**(4), 304–311 (1980)
72. Musen, P., Bailie, A., Upton, E.: Development of the lunar and solar perturbations in the motion of an artificial satellite. NASA Tech. Note D–494, 40 (1961)
73. Pardini, C., Anselmo, L.: Long-Term evolution of geosynchronous orbital debris with high area-to-mass ratios. Trans. Jpn. Soc. Aero. Space Sci. **51**(171), 22–27 (2008)
74. Petit, A., Casanova, D., Dumont, M., Lemaitre, A.: Dynamical lifetime survey of geostationary transfer orbits. Celest. Mech. Dyn. Astron. **130**(79) (2018)
75. Radtke, J., Dominguez-Gonzalez, R., Flegel, S.K., Sanchez-Ortiz, N., Merz, K.: Impact of eccentricity build-up and graveyard disposal Strategies on MEO navigation constellations. Adv. Space Res. **56**, 2626–2644 (2015)
76. Rosengren, A.J., Alessi, E.M., Rossi, A., Valsecchi, G.B.: Chaos in navigation satellite orbits caused by the perturbed motion of the Moon. Mon. Not. R. Astron. Soc. **449**, 3522–3526 (2015)
77. Rosengren, A.J., Amato, D., Bombardelli, C., Jah, M.K.: Resident space object proper orbital elements. AAS 19–557 (2019)
78. Rosengren, A.J., Daquin, J., Tsiganis, K., Alessi, E.M., Deleflie, F., Rossi, A., Valsecchi, G.B.: GALILEO disposal orbit strategy: resonances, chaos and stability. Mon. Not. R. Astron. Soc. **464**(4), 4063–4076 (2017)
79. Rosengren, A.J., Scheeres, D.J.: Long-term dynamics of high area-to-mass ratio objects in high-Earth orbit. Adv. Space Res. **52**, 1545–1560 (2013)
80. Rosengren, A.J., Scheeres, D.J., McMahon, J.W.: The classical Laplace plane as a stable disposal orbit for geostationary satellites. Adv. Space Res. **53**(8), 1219–1228 (2014)
81. Rossi, A.: Resonant dynamics of Medium Earth Orbits: space debris issues. Celest. Mech. Dyn. Astr. **100**, 267–286 (2008)
82. Rossi, A., Valsecchi, G.B.: Collision risk against space debris in Earth orbits. Celest. Mech. Dyn. Astron. **95**, 345–356 (2006)
83. Rossi, A., Valsecchi, G.B., Farinella, P.: Collision risk for high inclination satellite constellations. Planet. Space Sc. **48**, 319–330 (2000)
84. Sanchez, D.M., Yokoyama, T., de Almeida Prado, A.F.B.: Study of some strategies for disposal of the GNSS satellites. Math. Probl. Eng. **2015**, Article ID 382340, 14 pages (2015)
85. Sampaio, J.C., Neto, A.G.S., Fernandes, S.S., Vilhena de Moraes, R., Terra, M.O.: Artificial satellites orbits in 2:1 resonance: GPS constellation. Acta Astronautica **81**, 623–634 (2012)
86. Schettino, G., Alessi, E.M., Rossi, A., Valsecchi, G.B.: A frequency portrait of Low Earth Orbits. Celest. Mech. Dyn. Astron. **131**(35) (2019)
87. Skoulidou, D.K., Rosengren, A.J., Tsiganis, K., Voyatzis, G.: Dynamical lifetime survey of geostationary transfer orbits. Celest. Mech. Dyn. Astron. **130**(77) (2018)
88. Valk, S., Delsate, N., Lemaitre, A., Carletti, T.: Global dynamics of high area-to-mass ratios geosynchronous space debris by means of the MEGNO indicator. Adv. Space Res. **43**, 1509–1526 (2009)
89. Valk, S., Lemaitre, A.: Analytical and semi-analytical investigations of geosynchronous space debris with high area-to-mass ratios. Adv. Space Res. **41**, 1077–1090 (2008)
90. Valk, S., Lemaitre, A., Anselmo, L.: Semi-analytical investigations of high area-to-mass ratio geosynchronous space debris including Earth's shadowing effects. Adv. Space Res. **42**, 1429–1443 (2008)
91. Valk, S., Lemaitre, A., Deleflie, F.: Semi-analytical theory of mean orbital motion for geosynchronous space debris under gravitational influence. Adv. Space Res. **43**, 1070–1082 (2009)
92. Zhu, T.L., Zhao, C.Y., Wang, H.B., Zhang, M.J.: Analysis on the long term orbital evolution of Molniya satellites. Astrophys. Space Sci. **357**, 126 (2015)
93. https://www.esa.int/Safety_Security/Space_Debris/Space_debris_by_the_numbers
94. http://spacelaunchreport.com/index.html

The Unaccomplished Perfection of Kepler's World

Antonio Giorgilli

Abstract A short walk trough the main works of Kepler, notably the *Astronomia Nova*, trying to follow his search of the perfection of the World till the discovery of his celebrated laws. At the end of the road, the consciousness that the finish line had not yet been reached.

Keywords Celestial mechanics · Kepler's laws

1 Apology

It may be suprising to read a contributed paper like the present one among the proceedings of a school on contemporary methods of Celestial Mechanics. It provides an answer to a question that happened to me to ask myself after many years spent in teaching Mechanics to students, and in doing theoretical research on Dynamical Systems, including Classical and Celestial Mechanics:

How did Kepler discover his celebrated laws?

The question is quite simple; the answer is not that simple. One can hardly find a page in Kepler's works were the laws are stated together in a compact formulation, similar to what one can find in treatises or textbooks on Mechanics since at least a couple of centuries.

I felt embarrassed, and, *si parva licet componere magnis*, I could not resist the temptation to quote the incipit of the introduction of *Astronomia Nova*.

A. Giorgilli (✉)
Istituto Lombardo Accademia di Scienze e Lettere, Milano, Italy
e-mail: antonio.giorgilli@unimi.it

G. Baù et al. (eds.), *New Frontiers of Celestial Mechanics: Theory and Applications*,
Springer Proceedings in Mathematics & Statistics 399,
https://doi.org/10.1007/978-3-031-13115-8_9

Durissima est hodie conditio scribendi libros Mathematicos, præcipue Astronomicos. Nisi enim servaveris genuinam subtilitatem propositionum, instructionum, demonstrationum, conclusionum, liber non erit Mathematicus: sin autem servaveris, lectio efficitur morosissima, præsertim in Latina lingua, quæ caret articulis, & illa gratia quam habet græca, cum per signa literaria loquitur.

Very painful nowadays is the condition of writing mathematical books, particularly astronomical. For unless you maintain the innate rigor of propositions, constructions, proofs and conclusions then the book will not be mathematical; but if you respect that sequence then reading will be most laborious, especially in Latin language which lacks the articles and that gracefulness possessed by Greek when you communicate through written symbols.

The point is that I had many doubts on how to present the matter. The topic falls within the field of the history of science, and there are books devoted to the work of Kepler; e.g., the reader may see the comprehensive work of Julian B. Barbour [1], where he or she will find many further references. The problem is that I'm not an historian: I had to make a remarkable effort to read Kepler, and it took a long. I should warmly thank N. Swerdlow for directing me to Refs. [16] and [17], for kindly replying to my mails and also for sending me some preliminary work: without his help, I probably wouldn't have been able to find at least a partial answer to my question. The most natural thing to do would be to address the reader to the references above. But I hope that the present short exposition will be useful for people who are interested in the same question, but do not have enough time (may be not enough patience) to read long and exhaustive books.

Therefore I eventually decided to write the present note. It is nothing special, and I apologize in advance if the reader will find it scarcely useful. It is only a quick and dirty trace that may help as an introduction to the work of Kepler. I also present my apologies to true historians for this intrusion in their field. My only justification is that this note may satisfy or possibly awaken the curiosity of some of my friends astronomers.

2 Before Kepler

Astronomia Nova is written in a style that is definitely uncommon in mathematical books. It seems to be organized as a thorough time reconstruction of the long journey of Kepler through the lands of Astronomy. He mixes autobiographical pages with mathematical ones and with physical considerations that may appear bizarre to our eyes. Far from making a synthetic and organized exposition of his results, he describes all his attempts with plenty of details, including partial achievements, failures and moments of enthusiasm or frustration. It seems that, being aware of the definite revolutionary character of his discoveries, he wanted to make any effort in order to convince his contemporaries that his view was the right one.[1]

[1] This aspect is widely discussed by Stephenson [16].

Therefore it is convenient to proceed step by step by recapitulating the state of the art at the beginning of Kepler's work. This means that we should briefly recall the models of Ptolemæus, Copernicus and Tycho Brahe.

2.1 The Universe of Nicolaus Copernicus (1473–1543) Compared to Claudius Ptolemæus (c. 100–170)

There is no need here to devote a long discussion to the model of Ptolemæus. The model is a Geocentric one, with an elegant geometric arrangement. The Universe is composed by material spheres that enclose the Earth. The Earth remains fixed, and the spheres rotate around it with different velocities. The most external sphere is that of the fixed stars, which revolves with the highest velocity performing a complete revolution in a little less than one solar day. Inside the sphere of the fixed stars one finds the spheres of Saturn, Jupiter, Mars, Sun, Mercury, Venus and the Moon, which revolve with decreasing velocity. According to Copernicus there are different opinions concerning the order of Mercury and Venus, but this is not so relevant here.

The geometrical tools used by Ptolemæus are more relevant. The basic motion is the rotation of the spheres, because the circle is the perfect geometrical figure and the uniform rotation on a circle is the perfect natural motion of the heavens, according to the Aristotelian theory. The inequalities of the motion of the Planets are accounted for with three geometrical tools.

i. The *eccentric:* the center of the sphere is displaced with respect to the position of the Earth. This accounts partially for the non uniform angular motion of the Planet, in particular of the Sun.
ii. The *epicycle:* the Planet moves on a secondary sphere (the epicycle) that has its center on the primary sphere (the *deferent*). This accounts in particular for the retrograde motion of the Planets.
iii. The *equant point:* a point symmetric to the Earth's position with respect to the center. The motion of the Planet around that point appears to be uniform. Every Planet but the Earth have an equant point; more on this later.

The new model of Copernicus is developed in his main treatise *De revolutionibus orbium cœlestium.* As it is well known, the great novelty with respect to Ptolemæus is the rediscovery of the Heliocentric model that had been imagined by the Greek astronomers Philolaus of Croton (c. 470–385 BC) and Aristarchus of Samos (c. 310–230 BC), but had been abandoned. No original writings on that matter have been preserved. Copernicus knew about Aristarchus' theory thanks to some references found in writings of the biographer Plutarch (c. 46–119 AD) and of the philosopher Sextus Empiricus (c. 160–210 AD). Therefore Copernicus elaborated the geometric scheme of his model from scratch.

The plan of Copernicus is clearly made in seven *petitiones* that are declared in the *Commentariolus*, an unpublished manuscript that circulated among a few astronomers at that time. It is worth to look at them.

Prima petitio: *Omnium orbium cælestium sive sphaerarum unum centrum non esse.*	A unique common centre of all the celestial orbs or spheres does not exist.
Secunda petitio: *Centrum terrae non esse centrum mundi, sed tantum gravitatis et orbis Lunaris.*	The centre of the Earth is not the centre of the world, but only the centre of gravity and of the sphere of the Moon.
Tertia petitio: *Omnes orbes ambire Solem, tamquam in medio omnium existentem, ideoque circa Solem esse centrum mundi.*	All the spheres encircle the Sun, which appears to be in the middle of them all, so that the centre of the world is near the Sun.
Quarta petitio: *Minorem esse comparationem distantiarum Solis et terrae ad altitudinem firmamenti, quam semidimetientis terrae ad distantiam Solis, adeo ut sit ad summitatem firmamenti insensibilis.*	The ratio of the Earth–Sun distance to the height of the firmament is much smaller than the ratio of the Earth's radius to the Earth–Sun distance, to such a degree that the Earth-Sun distance is imperceptible compared with the height of the firmament.
Quinta petitio: *Quicquid ex motu apparet in firmamento, non esse ex parte ipsius, sed terrae. Terra igitur cum proximis elementis motu diurno tota convertitur in polis suis invariabilibus firmamento immobili permanente ac ultimo caelo.*	Whatever motion appears to be in the firmament is not a property of it, but of the Earth. Thus the whole Earth together with the elements close to it revolves around its fixed poles in a day–long revolution, while the firmament remains fixed, being the highest heaven.
Sexta petitio: *Quicquid nobis ex motibus circa Solem apparet, non esse occasione ipsius, sed telluris et nostri orbis, cum quo circa Solem volvimur ceu aliquo alio sidere, sicque terram pluribus motibus ferri.*	Whatever motion we observe in the Sun is due, not to its motion, but to the motion of the Earth and of our sphere, with which we revolve about the Sun, as any other Planet, and so the Earth undergoes many motions.
Septima petitio: *Quod apparet in erraticis retrocessio ac progressus, non esse ex parte ipsarum sed telluris. Huius igitur solius motus tot apparentibus in caelo diversitatibus sufficit.*	What appears in the Planets as retrograde and direct motion is due, not to their motion, but to the Earth's. Thus the motion of the Earth alone suffices to explain all apparent irregularities in the heaven.

The three first petitions make explicit the basic idea of Copernicus that the Earth is one of the Planets that revolve around the Sun, what everybody now knows (except for a few bizarre people who are still convinced that a flat Earth is the centre of the Universe). However the nowadays common belief that Copernicus places the Sun at the centre of the world is not quite correct: as stated by third petition the Sun is *near the centre*. The problem of the eccentricity of the orbits is still there, and Copernicus places the centre of the sphere of a Planet in the *mean Sun*, not the *true Sun*.

The fourth petition addresses the problem that had perhaps motivated the rejection of the heliocentric system by ancient astronomers: if the Earth revolves in an orbit around the Sun then we should observe the effect of that motion in the sky, i.e., the

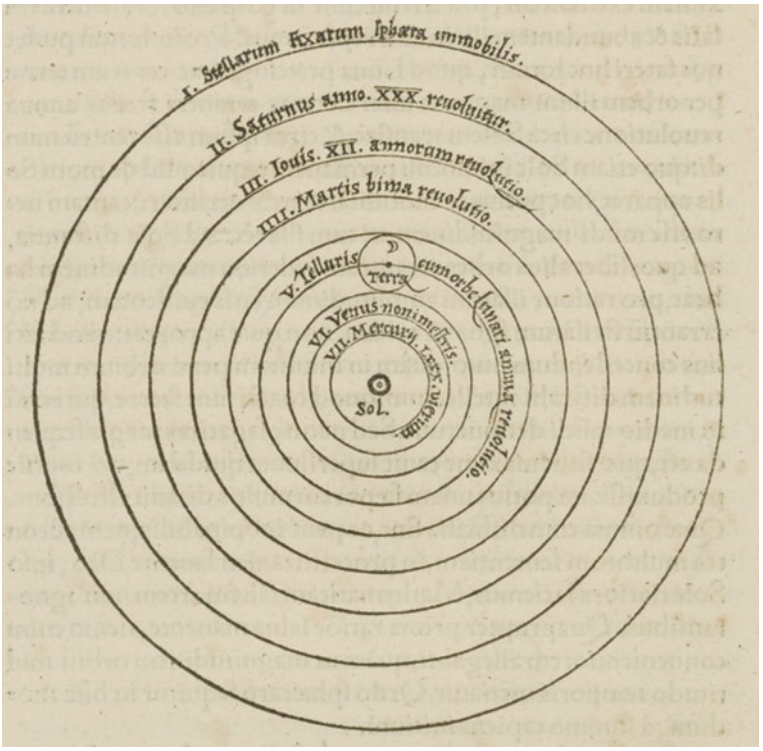


Fig. 1 The heliocentric model of Copernicus

phenomenon that we call *parallax*. The answer of Copernicus may appear obvious to us, but it was not so at his time: we can not observe the parallax because the radius of the sphere of fixed stars is exceedingly large with respect to the radius of the Earth's orbit. The parallax was indeed observed by Friedrich Bessel in 1838 [2], about three centuries later, using a heliometer.

The remaining three petitions claim that all inequalities that we observe in the sky, such as the annual motion of the Sun and the retrograde motion of the external Planets, are due to the motion of the Earth. Thus the celestial spheres are rearranged in a different order, as represented in Fig. 1, so as to put the Sun close to the centre. The Earth is placed between Venus and Mars, and the sphere of the Moon encircles the Earth. The highest sphere of stars is fixed, and the order of the planetary spheres is determined by the angular velocity of the revolution around the Sun: Saturn is the slowest one, Mercury the fastest. Thus the order of velocities is reversed with respect to the model of Ptolemæus.

Concerning the Earth, it is subject to three movements.

i. *Rotation:* explains the daily rotation of the Sun and of the stars.
ii. *Revolution:* explains the annual motion of the Sun with respect to the stars.
iii. *Declination:* an annual motion of the rotational axis of the Earth that explains the alternation of seasons.

The motion of declination may raise some perplexities: we never found it mentioned in our textbooks. The point is that Copernicus keeps the material spheres. Thus the Earth, as well as the other Planets, is fastened to a sphere that revolves with a period of one year. This causes the axis of rotation of the Earth to remain fixed with respect to the surface of the sphere, so that no change of season could occur. Therefore Copernicus introduces a motion of declination, namely a rotation of the axis of the Earth around the perpendicular to ecliptic with a period close to one year. This reminds the motion of the axis of the Earth that causes the precession of equinoxes, a phenomenon that had been discovered a long before by Hipparchus (c. 190–120 BC). The latter phenomenon may be justified in Copernicus' system by accepting a small difference between the period of revolution and that of declination. It was Galileo who pointed out that if the material spheres are removed and the rotation axis of the Earth is kept fixed but inclined with respect to the perpendicular to ecliptic then the alternation of the seasons is a straightforward consequence of the motion of revolution [5]. This is indeed what we read in nowadays textbooks.

The system of Copernicus differs from the ancient one also in two more points that are relevant for our discussion. The problem is to explain two main inequalities, namely: (i) the non uniform angular motion of the Planets with respect to the Sun, and (ii) the retrograde motion of the external Planets as seen from the Earth. As we have seen, the first problem is solved by Ptolemæus by introducing the equant point. Copernicus rejects that method, because he does not attribute any acceptable meaning to the equant. He rather introduces a pair of epicycles that produce the same effect. The second problem disappears by taking into account the revolution of the Earth: the retrograde motion occurs when the Sun and the Planet are in opposition, and is due to the angular velocity of the Earth being bigger than that of the external Planet. This fact is well illustrated in an image due to Galileo, Fig. 2.

2.2 The Universe of Tycho Brahe (1546–1601)

The model of Tycho Brahe, represented in Fig. 3, may be considered as the simplest one that represents the motion of the Planets as they are seen from the Earth. He refuses the heliocentric model of Copernicus, that he considers as an absurdity, as he claims in his inscription at the top of the figure. His simple but effective remark may be so stated: exchanging the deferent with the epicycle in the model of Ptolemæus makes things simpler, for the epicycle of Ptolemæus is nothing but the orbit of the Sun. Thus he claims that the Sun and the Moon revolve around the Earth, while all the other Planets revolve around the Sun.

His model has a main trouble: it conflicts with the existence of material spheres, because the sphere of Mars should intersect that of the Sun, as the figure clearly shows. Tycho rejects the spheres mainly on the basis of two arguments.

The first argument is concerned with the orbit of a great comet observed in 1577. With careful observations Tycho concluded that the comet was beyond the sphere of

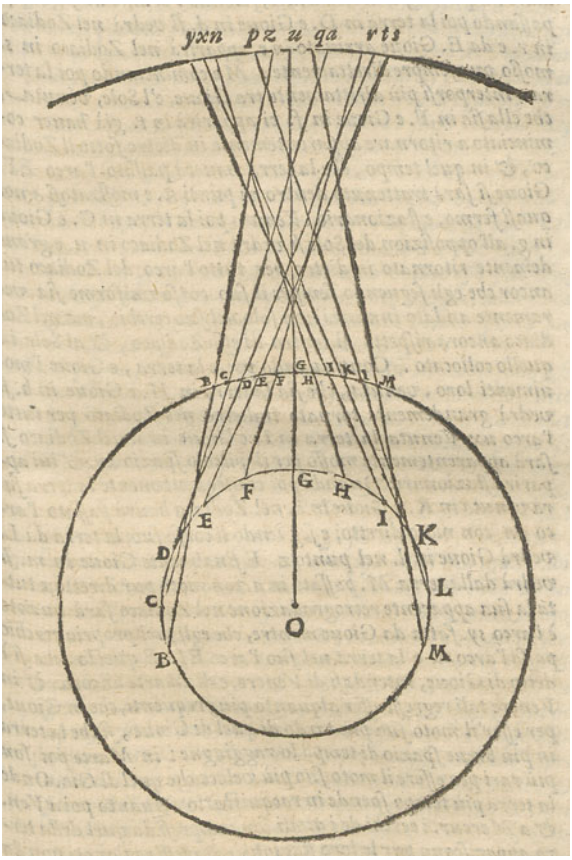


Fig. 2 The retrograde motion of the external Planets illustrated by Galileo ([5], giornata terza). Moving from the Sun at the centre one finds the orbit of the Earth, the orbit of the Planet and the fixed stars. When the Planet is close to opposition the projection of different position of the Planet on the sky as seen from the Earth appears as a retrograde motion

the Moon, and that the orbit had crossed the sphere of Venus. As to the first point we should not forget that in the representation of the world due to Aristotle the sphere of the Moon separates the lower region of the world from the upper one; the perfect motion of Sun, Planets and stars characterizes the upper region, while irregular motions and changing phenomena are bounded to the lower one. The comets are clearly changing objects, so they should be located inside the sphere of the Moon.[2] On the basis of his observations Tycho concluded not only that the comet was in the upper region, but also that the orbit was incompatible with the material nature of the sphere of Venus.

[2] It is mandatory here to recall the contributions of Galileo [5]. With his observations of the Moon and the discovery of satellites of Jupiter he had produced substantial support to his thesis that the Mechanics and the Physics of the Universe is the same as the Mechanics and the Physics on the Earth; an idea that was far for being accepted at his time and provided a basis for the achievements of Newton. The long discussion devoted by Galileo to the problem of the location of comets in "giornata terza" of the *Dialogo* is a direct consequence of his thesis.

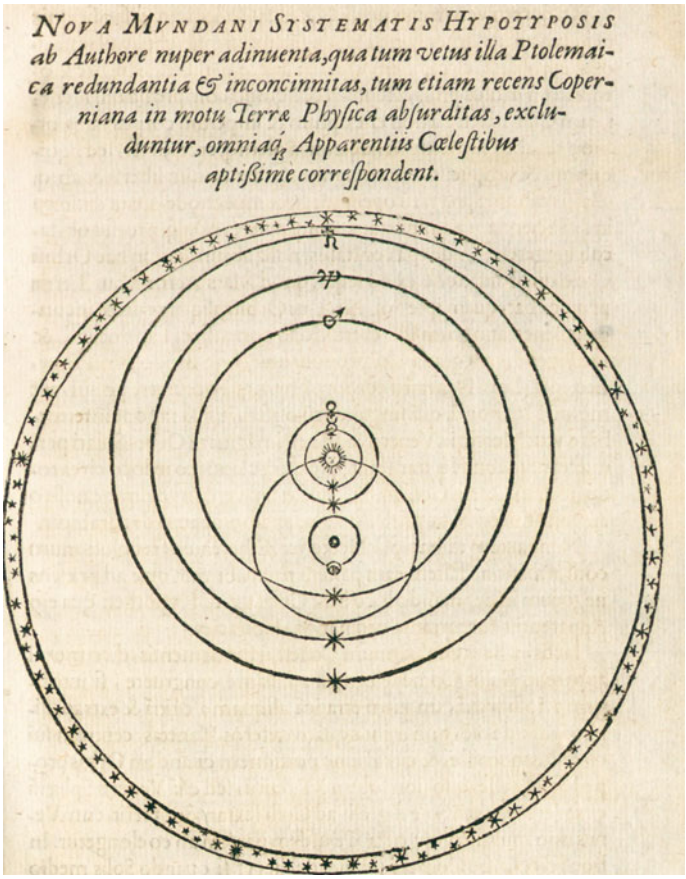


Fig. 3 The model of Tycho Brahe. The inscription on the top of the image says: *New hypothesis for the system of the world, recently proposed by the author, by which the old redundancy and inelegance of Ptolemæus and the recent absurdity of Copernicus concerning the motion of the Earth are excluded, and which is in excellent agreement with the Celestial Appearance.* (Figure from *De mundi ætherei recentioribus phænomenis*)

The second argument is concerned precisely with the sphere of Mars. Again, Tycho makes accurate observations in order to determine the distance of Mars from the Earth at opposition with the Sun. In 1582 he makes the conclusion that Mars at opposition is closer to the Earth than the Sun, while the opposite must obviously occur at conjunction. Thus he rejects the hypothesis that the spheres are material objects.

The major impact of the astronomical work of Tycho Brahe is connected both with observations and with the compilation of new astronomical tables. The extraordinary ability of Tycho in constructing new instruments for determining the position of celestial objects produced a substantial improvement of astronomical data. Visual

observations were affected by an error that may be evaluated in $10'$. Tycho could reduce the error to $1'$ or $2'$: such an improvement had a crucial impact on the calculations of Kepler. Moreover in 1600 Tycho Brahe convinced Kepler to move to Prague in order to collaborate to the compilation of the new tables. Kepler had initially the assignment of working on the orbit of Mars, that Plinius had named "inobservabile sidus" (the star impossible to trace). This, as narrated by Kepler himself in chapter VII of *Astronomia Nova*, is the first act in the long standing war of Kepler with Mars—the god of war, indeed.

3 The Discovery of Kepler's Laws

The guiding idea of most of the work of Johannes Kepler (1571–1630) is identified in an ambitious, deep question that he raises:

To discover the plan of God when He created the Universe.

The first attempt is found in the first great treatise *Mysterium Cosmographicum (1596)*, but it was only some twenty years later that Kepler could reach a conclusion; we read it in *Harmonice Mundi (1619).* The treatise *Astronomia Nova (1609)* is placed in between.

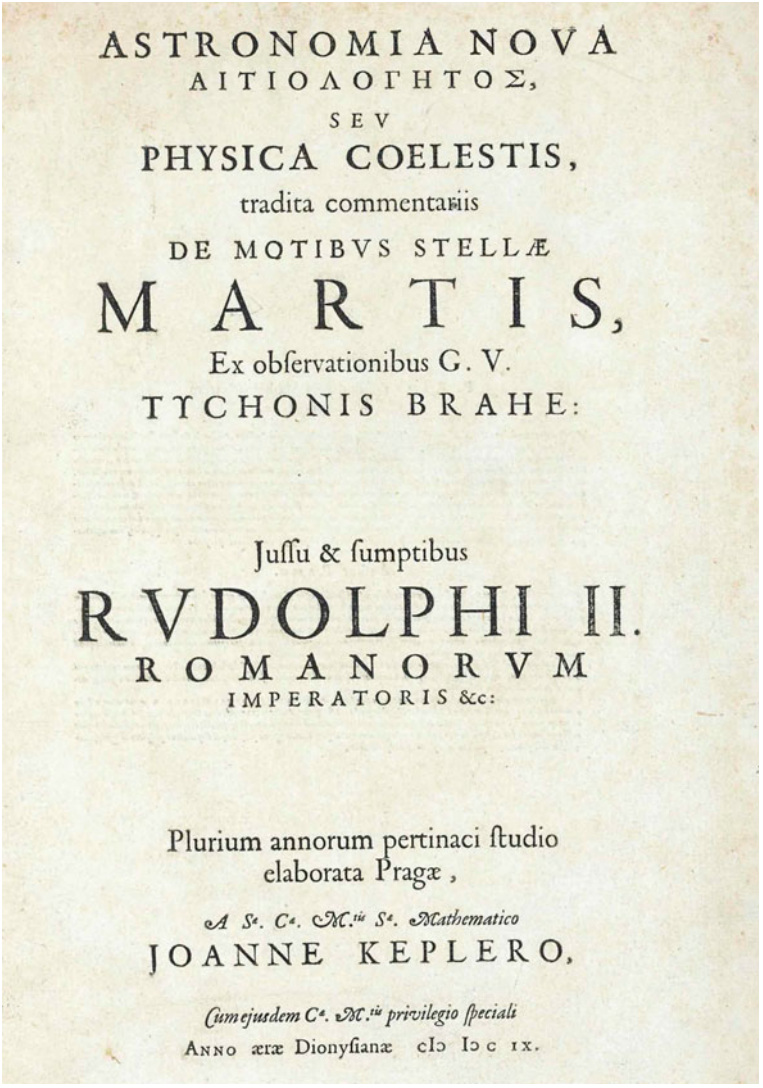
Kepler's laws that are the central subject of the present note are to be found in *Astronomia Nova* (first and second law) and in *Harmonice Mundi* (third law). However, in order to understand the revolutionary work performed by Kepler it is convenient to spend a few words on the first part of *Astronomia Nova*. Already in the title Kepler emphasizes his intention to investigate the *causes* of the motions of the Planets on the basis of a *physical model* (see Fig. 4).

3.1 Ad Imitationem Veterum

The first part of *Astronomia nova* fixes the model used by Kepler. For physical reasons the heliocentric model of Copernicus in preferred, but Kepler introduces some relevant differences.

The first point is that from a geometric viewpoint the three previous models of Ptolemæus, Copernicus and Tycho Brahe can be considered as equivalent. We may say that Kepler combines elements from the three models in a way to construct a more satisfactory basis.

Kepler, in agreement with Tycho Brahe, rejects the existence of solid spheres. Thus all ancient models are deprived of *physical* meaning: the Planets move in the ethereal space, but there is no physical reason that justifies the complicated geometric combinations of eccentrics, epicycles and equants. The intuition of Kepler is that the sole object that may offer a physical interpretation is the Sun.


ASTRONOMIA NOVA
ΑΙΤΙΟΛΟΓΗΤΟΣ,
SEV
PHYSICA COELESTIS,
tradita commentariis
DE MOTIBVS STELLÆ
MARTIS,
Ex obſervationibus G. V.
TYCHONIS BRAHE:

Juſſu & ſumptibus
RVDOLPHI II.
ROMANORVM
IMPERATORIS &c:

Plurium annorum pertinaci ſtudio
elaborata Pragæ,

A Sæ. Cæ. M.tis Sæ. Mathematico
JOANNE KEPLERO,

Cum ejusdem Cæ. M.tis privilegio ſpeciali
Anno æræ Dionyſianæ cIↃ IↃ c ix.

Fig. 4 The front page of *Astronomia nova*. The title puts the accent on the search of the author for the *causes* (αἰτιολογητός), namely for a *physical astronomy* (physica cœlestis)

Here comes the first successful conclusion: *using the Sun as reference point the orbits of all Planets turn out to be planar, and all planes have a common point in the Sun.* This is also the first main difference with respect to the Copernican system.

The second point is that Kepler proceeds *ad imitationem veterum* (imitating the ancients) by restoring the equant point in place of the double epicycle of Copernicus.

Both are geometrically equivalent, and both are unacceptable in a physical scheme, but using the equant makes things easier. Anyway both the equant and the epicycle remain a mere geometrical artifice.

After the first attempts to calculate the orbit of Mars with the ancient methods Kepler reverts to consider the motion of the Earth (equivalently: the motion of the Sun with respect to the Earth). Assuming the orbit to be circular, which is perfectly reasonable in view of the low eccentricity, he uses a clever choice of the observations of Mars by Tycho in order to calculate the orbit of the Earth with good precision. Doing so, he realizes that *also the Earth should be assigned an equant*, contrary to the models of Ptolemæus (for the Sun) and Copernicus. Here the problem of *bisection of eccentricity* shows up. In order to agree with the geometry of Ptolemæus the centre of the eccentric circle of the orbit of a Planet should be placed halfway between the equant and the Sun; this produces a satisfactory description of the non uniform motion on the deferent. Among his attempts Kepler also tries to replace the criterion of bisection with what he calls a *vicarious hypothesis*: he introduces a ratio 8 : 5 between the distances of the centre from the equant and the Sun. After a lot of calculations he concludes that the bisection of eccentricity reproduces the correct form of the orbit, hence it is the preferable one. On the other hand the vicarious hypothesis produces correct angles, but wrong distances. Hence he provisionally adopts the vicarious hypothesis as a useful tool for calculating the angles.

3.2 Towards the "Law of Areas"

Still focusing attention on Earth's orbit Kepler comes to a more accurate investigation of the observed fact that the velocity at perihelion is larger than at aphelion. His starting point is so stated.

Primum fit inductio: omnes in omnino Planetas uti Æquante circulo, seu bisectione Eccentricitatis puncti Æquatorii.	First the following precept is made: all Planets possess an equant circle, namely the bisection of the eccentricity of the equalizing point.

However, as we have seen, he does not attribute any physical meaning to the equant point. Here comes the first claim that Kepler recognizes as a milestone. He puts strong emphasis on the physical meaning of his principle.

Super hoc principium Geometrica demonstratione instruitur universale hoc, Moras Planetæ in æqualibus arcubus Eccentrici proportionari cum discessu Planetæ a puncto, unde consurgit Eccentricitas. Arrigite aures Physici, hic enim deliberatio suscipitur de impressione in vestram provinciam facienda.

On this precept the following universal principle is stated, through a geometrical demonstration: the time spent by the Planet on equal arcs is proportional to the distance of the Planet from the point where the eccentricity arises. Please keep wide open your ears, you Physicists, because here is the origin of the initiative to make an intrusion into your area.

Finally the main statement is declared, thus referring the origin of the non uniform angular velocity only to the Sun, with no need of the equant point.

Primum sciat in omni hypothesi Ptolemaica hac forma instructa, quantacunque eccentricitas fuerit, celeritatem in perihelio & tarditatem in aphelio proportionari quam proxime lineis ex centro mundi eductis in Planetam.

First, [the reader] should know that in whatever hypothesis that agrees with this Ptolemaic form, independently of the eccentricity, the rapidity at perihelion and the slowness at aphelion exhibit *quam proxime* a proportion with the lines drawn from the centre of the world to the Planet.

It must be stressed that no reference to areas is made here: only the distance from the Sun matters.

It may be interesting to briefly follow the proof of Kepler ([8], Chap. XXXII): this will provide an idea of the methods and of the language of *Astronomia Nova*. Let us make reference to Fig. 6, which is a simplified version of the original image Fig. 5. We consider two opposite arcs AB and CD of the orbit that are seen from the Sun under the same angle, and have initial point on the aphelion and the perihelion, respectively. The time spent on the arcs (called *mora* by Kepler) is represented by the arcs FG and HK, respectively. Assuming that the arcs AB and CD are small we approximate them with segments orthogonal to the line of apsides CF, and do the same for the arcs FG and HK. Then the sector SAB is similar to SCD, the sector EAB is similar to EFG, and the sector EHK is similar to ECD. By Euclidean geometry we get the relations

$$\frac{|AB|}{|FG|} = \frac{|EA|}{|EF|}\,, \quad \frac{|HK|}{|CD|} = \frac{|EH|}{|EC|}\,. \tag{1}$$

Notice that the right members contain only points on the line of apsides; therefore the quantities on the left do not depend on the choice of the points B and D, and moreover the initial choice that the arcs AB and CD are seen from the Sun under the same angle is not necessary.

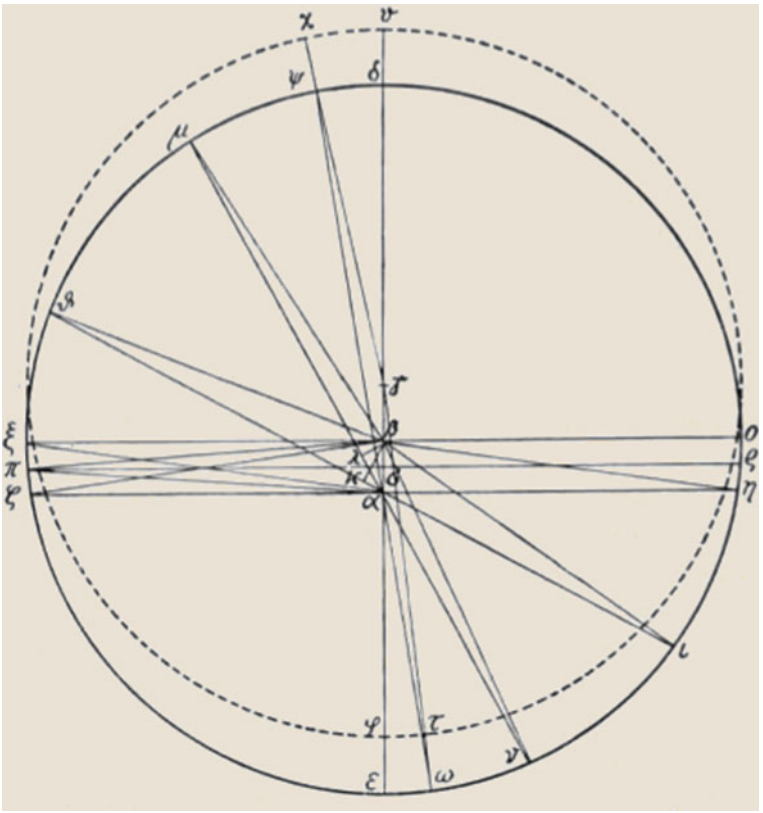


Fig. 5 The image of Kepler illustrating his discovery that the time spent by the Planet on equal arcs is proportional to the distance of the Planet from the Sun

Now, Kepler wants to prove that the relations

$$\frac{|AB|}{|FG|} \simeq \frac{|OA|}{|SA|}\,, \quad \frac{|HK|}{|CD|} \simeq \frac{|SC|}{|OC|} \tag{2}$$

hold true *quam proxime*, i.e., in good approximation. Remark that the latter relations allow us to focus attention on the distances $|SA|$ and $|SC|$ from the Sun, thus forgetting the distances from the equant point. Comparing the right members of Eqs. (1) and (2) and in view of $|OA| = |EG| = |OC| = |EH|$ we see that Kepler's relation (2) is true provided the relation

$$|OA|^2 \simeq |EA|\,|SA|\,, \quad |OC|^2 \simeq |EC|\,|SC|\,, \tag{3}$$

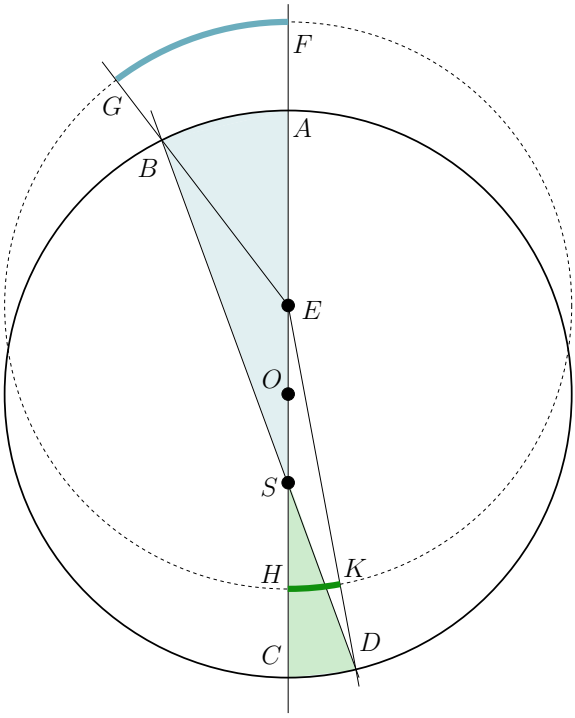


Fig. 6 Illustrating the proof of the inverse proportionality between the distance from the Sun and the time spent in an arc of the orbit. The labels of the relevant points are: S, Sun; E, Equant; O, centre of the eccentric orbit; CF, line of apsides. The eccentricity is $e = |SO|/r$, where $r = |OA|$ is the radius of the eccentric circle. Some obvious relations are $|OA| = |OC| = |EF| = |EH| = r$, $|EA| = |OH| = |SC| = r(1-e)$, and $|SA| = |EC| = |OF| = r(1+e)$

holds true, i.e., that the radius $|OA| = |OC|$ of the eccentric circle is the geometric mean between the segments on the right hand side of the equalities. This is untrue, because from the figure we get only the equalities between the arithmetic means

$$|OA| = \frac{|EA| + |SA|}{2}\,, \quad |OC| = \frac{|EC| + |SC|}{2}\,. \tag{4}$$

Here Kepler proposes a clever approximation. His remark is that if the difference between the segments $|EA|$ and $|SA|$ ($|EC|$ and $|SC|$ respectively) is small then the difference between the arithmetic and the geometric means becomes negligible. If we do the calculations with algebra (that was not known to Kepler) we say the same by remarking that the arithmetic mean between $1+e$ and $1-e$ (where e is the small eccentricity) is well approximated by the geometric mean $\sqrt{1-e^2} \simeq 1 - e^2/2 \simeq 1$ with an error $O(e^2)$. Therefore we may consider the relations (2) as true, and by multiplication we get

$$\frac{|AB|}{|CD|} \cdot \frac{|HK|}{|FG|} = \frac{|SC|}{|SA|}\,.$$

Finally, by recalling that the choice of the points B and D is arbitrary we may set $|AB| = |CD|$, so as to obtain the really interesting relation

$$\frac{|HK|}{|FG|} = \frac{|SC|}{|SA|},$$

namely that the *mora* on the arcs is proportional *quam proxime* to the distance from the Sun, as Kepler wanted to prove. For instance, let us consider the case of the Earth. Based on Tycho's observations the eccentricity is evaluated to be $e \sim 1.8 \times 10^{-2}$. If the difference is estimated to be e^2 we get an error $\sim 3.24 \times 10^{-4}$, a very small quantity; actually the evaluation of the next section shows that we could divide the difference by a further factor 4.

The argument of Kepler is not fully rigorous: it does not apply directly to any point of the orbit, because the arc is no more orthogonal to the diameter of the circle. Kepler is well aware of this fact, but he says that *tenuissima apparet diversitas* (the difference is clearly very small), with some elaborated justifications.

We could merely conclude that what Kepler has actually seen is that the uniform motion with respect to the equant point is *quam proxime* equivalent to the claim that the *mora* on small arcs is proportional to the distance from the Sun. But we should consider the question from Kepler's viewpoint: he looks for a model that has a *physical meaning*. Now, the equant point is merely a useful geometric artifice: there is nothing there that can exert a physical action on a Planet. On the other hand there is no *a priori* justification for using the equant: there is only a long standing tradition and a general agreement among astronomers for using it. We may instead attribute the Sun the ability of exerting a physical action. This was a challenging question for Kepler. In rough terms, what he imagines is that the Sun rotates around its axis, and drags the Planets around the orbit by exerting an action that maintains the motion; he was unaware of the principle of inertia.[3] Therefore Kepler considers his principle as preferable with respect to the use of an equant.

The problem now is: *to calculate the time spent on an arc using the principle established by Kepler*. This is a difficult matter: we could say that as a geometrical method the equant is easier to use. Kepler's idea is that we may consider an arc as composed by a sequence of small arcs to which the principle can be applied; an attempt that he works out with a lot of patience, because the actual calculation turns out to be very hard. We should not forget that Kepler could not rely on integral calculus, not yet invented. But he introduces a further ingenious idea: by a rather elaborated argument he gets convinced that the sum on infinitely many distances may be replaced by the area of the sector filled in by the distances. This is essentially the seed of the second Kepler's law, though till now Kepler seems to consider also the area as a geometrical tool: his point remains that the relevant fact is the proportionality between the time and the radius, as stated at the beginning of the present section.

[3] As to the physical nature of the action of the Sun, in Chap. XXXIII Kepler makes a long discussion invoking, e.g., a magnetic force that propagates in a similar way to light.

Our usual statement of the *law of areas*, that we call second Kepler's law, has been enunciated by Newton (*Principia* [15], Sect. II, Proposition I theorema I).

3.3 Kepler's Principle Revisited

The geometric argument of Kepler, as we see, is ingenious, but perhaps unfamiliar for us. We may reexamine it forcing the use of areas and exploiting differential calculus, just in order to see how the calculation can be simplified by our formalism (and perhaps to understand it better).

Let us refer to Fig. 7. Consider the arcs $A'A''$ and $P'P''$ seen from the equant point under the same angle $\delta\theta$. Then the length of the arcs is

$$|A'A''| = (1-e)a\delta\theta \ , \quad |P'P''| = (1+e)a\delta\theta \ ,$$

respectively. Consider now the sectors $SA'A''$ and $SP'P''$ with vertex in the Sun. We have

$$\left.\begin{aligned} \text{area}(SP'P'') &= \frac{1}{2}|P'P''|\cdot|SP| \\ \text{area}(SA'A'') &= \frac{1}{2}|A'A''|\cdot|SA| \end{aligned}\right\} = \frac{1}{2}(1-e)(1+e)a^2\delta\theta.$$

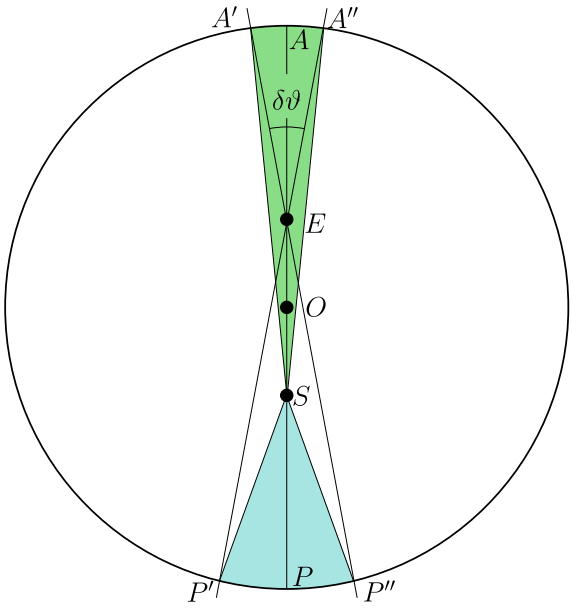


Fig. 7 Revisiting Kepler's proof in terms of areas. The labels of the relevant points are: S, Sun; E, Equant; O, centre of the eccentric orbit; P, perihelion; A, aphelion; AP, line of apsides. The relevant quantities are $|OA| = |OP| = a$, the semimajor axis; $|OS| = |OE| = ea$; $|SP| = |EA| = (1-e)a$; $|SA| = |EP| = (1+e)a$

Denoting now by $\delta\varphi$ the angle of either arc AB or CD as seen from the Sun and by r the distance from the Sun we may write the length of the arc as $\delta s = r\,\delta\varphi$; thus we recover the claim of Kepler in either form

$$\frac{\delta s}{\delta t} = \frac{1}{r} \quad \text{or} \quad \delta t = r\,\delta s\ .$$

The argument reproduces that of Kepler in the previous section, including its defect: it is perfect for apsides; only approximate for all other points. So let us recast it in a more general form, using nothing more than Euclidean geometry.

Let us refer to Fig. 8. The angle θ taken from the equant evolves uniformly, and determines the line EP that gives the position P of the Planet on the circle with

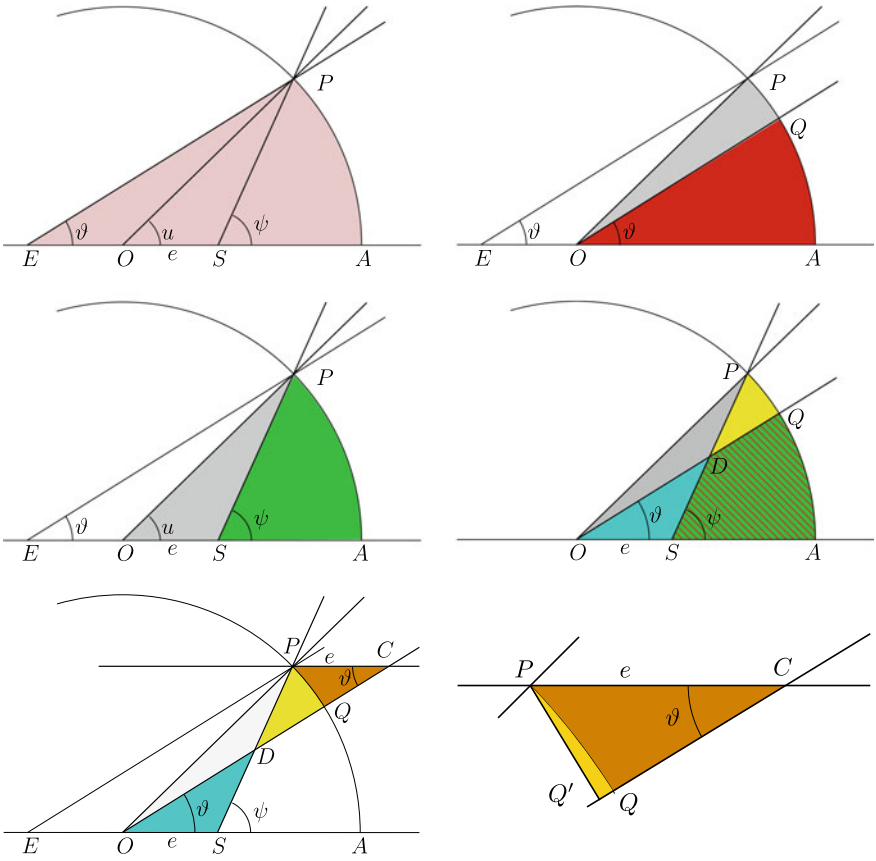


Fig. 8 Comparing the uniform motion around the equant point and the law of areas. The labels of the relevant points are: S, Sun; E, Equant; O, centre of the eccentric orbit; A, perihelion; u, eccentric anomaly; ψ, true anomaly. θ is the angle measured from the equant, which evolves uniformly in time according to Ptolemæus

centre O (upper left panel). Translate the angle θ so that its vertex is in O (upper right panel). Then the area of the sector QOA evolves uniformly in time, thus replacing the angle θ. Taking the Sun as the reference point we look instead to the area of the sector PSA (middle left panel) which is swept by the radius joining the Sun with the Planet. We want to compare the areas of the two sectors, and show that they are *quam proxime* equal. Let us superimpose the two areas (middle right panel); forgetting the common part $ASDQ$, we should compare the areas of the triangles ODS and PDQ.

Drawing the line PC parallel to OA (lower left panel) we construct the triangle PDC, whose area is clearly equal to the area of the triangle ODS; this is just Euclidean geometry. Therefore the difference between the areas is the small curvilinear triangle PQC. The latter triangle is enlarged in the lower right panel, and the line PQ' is drawn orthogonal to QC. The resulting right triangle has sides $|PC| = e$, $|PQ'| = e\sin\theta$ and $|Q'C| = e\cos\theta$; hence we get $\text{area}(DSO) - \text{area}(DPQ) \sim \frac{e^2 \sin 2\theta}{4}$. We conclude

$$\left|\text{area}(SAP) - \text{area}(OAQ)\right| \sim \frac{e^2 \sin 2\theta}{4} .$$

This shows that for small eccentricity the uniform motion around the equant and the uniform evolution of the area swept by the radius from the sun are *quam proxime* equivalent, with a difference not exceeding $e^2/4$.

3.4 The Elliptic Orbit of Mars, the Inobservabile Sidus

Coming to Mars, a careful examination of the observational data convinces Kepler that the method of equants produces the correct angles with respect to the centre of the orbit, but fails to produce the correct distances. This is a major trouble, widely narrated in part IV of *Astronomia Nova*, starting with Chap. XLI.

After a long examination Kepler finds that the orbit exhibits an oval shape, represented by the dashed curve on the left side of Fig. 9. The true orbit and the circle differ by a *lunula* (tiny moon) that appears as generated by an oscillation along the diameter of an epicycle, in a direction orthogonal to the line of apsides; the epicicle is represented in the lower part of Fig. 10. The latter fact conflicts with any physical interpretation. The hard problem for Kepler is to justify the existence and the width of that *lunula*. He seems to suspect that the true shape could be an ellipse, but lacks a convincing justification for such a claim.

After many attempts, the breakthrough is reached in Chap. LVI. Here the description takes on an enthusiastic tone. Kepler's considerations refer to Fig. 10.

Cum igitur duobus argumentis (. . .) non obscure colligerem, lunulæ illius latitudinem dimidiam tantum assumendam, scilicet 429, *correctius* 432, (. . .); *cœpi de causis & modo cogitare, quibus tantæ latitudinis lunula rescinderetur.*
Qua in cogitatione dum versor anxie, . . . forte fortuito incido in secantem anguli 5°18′ *quæ est mensura æquationis Opticæ maximæ. Quem cum viderem esse* 100429, hic quasi e somno expergefactus, & novam lucem intuitus, *sic cœpi ratiocinari.*

Thus, having clearly concluded, from two different arguments, (. . .) that the width of that lunula should be halved, i.e., 429, or more correct 432, (. . .); I began to thoroughly investigate how and why such a wide lunula should be subtracted.
While reflecting with anxiety, by fortuitous chance I fall on the secant of the angle 5°18′, which is the maximal amplitude of the optical equation. When I saw it to be 100429, *it was like being suddenly awakened from sleep, and seeing a new light.* Then I began to argue as follows.

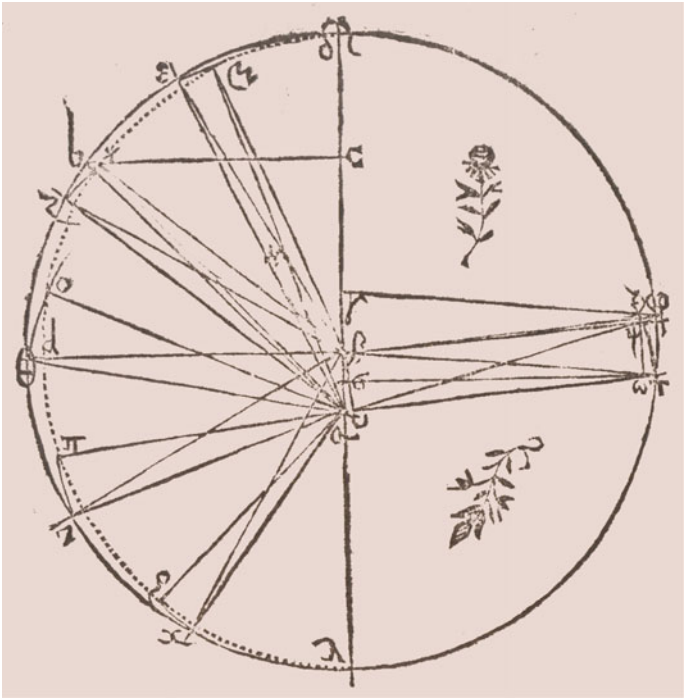


Fig. 9 The image of Kepler showing the deviation of the orbit of Mars from a circle: the *lunula* between the circle and the observed orbit

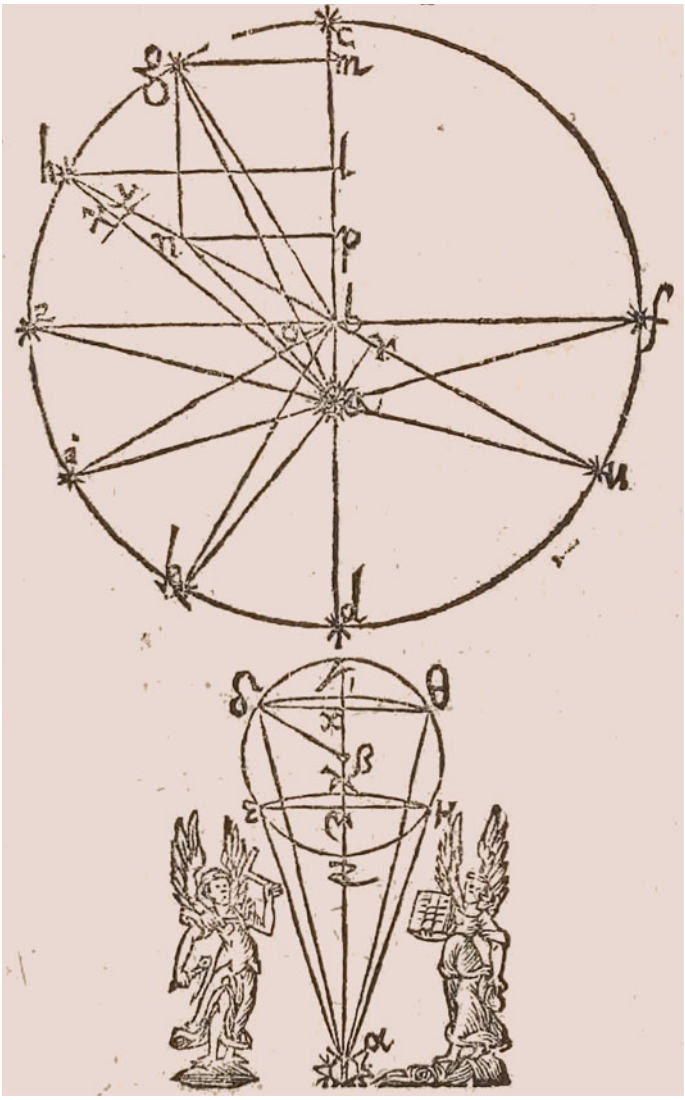


Fig. 10 The image by which Kepler illustrates his discovery of the true shape of the orbit, as quoted in the text. Figure 11 provides a simplified version. The lower part of the figure represents the epicycle that is conjectured to provide the width of the lunula, taking the projection on the horizontal diameter

As a personal note, I felt impressed by the strength of Kepler's expressions. *"Forte fortuito"* stresses the quite unexpected casualty of the crucial lucky event; *"quasi e somno expergefactus"* puts a strong accent on the surprise that strikes the mind of Kepler, like a thunderbolt.

In longitudinibus mediis, æquationis pars Optica fit maxima. In longitudinibus mediis lunula seu curtatio distantiarum est maxima, etque tanta, quantus est excessus secantis æquationis opticæ maximæ 100429 *supra radium* 100000. *Ergo si pro secante usurpetur radius in longitudine media, efficitur id, quod suadent observationes.*

Close to the average value of the longitude the optical equation is close to a maximum. The amplitude of the lunula takes a maximum there, and is the same as the excess of the secant of the optical equation, namely 100429 over the radius 100000. Therefore, if for the mean longitude you replace the secant with the radius the result of the observations will be recovered.

Et in schemate cap. XL conclusi generaliter, si pro HA usurpes HR, pro VA vero VR, & pro EA substituas EB, & sic in omnibus, fiet idem in locis cæteris eccentrici, quod hic factum est in longitudinibus mediis.

According to the scheme of Chap. XL, I made the general conclusion that if you replace HA with HR, VA with VR and EA with EB, and similarly for all points, the same that occurred for average longitude will happen at any point of the eccentric circle.

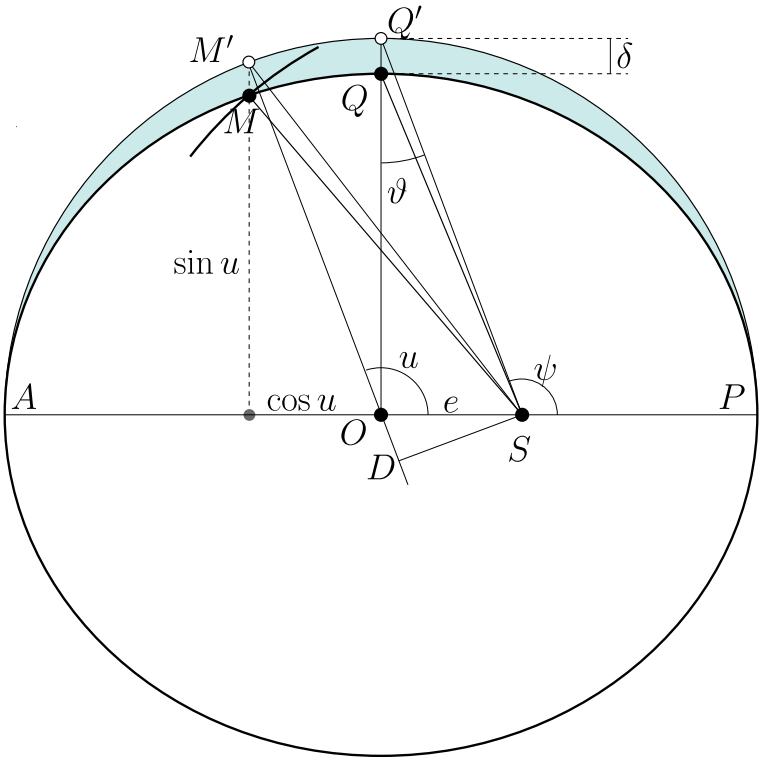


Fig. 11 A simplified drawing of Kepler's Fig. 10, illustrating the deduction of the equation of an ellipse. The labels of the relevant points are: S, Sun; O, centre of the eccentric orbit; A, perihelion; u, eccentric anomaly (here of M'); ψ, true anomaly (here of Q)

It may be useful to add a few comments, with the help of Fig. 11. The optical equation is the angle $\theta = \widehat{OQ'S}$ or $\widehat{OM'S}$; roughly, it is the difference between the true anomaly ψ and the eccentric anomaly u. It is clearly close to a maximum when $u = 90°$; according to Kepler, for Mars that angle is evaluated as $\theta \simeq 5°18'$. Still with $u = 90°$ Kepler evaluates the distance $\delta = |QQ'|$, which is the width of the lunula, as $\delta \simeq 0.00429$ when using the radius of the circular orbit as unit of length. That quantity is precisely $\sec\theta - 1$, i.e., the excess of the secant over the radius: this is what strikes Kepler. His final argument is that for a generic point M' on the circle with eccentric anomaly u he should recalculate the distance from the Sun the same way, i.e., taking the projection of the segment SQ' on the line OQ'. For $u = 90°$ this turns out to be precisely the radius. For a generic point M' on the circle we should project the segment SM' on the line OM', thus getting the segment DM', and take the length of the projection as the distance from the Sun; i.e., we find the true position M of the Planet by setting $|SM| = |DM'|$. That way he obtains the arc with radius SM in the figure, which fits the observations. The position M of the planet is at the intersection of the arc with the line from M' orthogonal to the line AP of apsides.

Now the road is traced. In the next Chaps. LVII–LIX Kepler exploits his intuition, and proves that the orbit so obtained is a perfect ellipse. With his Euclidean methods the construction is rather elaborated: it takes several pages. So let us make it simpler with our methods of analytical geometry, again with the help of Fig. 11. Taking O as the origin, the line of apsides AP as x axis and setting for a moment $|OA| = 1$ we find that $SM' = (\cos u - e, \sin u)$ and $OM' = (\cos u, \sin u)$. The projection $|DM'| = 1 - e\cos u$ is calculated by a scalar product. Let us now rescale all the distances by restoring $a = |OA|$, the radius of the eccentric circle. Following the trace suggested by Kepler we readily get

$$r = a(1 - e\cos u)\ ;$$

this is the equation of an ellipse with semimajor axis a and eccentricity e, written in terms of the distance r from the focus S and of the eccentric anomaly u.

The elliptic form of the orbit has been thus discovered, which establishes the first Kepler's law.[4]

3.5 *Kepler's Equation*

Having determined the shape of the orbit, in Chap. LX Kepler comes to the problem of calculating the position of the Planet on its orbit at any given time. Again, his construction using Euclidean methods is rather elaborated, and we are able to simplify it.

The problem is to find a *mean anomaly* ℓ, namely a quantity (or possibly an angle) that evolves uniformly in time. The method is based on the principle that we

[4] Remark that the orbit is planar, as determined in advance, see Sect. 3.1.

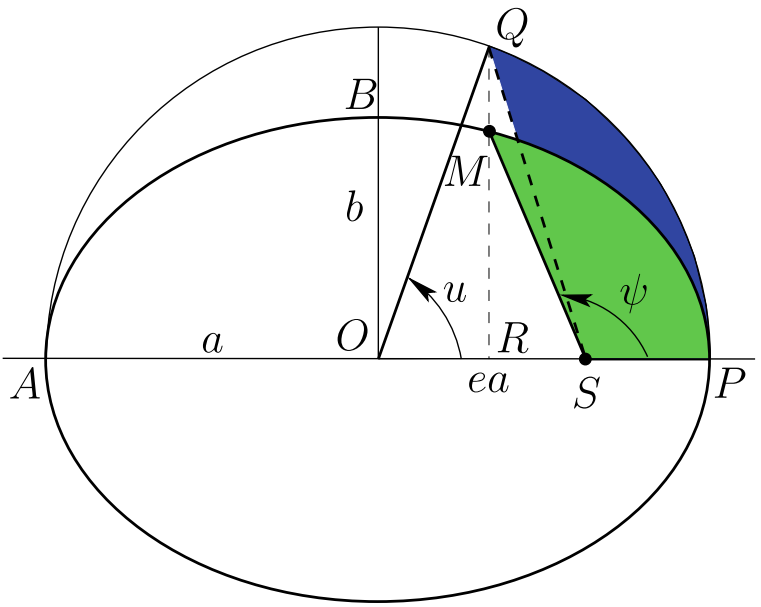


Fig. 12 Illustrating the deduction of Kepler's equation. The labels of the relevant points are: S, Sun; O, centre of the eccentric orbit; A, perihelion; a, semimajor axis; b semiminor axis; u, eccentric anomaly; ψ, true anomaly

have seen established in Sect. 3.2, using again the idea of evaluating the sum of the distances from the Sun through the area swept by the radius joining the Planet with the Sun, namely of the elliptic sector SPM. A direct calculation of that area is a hard task, but Kepler uses the ingenious method represented in Fig. 12. Assuming that the Planet is in M draw the line RQ orthogonal to the line of apsides, with Q on the circle of radius a. Then we easily calculate

$$\text{area}SPM = \frac{b}{a}\big(\text{area}(OPQ) - \text{area}(OSQ)\big) = \frac{b}{a}\left(\frac{1}{2}a^2u - \frac{1}{2}a^2e\sin u\right) .$$

The area evolves uniformly as $\frac{\pi ab}{T}t$; hence the mean anomaly ℓ at time t is

$$\ell = nt \quad \text{with} \quad n = \frac{2\pi}{T} ,$$

where n is the *mean motion*, T is the period, and the relation between ℓ and the eccentric anomaly u is given by the celebrated Kepler's equation

$$u - e\sin u = \ell .$$

The final problem is to calculate the eccentric anomaly u for any given value of ℓ, since ℓ is easily calculated. Kepler could not find a solution of his equation: it can not be solved by the methods of Euclidean geometry, indeed. Thus Kepler closes Chap. LX with a question to geometers:

Data area partis semicirculi, datoque puncto diametri, invenire arcum, & angulum ad illud punctum: cujus anguli cruribus, & quo arcu, data area comprehenditur. Vel: Aream semicirculi ex quocunque puncto diametri in data ratione secare.	Let the area of a part of a semicircle and a point on the diameter be given; to find the arc and the angle at that point, such that the sides of that angle and that arc enclose the given area. Or, similarly: to divide the area of a semicircle in a given ratio from any given point on the diameter.
Mihi sufficit credere, solvi a priori non posse propter arcus & sinus ἑτερογένειαν. *Erranti mihi, quicumque viam monstraverit, is erit mihi magnus Apollonius.*	I think that a priori it can not be solved, for the arc and the sine are *heterogeneous* quantities. I'm wandering here, and if anyone will show me the way, he will be for me a great Apollonius.

A solution by series (that now we call *Bessel functions*) has been published in 1771 by Lagrange [12, 13].

3.6 The Perfection of the World

The third Kepler's law seems to represent the conclusion of a continuous and perseverant reflection that had started with *Mysterium Cosmographicum* [7], and had continued for more than twenty years: Kepler was obsessed with the pursuit of the world's perfection.

The main thesis of *Mysterium Cosmographicum* is that the architecture of the world is deeply connected with the existence of five regular polyhedra. The scheme is represented in Fig. 13, a beautiful and famous image. Kepler distinguishes two different regions, separated by Earth. The external spheres and polyhedra are arranged as

Saturn — *Cube* — Jupiter — *Tetrahedron* — Mars

The separating sphere is that of the Earth and the Moon, enclosed inside two polyedra in the order

Dodecahedron — Earth and Moon — *Icosahedron*

The internal region is arranged as

Venus — *Octahedron* — Mercury

One can hardly expect that the dimensions of the polyhedra fit exactly with the spheres; but Kepler makes a strong effort in order to account also for the epicycles

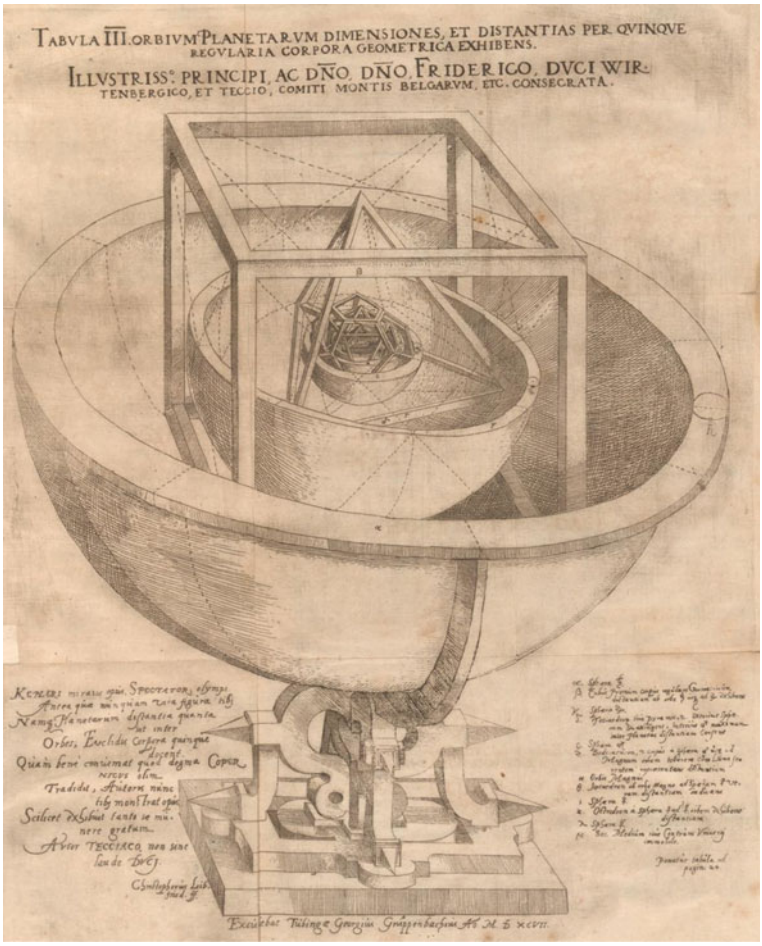


Fig. 13 A famous image from *Mysterium Cosmographicum* of Kepler, drawn by Georg Gruppenbach. The title says: "The sizes of the orbits of the Planets, exhibiting the distances ruled by five regular polyedra." The inscription on the left lower corner says: "Here you will admire, spectator, the work of the Olimpian Kepler, an image that you never saw before. Indeed the five solids of Euclid rule the distance between the orbs. The author's work now reveals how good was the doctrine transmitted to us by Copernicus"

of the Planets. Moreover, he adds plenty of reasons, based on geometrical and arithmetical considerations close to the border of mysticism (or even beyond), in order to establish that the order of polyhedra can not be different from the actual one; consequently, the same applies for the order of the Planets. Anyway, he comes to the conclusion that there are six Planets because there are five regular polyhedra, and that the order of Planets and polyhedra reflects a necessary and perfect plan.

The geometry of polyhedra is widely revisited in *Harmonice Mundi*. Moreover in Liber III Kepler comes to establish a connection between the motions of the Planets and the rules of musical harmonics. What he wants to discover is the music played by Planets in honor of the Creator. Figure 14 is such an example.

The perfect musical harmony of the Planets leads Kepler to reconsider the problem of finding a relation between the radii of the orbits and the periods of revolution. He had already raised the question in *Mysterium Cosmographicum*, but without being able to reach a conclusion. The question is achieved in Liber V, Chap. III of *Harmonice mundi*.

Hactenus egimus de diversis moris vel arcubus unius et eiusdem Planetæ. Jam etiam de binorum Planetarum motibus inter se comparatis agendum. (. . .)

So far we have considered either time intervals or arcs of one and the same Planet. Now we should consider the motions of pairs of Planets, compared with each other. (. . .)

Rursum igitur hic aliqua pars mei Mysterii Cosmographici, suspensa ante 22 annos, quia nondum liquebat, absolvenda et huc inferenda est. Inventis enim veris orbium intervallis per observationes Brahei, plurimi temporis labore continuo, tandem, tandem genuina proportio temporum periodicorum at proportionem orbium—

sera quidem respexit inertem,
Respexit tamen et
longo post tempore venit;

Therefore I may achieve another part of my *Mysterium Cosmographicum*, suspended twenty-two years ago, because it was not yet clear enough. For having calculated the distances of the orbits, thanks to Brahe's observations, and having spent a great amount of time working hard, at last!, at last the true proportion of the periodic times to the orbits —

late she saw me, lying helpless,
yet she came and gazed at me
after a long time;

eaquem si temporis articulos petis, 8 Mart. hujus anni millesimi sexcentesimi decimi octavi animo concepta, sed infeliciter ad calculos vocata, eoque pro falsa rejecta, denique 15 Maji reversa, novo capto impetu expugnavit mentis meae tenebras, tanta comprobatione et laboris mei septendecennalis in observationibus Braheanis et meditationis hujus in unum conspirantibus, ut somniare me et præsumere quæsitum inter principia primo crederem.

and if you want to know the precise moment, the first idea came across me on 8 March of this year 1618; but it was rejected as false, due to inappropriate reduction to calculation. But later it fell again upon me on 15 May, and conquered with outstanding power the darkness of my mind, thanks to the alliance between this idea and my seventeen years labour on Brahe's observations. I thought I was dreaming, and eventually I took my result for granted in my first assumptions.

I can not refrain from adding a further personal remark. I felt again impressed by the attitude of Kepler. With the verses in the middle, taken from Vergilius (*Bucolica*, Ecloga I), he fully expresses all his relief, almost a sense of liberation from an obsession that has dogged him for so many years. At the same time he adds an

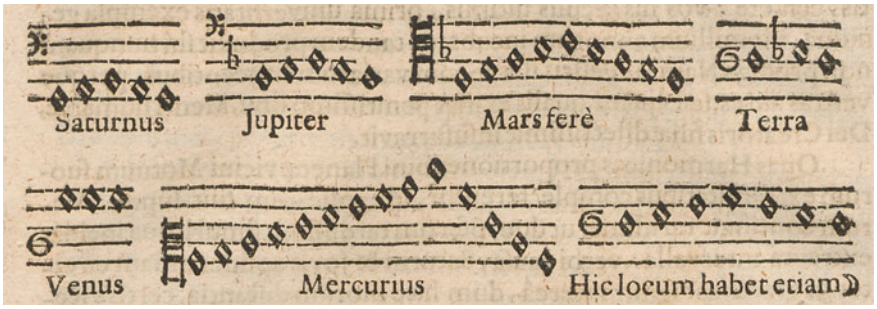


Fig. 14 The music played by Planets in honor of their Creator. Image from *Harmonice Mundi*

utterance that is almost a confession: it is not Kepler who catches the truth; it is the truth who catches Kepler. I never found anything similar in a book of Analysis.

But let us come to the conclusion of Kepler.

> *Sed res est certissima exactissimaque, quod* **proportio quæ est inter binorum quorumcunque Planetarum tempora periodica, sit præcise sesquialtera proportionis mediarum distantiarum, id est orbium ipsorum**; *attento tamen hoc, quod medium arithmeticum inter utramque diametrum ellipticæ orbitæ sit paulo minus longiore diametro. Itaque si quis ex periodo, verbi causa Telluris, quæ est annus unus, et ex periodo Saturni triginta annorum, sumserit tertiam proportionis partem, id est, radices cubicas, et huius proportionis duplum fecerit, radicibus quadrate multiplicatis, is habet in prodeuntibus numeris intervallorum Terræ et Saturni a Sole mediorum proportionem justissimam. Nam cubica radix de 1 est 1, ejus quadratum 1. Et cubica radix de 30 est major quam 3, eius igitur quadratum majus quam 9. Et Saturnus mediocriter distans a Sole, paulo altior est noneuplo mediocris distantiæ Telluris a Sole.*
>
> For it is a definitely certain and absolutely exact truth that **the actual proportion between the periodic times of any two Planets is precisely the sesquialtera proportion of the mean distances of the orbits, themselves;** this taking into account that the arithmetic mean between the two diameters of the elliptic orbit is a little less than the longer diameter. Thus if one considers, e.g., the period of the Earth, which is 1 year, and the thirty years period of Saturn, and takes one third of the proportions, that is the cubic roots, and doubles that proportion, making the square of the roots, he will get from the resulting numbers the correct proportion of the mean distances of the Earth and of Saturn from the Sun. For the cubic root of 1 is 1, and its square is 1; and the cubic root of 30 is greater than 3, and its square greater than 9. Indeed the average distance of Saturn from the Sun is a little bigger than nine times the average distance of the Earth from the Sun.

This is the third Kepler's law: the sole part of *Harmonice Mundi* that has survived in our Astronomy. Anyway now Kepler can finally claim that:

The world is perfect. The plan of God has been unveiled.

But . . . is it really so?

4 The "Tabulæ Rudolphinæ"

In 1564 Tycho Brahe (17 years old) conceived the project of compiling new astronomical tables that should replace the old and obsolete Alphonsine ones. The work was undertaken in 1572, and in 1600 Kepler initiated his collaboration for working out all the necessary calculations. After Tycho's death in 1601 Kepler took the job of completing the compilation, a work that continued for several years and was terminated in 1623. The new tables, called *Rudolphinæ*, were published in 1627 [4].

Kepler was very proud of his work. In the *præfatio* he writes [11]:

> *Et de certitudine quidem calculi testabuntur observationes præsentium temporum, imprimis Braheanæ; de futuris vero temporibus plura præsumere non possumus, quam vel observationes veterum, quibus usus sum, vel ipsa motuum mediorum conditio, nondum penitus explorata, concursusque causarum physicarum præstare possunt, cum observationes Regiomontani et Waltheri testentur, omnino de æquationibus secularibus esse cogitandum, ut singulari libello reddam demonstratum suo tempore; quæ tamen æquationes quales et quantæ sint, ante plurimum sæculorum decursum observationesque eorum, a gente humana definiri nequaquam possunt.*
>
> And the reliability of our calculations will be proved by the observations made in our epoch, in particular by Brahe. Yet, concerning the future we can not expect so much. The validity may be questioned by ancient observations, that I'm well aware of, by the knowledge of the mean motions, that have not yet been fully explored, and by the concurrence of physical actions. The observations of Regiomontanus and Walther do indeed show that we should definitely think about secular equations, as I will explain in a specially devoted booklet. Which and how many equations we need, humanity will be unable to decide it before many centuries of observations have been accumulated.

4.1 The Need for "Secular Equations"

The booklet promised by Kepler remained an unpublished manuscript for more than two centuries. Very little was known about it, because Kepler had only communicated in a few letters to friends that he had discovered an inequality in the motions of the Planets, in particular of Jupiter and Saturn: Jupiter seems to accelerate, Saturn to slow down. Kepler's manuscript was published in the complete collection of the works of Kepler in 1860 [10].

In his manuscript Kepler compares the results of his calculations with a series of observations collected between 1461 and 1504 by Johannes Müller der Könisberg (1436–1476), also named Regiomontanus, and Bernard Walther (1430–1504). A few ancient observations reported by Ptolemæus are also included. A report on Kepler's manuscript may be found in [6]; here is a short summary of the main points.

A comparison between the observed longitudes of Jupiter and Saturn and the predicted ones through Rudolphinæ tables is reported in Fig. 15. The dispersion of

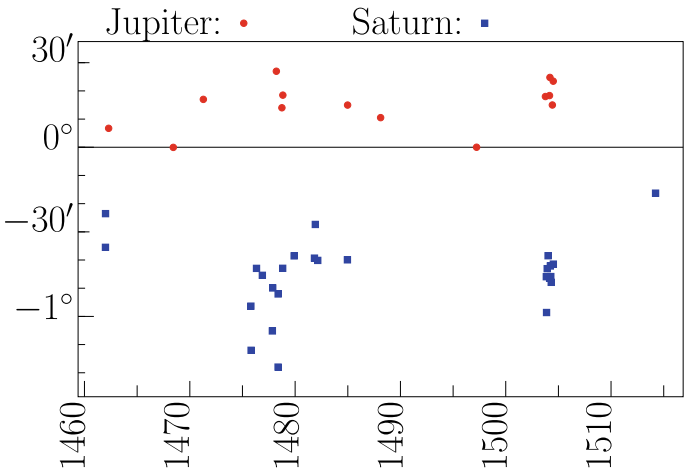


Fig. 15 Difference between the observed and calculated longitude of Jupiter and Saturn, according to Kepler [10]

data is quite big, but we should keep in mind that they come from visual observations or at most using armillæ, that could not assure a very high precision. Nevertheless, it is evident that there is a systematic deviation: Jupiter is always in advance with respect to the predictions; Saturn is always delayed. This is the deviation mentioned by Kepler in the passage quoted above. Kepler's note includes data also for Mars, Venus and Mercury, but the corresponding plot does not suggest a systematic deviation, though Kepler inclines to admit it (see [6]).

It took quite a long for Kepler to get convinced that there was some unaccounted phenomenon that could not be explained on the sole basis of the elliptic shape of the orbits. He refers that he eventually got convinced of the necessity of accepting the existence of inequalities when he went to examine the event of a great conjunction observed by Walther on 24 May 1504: the calculation had produced a discrepancy of 58′. Thus Kepler writes:

> *Et hic dissensus calculi in Jupiter and Saturn, excurrens ad integrum gradum, est remora illa, quæ me, plurima perplexitate circumventum, per solidos quinque menses in observationibus Waltherianis exercuit tandemque ad nova consilia circa motuum mediorum speculationem adegit, deprehensa manifesta inæequalitate motuum seculari. (Absolvi hucusque 18 Junii 1624.)*
>
> And this discrepancy in the calculation for Jupiter and Saturn, which amounts to a whole degree, is such a hindrance that it caused me to be assailed by many perplexities, and for five solid months I have been troubled until I eventually changed my advice concerning average motions, having accepted the manifest secular inequality of the motions. (I concluded this on 18 June 1624.)

The expression "secular inequality" coined here by Kepler means that he is forced to admit that the orbital elements may be subject to a slow change that becomes visible after centuries. Thus he begins to look for a *secular equation* to be included in his calculation in order to cover that phenomenon. He tries to adjust the position of the perihelion, assuming that the orbit of the Earth remains uniform and refusing to modify the eccentricities, accurately determined by Tycho Brahe. He reaches a partial success by fitting the observations of Ptolemæus, but he fails to fit them also with the data of Regiomontanus and Walther. A more detailed explanation of Kepler's attempts would take too much space here; the interested reader may want to look at [6].

As to the nature of the secular equations to be included Kepler is not very explicit, but reading his note I got convinced that he was trying to introduce *periodic* corrections. That is, to reintroduce the epicycles that he had successfully removed.

The final conclusion of Kepler that we read in the preface to *Tabulæ Rudolphinæ* is that he could only leave the problem for the future of astronomy, perhaps after many centuries of systematic observations. Such a claim is fully justified in the framework of ancient Astronomy, including Copernicus and Kepler: the sole way for determining the periodic motions was to observe them over many periods. A dynamical model was not available.

4.2 After Kepler

The problem pointed out by Kepler received increasing attention during the XVII and XVIII centuries. A first step was made by Halley in 1700 with the publication of tables that included secular corrections *linear in time*. This was a pragmatic approach, actually an interpolation, that in Halley's intentions could produce correct predictions over 6000 years before and after 1700. After Halley the expression "secular terms" began to identify non periodic terms in perturbation expansions.

As we all know, Newton's gravitational theory suggested the possibility that the secular inequalities observed by Kepler could be explained as the effect of gravitational interactions among the Planets. This has been indeed the job undertaken by great mathematicians and astronomers of the XVIII century, among them D'Alembert, Clairaut, Lalande, Euler, Boskovich, Lagrange and Laplace. A report about the progress of the problem raised by Kepler paying particular attention to the work of Lagrange and Laplace may be found in [14]. The remarkable outcome of that effort has been the birth of perturbation theory as was later developed, and still is in use. The XIX century Astronomy (or Celestial Mechanics, as we may prefer to call it) has seen a strong development of perturbation methods (and an accumulation of troubles due to resonances), culminating with the discovery of chaos in the problem of three bodies, by Poincaré.

In the second half of the past century the availability of computers and the beginning of the exploration of space, together with the announcement of strong theoretical results by Kolmogorov, Arnold, Moser and Nekhoroshev, launched the development

of new methods and results that is still in progress. A detailed exposition of the matters listed in this section exceeds both the limits and the scope of the present note. Thus we should stop here.

4.3 A Final, Strictly Personal Remark

The discovery of chaos may appear as the downfall of Kepler's dream that the perfection of the Universe could be discovered. In this sense his project should be considered as unaccomplished. But there are many ways of looking at that question.

We know that on May 29, 1919, Frank Watson Dyson and Arthur Stanley Eddington observed the deflection of light by the Sun during an eclipse as predicted by Einstein some years before. It is reported that when Albert Einstein was informed of that result, he replied that had the experiment failed: "Then I would feel sorry for the dear Lord. The theory is correct anyway". We may want to rephrase the sentence by saying:

"The theory is perfect. The world is wrong."

But a quite different attitude may be possibly taken. We could say as well:

"The world is perfect. What is wrong is our idea of perfection."

> *Oscar Wilde said that sunsets were not valued because we could not pay for sunsets.*
> *But Oscar Wilde was wrong; we can pay for sunsets.*
> *We can pay for them by not being Oscar Wilde.*
>
> (G. K. Chesterton.)

References

1. Barbour, J.B.: The Discovery of Dynamics. Oxford University Press (2001)
2. Bessel, F.W.: On the parallax of 61 Cygni. Mon. Not. R. Astron. Soc. **4**(17), 152–161 (1838)
3. Brahe, T.: De mundi aetherei recentioribus phaenomenis liber secundus de cometa anni 1577. Pragae Bohemiae : typis inchoatus Uraniburgi Daniae absolutus (1603). Digital copy available from ETH-Bibliothek Zürich, Rar 4458. Public Domain Mark. http://doi.org/10.3931/e-rara-1303
4. Rudolphinae, T., Jonas Saur, U.: Digital copy available from ETH-Bibliothek Zürich, Rar 8895 q (1627). Public Domain Mark. http://doi.org/10.3931/e-rara-8742
5. Dialogo di Galileo Galilei Linceo, Matematico Sopraordinario dello Studio di Pisa, e Filosofo, e Matematico primario del serenissimo Gr.Duca di Toscana, dove ne i congressi di quattro giornate si discorre sopra i due Massimi sistemi del mondo, Tolemaico, e Copernicano In Fiorenza, per Gio:Batista Landini MDCXXXII. Reprinted in: Le Opere di Galileo Galilei, Edizione Nazionale sotto gli auspicii di Sua maestà il Re d'Italia, Vol. VII. Firenze, Tipografia di G. Barbera (1897)

6. Giorgilli, A.: A Kepler's note on secular inequalities. Rendiconti dell'Istituto Lombardo Accademia di Scienze e Lettere, Classe di Scienze Matematiche e Naturali **145**, 97–119 (2011)
7. Kepler, J.: Mysterium Cosmographicum de admirabili proportione orbium cœlestium; deque causis cœlorum numeri, magnitudinis, motuumque periodicorum genuinis et propriis, demostratum per quinque regularia corpora geometrica, a M. Johanne Keplero, Wirtembergico, Illustrium Styriæ Provincialium Mathematico. Tubingæ excudebat Georgius Gruppenbachius, Anno MDXCVI. Reprinted in: Prodromus dissertationum Cosmographicarum, continens Mysterium Cosmographicum, 1621. See also: Johannis Kepleri astronomi opera omnia, edidit Dr. CH. Frisch, Frankfurti A.M. et Erlangæ Heyder & Zimmer, MDCCCLX, Vol. I. Also reprinted in: Johannes Kepler Gesammelte Werke, Bayerische Akademie der Wissenschaften München, Band I (1990)
8. Kepler, J.: Astronomia Nova, seu Physica Cœlestis tradita commentariis de motibus Stellæ Martis ex observationibus G.V. Tychonis Brahe. Jussu & sumptibus Rudolphi II, Romanorum Imperatoris &c. Plurium annorum pertinaci studio elaborata Pragæ, A.S.C.M.S. Mathematico Johanne Keplero, cum ejusdem C.M. privilegio speciali, Anno æræ Dionysianæ MDCIX. Reprinted in: Johannis Kepleri astronomi opera omnia, edidit Dr. CH. Frisch, Frankfurti A.M. et Erlangæ Heyder & Zimmer, MDCCCLX, Vol. III. Also reprinted in: Johannes Kepler Gesammelte Werke, Bayerische Akademie der Wissenschaften München, Band III (1990)
9. Kepler, J.: Harmonices Mundi libri V. sumptibus Godofredi Tampachii Bibl. Francof. excudebat Ioannes Plancus, anno MDCXIX. Reprinted in: Johannis Kepleri astronomi opera omnia, edidit Dr. CH. Frisch, Frankfurti A.M. et Erlangæ Heyder & Zimmer, MDCCCLX, Vol. V. Also reprinted in: Johannes Kepler Gesammelte Werke, Bayerische Akademie der Wissenschaften München, Band VI (1990)
10. Kepler, J.: Consideratio observationum Regiomontani et Waltheri. In Johannis Kepleri astronomi opera omnia, edidit Dr. CH. Frisch, Frankfurti A.M. et Erlangæ Heyder & Zimmer, MDCCCLX, Vol. VI, pp. 725–774. Also reprinted in: Johannes Kepler Gesammelte Werke, Bayerische Akademie der Wissenschaften München, Band XX, pp. 393–455 (1990)
11. Kepler, J.: In tabulas Rudolphi præfatio. In Johannis Kepleri astronomi opera omnia, edidit Dr. CH. Frisch, Frankfurti A.M. et Erlangæ Heyder & Zimmer, MDCCCLX, Vol. VI, pp. 666–674. See also: Johannes Kepler Gesammelte Werke, Bayerische Akademie der Wissenschaften München, Band I (1990)
12. Lagrange, J.L.: Nouvelle méthode pour résoudre les équations littérales par le moyen des séries. Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin, t.XXIV (1770). Reprinted in: Oeuvres de Lagrange, Gauthier–Villars, Paris (1869), tome III, pp. 5–73
13. Lagrange, J.L.: Sur le Problème de Képler. Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin, t.XXV (1771). Reprinted in: Oeuvres de Lagrange, Gauthier–Villars, Paris (1869), tome III, pp. 113–138
14. Laskar, J.: Lagrange et la Stabilité du Sytème Solaire. In: Sacchi Landriani, G., Giorgilli, A. (eds.) Sfogliando la Méchanique Analitique, LED edizioni, Milano (2008)
15. I. Newton Philosophiæ Naturalis Principia Mathematica, autore Is. Newton, Trin. Coll. Cantab. Soc. Matheseos Professore Lucasiano, & Societatis Regalis Sodali, Londini. Jussu Societatis Regiæ ac typis Josephi Streater, Anno MDCLXXXVII
16. Stephenson, B.: Kepler's Physical Astronomy. Springer (1987)
17. Swerdlow, N.M.: Astronomy in the renaissance. In: Walker, C. (ed.) Astronomy Before the Telescope, pp. 187–230. British Museum Press (1996)