

# WHY ARE INNER PLANETS NOT INCLINED?

by ANDREW CLARKE, JACQUES FEJOZ, and MARCEL GUARDIA

## ABSTRACT

Poincaré’s work more than one century ago, or Laskar’s numerical simulations from the 1990’s on, have irrevocably impaired the long-held belief that the Solar System should be stable. But mathematical mechanisms explaining this instability have remained mysterious. In 1968, Arnold conjectured the existence of “Arnold diffusion” in celestial mechanics. We prove Arnold’s conjecture in the planetary spatial 4-body problem as well as in the corresponding hierarchical problem (where the bodies are increasingly separated), and show that this diffusion leads, on a long time interval, to some large-scale instability. Along the diffusive orbits, the mutual inclination of the two inner planets is close to  $\pi/2$ , which hints at why even marginal stability in planetary systems may exist only when inner planets are not inclined. More precisely, consider the normalised angular momentum of the second planet, obtained by rescaling the angular momentum by the square root of its semimajor axis and by an adequate mass factor (its direction and norm give the plane of revolution and the eccentricity of the second planet). It is a vector of the unit 3-ball. We show that any finite sequence in this ball may be realised, up to an arbitrary precision, as a sequence of values of the normalised angular momentum in the 4-body problem. For example, the second planet may flip from prograde nearly horizontal revolutions to retrograde ones. As a consequence of the proof, the non-recurrent set of any finite-order secular normal form accumulates on circular motions – a weak form of a celebrated conjecture of Herman.

## CONTENTS

1. Introduction . . . . .	
2. Main results in Deprit coordinates . . . . .	
3. The averaging procedure and the secular Hamiltonian . . . . .	
4. Analysis of the first-order Hamiltonian . . . . .	
5. The inner dynamics . . . . .	
6. The outer dynamics . . . . .	
7. Reduction to a Poincaré map and the shadowing argument . . . . .	
8. Continuation in the planetary regime . . . . .	
9. From Deprit coordinates to elliptic elements: proof of Theorem 4 . . . . .	
Appendix A: Deprit’s coordinates . . . . .	
Appendix B: The scattering map of a normally hyperbolic invariant manifold . . . . .	
Appendix C: A general shadowing argument . . . . .	
Appendix D: Computation of the phase shift in $\tilde{\psi}_1$ . . . . .	
Appendix E: Expansion of the inner Hamiltonian . . . . .	
Appendix F: Corrigendum of [41] . . . . .	
References . . . . .	

## 1. Introduction

**1.1. *A case for instability in the solar system.*** — Hook’s and Newton’s discovery of universal attraction in the XVII century masterly reconciles two seemingly contradictory physical principles: the principle of inertia, put forward by Galileo and Descartes in terrestrial mechanics, and the laws of Kepler, governing the elliptical motion of planets around the Sun [1, 61, 86]. The unforeseen mathematical consequence of Hook’s and Newton’s discovery was to question the belief that the solar system be stable: it was no longer obvious that planets kept moving immutably, without collisions or ejections, because of their mutual (“universal”) attraction. Newton himself, in an additional and staggering tour de force, estimated the first order effect on Mars of the attraction of other planets.



But infinitesimal calculus was in its infancy and the necessary mathematical apparatus to understand the long-term influence of mutual attractions did not exist.

In their study of Jupiter's and Saturn's motions, Lagrange, Laplace, Poisson and others laid the foundations of the Hamiltonian theory of the variation of constants. They managed to compute the secular dynamics, i.e. the slow deformations of Keplerian ellipses, at the first order with respect to the masses, eccentricities and inclinations of the planets. This level of approximation is still integrable, and Keplerian ellipses have slow but non-vanishing precession and rotation frequencies, thus departing from the dynamical degeneracy described by Bertrand's theorem. Besides, the analysis of the spectrum of the linearised vector field entailed a resounding stability theorem for the solar system: the observed variations of action variables in the motion of Jupiter and Saturn come from resonant terms of large amplitude and long period, but with zero average ([66, p. 164], [68]). Yet it is a mistake, which Laplace made, to infer the topological stability of the non-truncated planetary system.

In the XVIII and XIX centuries, mathematicians spent an inordinate amount of energy trying to prove the stability of the Solar system... until Poincaré discovered a remarkable set of arguments strongly speaking *against* stability: generic divergence of perturbation series, non-integrability of the 3-body problem, and entanglement of the stable and unstable manifolds of the Lagrange relative equilibrium in the restricted 3-body problem [89].

In the mid XX century, Siegel and Kolmogorov still proved that, respectively for the linearisation problem of a one-dimensional complex map and for the perturbation of an invariant torus of fixed frequency in a Hamiltonian system, perturbation series do converge, albeit non-uniformly, under some arithmetic assumption of Diophantine type, ensuring that the frequencies of the motion are far from low order resonances, in a quantitative way. The obtained solutions are quasiperiodic and densely fill Lagrangian invariant tori. They form a large set in the measure theoretic sense, but a small set from the topological viewpoint. Besides, starting from dimension 6, invariant tori do not separate energy levels and thus do not confine neighboring motions, so, outside invariant tori, nothing prevents action variables from drifting. Kolmogorov's theorem was successfully adapted to the planetary system, despite the numerous degeneracies of the latter, and assuming that the masses of the planets are very small [3, 23, 38, 90]. This result is sometimes referred to as Arnold-Herman's invariant tori Theorem. The obtained solutions are small perturbations of (Diophantine) Laplace-Lagrange motions.

Soon afterward, Arnold imagined an example of dynamical instability in a near-integrable Hamiltonian system with many degrees of freedom, where action variables may drift (for some well-chosen orbits), by an amount uniform with respect to the smallness of the perturbation [4]. Of course, the drifting time tends to infinity as the size of the perturbation tends to 0, consistently with the continuity of the time- $t$  map of the flow with respect to parameters. Drifting orbits shadow the stable and unstable manifolds of a chain of hyperbolic invariant tori ("transition chain"). This phenomenon has been called *Arnold*

*diffusion*, since Chirikov coined the phrase, referring to the (in part conjectural) stochastic properties of such a dynamics [24]. In fact, in this seminal paper, Arnold conjectured the following.

*Conjecture 1 (Arnold [4]). — The mechanism of transition chains [...] is also applicable to the case of general Hamiltonian systems (for example, to the problem of three bodies).*

Arnold’s example has proved difficult to generalise because of the so-called *large gap problem*: usually the transition chain is a (totally disconnected) Cantor set of hyperbolic tori and it is not obvious whether there exist orbits shadowing these tori. A better strategy has emerged, consisting in shadowing normally hyperbolic cylinders (whether they contain invariant tori or not). Nearly integrable Hamiltonian systems are usually<sup>1</sup> classified as *a priori unstable* and *a priori stable* [21]. A priori unstable models are those whose integrable approximation presents some hyperbolicity (the paradigmatic example being a pendulum weakly coupled with several rotators). In this case, the unperturbed model has a normally hyperbolic invariant manifold with attached invariant manifolds that one can use, for the perturbed model, as a “highway” for diffusing orbits. The existence of Arnold diffusion generically in these models is nowadays rather well understood, at least for two and a half degrees of freedom (see [7, 19, 28, 31, 52, 78, 92], or [34, 53, 93] for results in higher dimension).

A priori stable systems are those whose integrable approximation is foliated by quasiperiodic invariant Lagrangian tori. Since the unperturbed Hamiltonian does not possess hyperbolic invariant objects, in order to construct the diffusing “highway” one has to rely on a first perturbation and face involved singular perturbation problems. One of the difficulties is that one cannot avoid double resonances, where the system is intrinsically non-integrable. Arnold’s conjecture refers to these models. The work of Mather on minimizing measures has been deeply influential. In the finite smoothness category, the papers [8, 18, 64] show the typicality (in the cusped residual sense as defined by J. Mather) of Arnold diffusion in a priori stable Hamiltonian systems of 3 degrees of freedom. Yet, many questions remain unsolved. In particular, the original Arnold conjecture on the typicality of Arnold diffusion for analytic non degenerate nearly integrable Hamiltonian systems of 3 or more degrees of freedom remains open (see however [47, 48]).

In the 1990s, with extensive numerical computations Laskar showed that over the physical life span of the Sun, or even over a few hundred million years, collisions and ejections of inner planets occur with some probability [67, 70].<sup>2</sup> Our solar system is now

---

<sup>1</sup> One can also consider the so called *a priori chaotic case*, where the unperturbed Hamiltonian presents “local non-integrability”. In particular, it has a first integral and a periodic orbit with transverse homoclinics at each energy level. Examples of such settings are certain geodesic flows with a time dependent potential, see [10, 25, 29, 30, 50, 51].

<sup>2</sup> Such long term computations are checked to pass various consistency tests (e.g. the preservation of first integrals). But due to the exponential divergence of solutions, they are statistical in nature: an uncertainty of a few centimeters on the initial position of the Earth leads to an uncertainty of the size of the Solar System after a few hundred millions years. But one likes to believe that such Hamiltonian systems have good shadowing properties, i.e. that any finite-time pseudo-orbit (as computed numerically) is shadowed by orbits.

believed only marginally stable. This has been corroborated by abundant numerical evidence, as overviewed in Morbidelli's book [83]. In particular, the effect of mean motion (Keplerian) resonances in the asteroid belt has been described by [73]. Numerical evidence has also been suggesting that secular resonances are a major source of chaos in the Solar system [49, 69, 71]. For example, astronomers have established that Mercury's eccentricity is chaotic and can increase so much that collisions with Venus or the Sun become possible, as a result from an intricate network of secular resonances [12]. On the other hand, that Uranus's obliquity ( $97^\circ$ ) is essentially stable, is explained, to a large extent, by the absence of any low-order secular resonance [11, 71]. The effects of secular resonances of the inner planets have later been studied systematically using both computer algebra and numerics and the main "sources of chaos" in the inner solar system have been identified [6, 80].

The mathematical theory of instability remains in its infancy and the Astronomers' instability mechanisms are still mysterious. A matter of discontent with Arnold diffusion is that the time needed for actions to drift looks larger than in other, far from integrable, instability mechanisms that astronomers observe. Resonance overlapping, a phenomenon described by Chirikov [24], would be a fantastic competing mechanism. But to our knowledge it lacks mathematical explanation (see however a simple example in [42]).

Regarding "the oldest problem in dynamical systems", in his ICM lecture [59] Herman formulated the following two precise conjectures. Consider the  $N$ -body problem in space, with  $N \geq 3$ . Assume that the center of mass is fixed at the origin and that on the energy surface of level  $e$  we  $C^\infty$ -reparametrise the flow by a  $C^\infty$  function  $\varphi_e$  such that the collisions now occur only in infinite time ( $\varphi_e > 0$  is a  $C^\omega$  function outside collisions).<sup>3</sup>

*Conjecture 2 (Global instability).* — *Is for every  $e$  the non-wandering set of the Hamiltonian flow of  $H_e$  on  $H_e^{-1}(0)$  nowhere dense in  $H_e^{-1}(0)$ ?*<sup>2</sup>

This would imply that bounded orbits are nowhere dense and no topological stability occurs.<sup>4</sup> The conjecture is wide open, and we are still at the stage of looking for non-recurrent orbits having negative energy (due to the Lagrange-Jacobi identity, non-wandering orbits have negative energy) [40].

Herman further argues that *What seems not an unreasonable question to ask (and possibly prove in a finite time with a lot of technical details) is that:*

*Conjecture 3 (Planetary instability).* — *If one of the masses is fixed ( $m_0 = 1$ ) and the other masses  $m_j = \rho \tilde{m}_j$ ,  $1 \leq j \leq n-1$ ,  $\tilde{m}_j > 0$ ,  $\rho > 0$ , then in any neighbourhood of fixed different circular orbits around  $m_0$  moving in the same direction in a plane, when  $\rho$  is small, there are wandering domains.*

---

<sup>3</sup> Here indeed it is natural to count the non-wandering set without collision orbits, since positions do not necessarily go to infinity at collisions. Herman furthermore claims that this reparametrised flow is complete. This is not clear due to the potential presence of non-collision singularities. But conjecture 2 remains relevant with a possibly incomplete flow.

<sup>4</sup> This incidentally contradicts a conjecture of Poincaré, that periodic orbits are dense [89, End of Section 36].

This conjecture is also wide open. Observed unstable motions in Celestial Mechanics in the last decades has typically fallen into one of two categories: chaotic dynamics and the previously mentioned Arnold diffusion which, even if related, are different in nature, and each can occur independently of the other.

By chaotic motions we mean the existence of invariant sets with symbolic dynamics (i.e. dynamics (semi-)conjugated to the shift map). Such sets can exist even if the orbital elements are almost constant, and therefore even if Arnold diffusion is not present. Such dynamics imply positive topological entropy and can be used to establish the existence of oscillatory motions<sup>5</sup> as conjectured by Chazy [17], see [2, 35, 54–57, 75, 76, 84, 91]. Other regions of phase space where chaotic motions have been proven to exist are close to triple collisions [77, 79], close to Poincaré’s periodic orbits of second species [9], close to the Euler-Lagrange points [13], on the submanifold of zero angular momentum of the 3-body problem [81, 82], and so on. It is part of the richness of the N-body problem to include such diverse kinds of behavior. Note that none of these results apply to models of solar systems, where all the bodies revolve (approximately) on well-separated ellipses. In fact, there are no results on Smale horseshoes in such regimes even in reduced models such as the Restricted Planar Circular 3-Body Problem.<sup>6</sup>

Indeed, scarce mathematical mechanisms regarding instabilities of any kind have been described regarding more astronomical regimes, which would be plausible for sub-systems of solar or extra-solar systems. Within the domain of negative energy, of special astronomical relevance is the planetary problem, where planets with small masses revolve around the Sun.<sup>7</sup> Another well-known problem is the hierarchical problem, an extension to N-bodies of the so-called Hill problem or lunar problem, where one body (the Sun) revolves far away around the other two (the Earth and the Moon).<sup>8</sup> These problems are rendered difficult by the proximity to a degenerate (“super-integrable”) integrable system of two uncoupled Kepler problems. As mentioned above, the recurrent set has positive Lebesgue measure due to Arnold-Herman’s invariant tori theorem, while the existence of unstable orbits relies on Arnold diffusion.

Even if Arnold in his seminal paper conjectured that his Arnold diffusion mechanism via transition chains should be present in the 3 body problem, even nowadays the results in this direction are rather scarce and, up to the present paper, nonexistent in planetary regimes. Indeed, as far as the authors know, the *only complete analytical* proof

---

<sup>5</sup> Oscillatory motions are those such that the superior limit of the distances between bodies is unbounded while the inferior limit is bounded. This terminology typically refers to the motion of a comet in a Sun-Planet-Comet model, where the comet makes long excursions far away from the other two bodies. Such regime is radically different to planetary regimes where all bodies make approximate ellipses. The most complete result up to now is [57], where such motions are proven to exist in the 3-Body Problem for almost any value of the masses. Note that the orbital elements along such orbits remain almost constant for all time, and so there is no Arnold diffusion.

<sup>6</sup> The paper [41] proves the existence of Smale horseshoes for the secular 3-body problem. This certainly does not lead to chaotic dynamics for the full 3-body problem since one model only approximates well the other over certain long time scales.

<sup>7</sup> Jupiter, the largest planet of the solar system, weighs roughly 1/1000 the mass of the Sun.

<sup>8</sup> The lunar distance is roughly 1/400 the distance from Earth to Sun.

of Arnold diffusion in Celestial Mechanics, prior to the present work, is contained in an article by Delshams, Kaloshin, de la Rosa and Seara [35]. In this paper, the authors consider the Restricted Planar Elliptic 3 Body Problem and construct orbits with large drift in angular momentum, assuming the mass ratio and the eccentricities of the primaries are sufficiently small. A key point of this paper is that the body of zero mass is close to the so called parabolic motions. That is, it relies on the invariant manifolds of infinity which are already present in the two body problem. They then perform a delicate perturbative analysis of the model.

Some works have uncovered instability mechanisms in celestial mechanics related to Arnold diffusion, relying on computer assisted computations (see [16, 43] and also [14] which relies on computer assisted proofs) or conditionally to a plausible transversality hypothesis [95].

Regarding Conjecture 3, the construction of wandering domains using Arnold diffusion-like mechanisms is a difficult problem. As far as we know, the only positive result in this direction is in [72], where wandering domains for Gevrey nearly integrable symplectic maps are constructed. The methods used in that paper to construct such domains do not admit an immediate extension to the analytic category.

The present work

- proves Arnold’s Conjecture 1 in the planetary spatial 4-body problem (see Theorem 4 below)
- and proves a weak local version of Herman’s Conjecture 3, dealing with non-recurrent orbits instead of wandering orbits (Theorem 6), for the corresponding secular dynamics.

**1.2. Main results.** — Consider the 4-body problem, that is 4 bodies numbered from 0 to 3 moving in 3-dimensional space according to the Newtonian gravitational law,

$$\ddot{x}_j = \sum_{\substack{0 \leq i \leq 3 \\ i \neq j}} m_i \frac{x_i - x_j}{\|x_i - x_j\|^3},$$

where  $x_j \in \mathbf{R}^3$  is the position and  $m_j > 0$  is the mass of body  $j$  for  $j = 0, 1, 2, 3$ .

For the sake of simplicity, let us first focus on the “hierarchical regime” where body 2 revolves around and far away from bodies 0 and 1, while body 3 revolves around and even farther away from bodies 0, 1 and 2. Each body thus primarily undergoes the attraction of one other body: bodies 0 and 1 are close to being isolated, body 2 primarily undergoes the attraction of a fictitious body located at the center of mass of 0 and 1, and body 3 primarily undergoes the attraction of a fictitious body located at the center of mass of 0, 1 and 2. We think of body 0 as the Sun and of the three other bodies as planets. The Jacobi coordinates are well suited for this regime (Figure 2), but we defer their definition to a later stage. Assuming that the center of mass is fixed, the small displacements of the Sun may be recovered from the positions of the planets.

The fast dynamics consists in those planets moving along Keplerian ellipses according to the above approximation, with elliptical elements as first integrals, in addition to the total energy and the total angular momentum  $\mathbf{C} = \mathbf{C}_1 + \mathbf{C}_2 + \mathbf{C}_3$ , where  $\mathbf{C}_i$  is the angular momentum of planet  $i$ . Let  $a_1$ ,  $a_2$  and  $a_3$  be the semimajor axes. In our regime,

$$a_1 \ll a_2 \ll a_3$$

(we will later make a more specific hypothesis on how the ratios  $a_i/a_{i+1}$  of semimajor axes compare to each other). Let also  $e_1$ ,  $e_2$  and  $e_3$  be the eccentricities. The angular momentum  $\mathbf{C}_i$ , seen as a vector in space, is normal to the plane of ellipse  $i$  and its length is  $\sqrt{a_i(1 - e_i^2)}$ , up to a mass factor (see (1) below). Let  $\theta_{ij}$  be the mutual inclinations, i.e. the oriented angle  $\alpha_{\mathbf{C}_i \times \mathbf{C}_j}(\mathbf{C}_i, \mathbf{C}_j)$  between the angular momenta of planets  $i$  and  $j$ , the orientation being defined by the normal vector  $\mathbf{C}_i \times \mathbf{C}_j$  (here assumed non-zero).

The so-called secular dynamics describes the slow evolution of the three Keplerian ellipses. At the first order of approximation, it is governed by the vector field obtained by averaging out the mean anomalies, thus defining a dynamical system on the “secular space” of triples of Keplerian ellipses with fixed semimajor axes. In the hierarchical regime, the dominating term is what is usually called the “quadrupolar” Hamiltonian  $F_{\text{quad}}^{12}$  of the two inner planets. It was introduced in various particular cases by Lidov and then Kozai [65, 74, 97] (it had been previously analysed by H. von Zeipel [94], see also the survey [62]). It may come as a surprise that  $F_{\text{quad}}^{12}$  is integrable (defined in (26) in Section 3), as noticed by Harrington [58], due to the fact that it does not depend on the argument of the outer pericenter  $g_2$ . This dynamics was later studied more globally in the secular space by Lidov, Kozai and others (see a review in [85]).

Our analysis follows from a higher order, non-integrable approximation of the system (it will rely also on the quadrupolar Hamiltonian of planets 2 and 3, and the octupolar term of planets 1 and 2, as introduced later). Precisely understanding the various time scales within the secular dynamics will be key to our analysis. At this stage, let us only loosely describe the different roles played by the three planets:

- Since the semimajor axis  $a_3$  is so large, it is planet 3 which most contributes to the total angular momentum  $\mathbf{C}$ , so planet 3 cannot change substantially in eccentricity or inclination. Yet, planet 3 is a source of angular momentum for the two inner planets, and a minor change of  $\mathbf{C}_3$  results in major changes of the elliptical elements of planets 1 or 2.
- The elliptical elements of the first (inner) planet vary faster than the elements of the second one. So the approximate conservation of  $F_{\text{quad}}^{12}$  introduces a coupling between the eccentricity and the inclination of the first planet. It can drive an initially near-circular orbit to arbitrarily high eccentricity, and flip an initially moderately inclined orbit between a prograde and a retrograde motion. This is an integrable, quasiperiodic dynamics, reportedly discovered by Lidov and thus



called the *Kozai mechanism*.<sup>9</sup> After reduction by the symmetry of rotations, the existence of an elliptic secular fixed point, which is the continuation (for non-zero eccentricities and inclinations) of the singularity known to Lagrange and Laplace, explains the oscillation of the argument of the inner pericenter.<sup>10</sup>

For our part, we will instead localise in a region where the two inner planets are mutually highly inclined, i.e.  $\theta_{12}$  is close to  $\pi/2$ . In this region, there is a hyperbolic secular singularity, which was used for instance by Jefferys-Moser to prove the existence of normally hyperbolic invariant tori [63]. This singularity, when lifted to the full phase space, yields a normally hyperbolic cylinder which will be crucial to our construction. In particular we will show that its codimension-1 stable and unstable invariant manifolds split, and we will need to control the splitting of the underlying foliations more or less carefully depending on directions. The motions of interest to us will occasionally shadow the stable and unstable foliations of the cylinder, thus moving away from the cylinder itself, to a homoclinic channel and then back close to the cylinder (along a so-called homoclinic excursion). So  $\theta_{12}$  will vary substantially but with no drift. On the other hand, it is planet 1 which most contributes to the secular Hamiltonian, so, because of the near conservation of the latter, the eccentricity  $e_1$  will be bounded within a small interval.

- In contrast, no first integral prevents the eccentricity  $e_2$  or the mutual inclination  $\theta_{23}$  to vary, even dramatically. It is the goal of this article to prove that these two quantities do vary arbitrarily, since any finite sequence of points in the  $(e_2, \theta_{12})$ -cylinder is shadowed by the projection of some solutions of the 4-body problem.

The following theorem is a more precise statement of some of these assertions, in terms of the normalised angular momentum vector

$$(1) \quad \tilde{C}_2 = \frac{\sqrt{m_0 + m_1 + m_2}}{m_2(m_0 + m_1)\sqrt{a_2}} C_2$$

of planet 2, which lies in the unit Euclidean ball  $B^3$  since  $|\tilde{C}_2| = \sqrt{1 - e_2^2}$  and  $e_2 \in (0, 1)$ . An even more precise statement will be given in Section 2. Fix  $0 < \eta \ll 1$  and choose masses in the set

$$(2) \quad \mathcal{M}_\eta = \left\{ (m_0, m_1, m_2, m_3) \in (0, +\infty)^4, \quad \begin{array}{l} m_0 + m_1 + m_2 + m_3 = 1, \\ |m_0 - m_1| \geq \eta, \quad |m_0 + m_1 - m_2| \geq \eta \end{array} \right\};$$

<sup>9</sup> We will call it the Lidov-Kozai effect, as suggested by A. Neishtadt and [85].

<sup>10</sup> The Lidov-Kozai mechanism has had useful many applications to a variety of systems from planetary and stellar scales to supermassive black holes. The orbits' eccentricity can reach extreme values, leading to a nearly radial motion, which can further evolve into short orbit periods and merging binaries. Furthermore, the orbits' mutual inclinations may change dramatically from pure prograde to pure retrograde, leading to misalignment and a wide range of inclinations. These dynamics are accessible from a large part of the triple-body parameter space and can be applied to a diverse range of astrophysical settings and used to gain insights into many puzzles [85]. But since it is an integrable behavior, it cannot bring light to Herman's conjectures.



the sum of masses is fixed in order to avoid non-compactness issues, which are artificial due to the invariance of the dynamics with respect to a change of mass units; and the two inequalities are meant to avoid some degeneracies in the secular dynamics. (Note that in the planetary regime these degeneracies do not exist.)

**Theorem 4 (Main result).** — Fix  $\eta > 0$ . Consider  $(m_0, m_1, m_2, m_3) \in \mathcal{M}_\eta$  and any finite sequence of points in  $\mathbf{B}^3$ . For every  $0 < \delta \ll 1$ , there exists an orbit whose normalised angular momentum  $\tilde{\mathbf{C}}_2$  passes successively  $\delta$ -close to each point of the prescribed itinerary.

The statement calls for a few comments:

- (1) The orbits of interest will be found first in the *hierarchical regime*  $a_1 \ll a_2 \ll a_3$ , with fixed masses such that

$$m_0 \neq m_1 \quad \text{and} \quad m_0 + m_1 \neq m_2.$$

- (2) Now consider the *planetary regime*, where  $m_0 = 1$  and  $m_i = \rho \tilde{m}_i$  ( $i = 1, 2, 3$ ), with  $\rho \rightarrow 0$ . A more precise statement of the Theorem 4 in Section 2 will show that the semimajor axes may be chosen independently from  $0 \leq \rho \leq 1$ , i.e. the conclusion holds in the planetary problem.
- (3) Let us be more specific on the orbit of the conclusion. Let  $\tilde{\mathbf{C}}_2^0, \dots, \tilde{\mathbf{C}}_2^N \in \mathbf{B}^3$  be the prescribed itinerary. These points determine values

- $e_2^k$  of the eccentricity
- $\theta_{23}^k$  of the inclination between planets 2 and 3 (the inclination of planet 3 being nearly fixed)
- and  $h_2^k$  of the longitude of the node of planet 2.

The proof will show that there exist times  $t_0 < t_1 < \dots < t_N$  such that the osculating orbital elements satisfy, as stated,

$$(3) \quad \begin{cases} |e_2(t_k) - e_2^k| & \leq \delta \\ |\theta_{23}(t_k) - \theta_{23}^k| & \leq \delta \\ |h_2(t_k) - h_2^k| & \leq \delta \end{cases} \quad (k = 0, \dots, N).$$

Moreover, the orbit can be chosen so that

$$(4) \quad \begin{cases} |e_1(t_k)| & \leq \delta \\ |\theta_{12}(t_k) - \theta_{12}^k| & \leq \delta \end{cases} \quad (k = 0, \dots, N),$$

where  $\theta_{12}^k \in (0, \pi)$  is determined by the relation

$$(5) \quad (1 - (e_2^0)^2)^{3/2} \cos \theta_{12}^k = (1 - (e_2^k)^2)^{3/2} \cos \theta_{12}^0$$

whereas

$$(6) \quad \begin{cases} |e_3(t) - e_3^0| & \leq \delta \\ |a_j(t) - a_j^0| & \leq \delta \quad \text{for } (j = 1, 2, 3, t \in [0, t_N]). \end{cases}$$

Note that we can choose any initial condition  $e_3^0 \in (0, 1)$  and  $e_3$  remains almost constant along the trajectories we consider.

- (4) We also obtain estimates for the drifting time  $T = t_N$ . In the hierarchical regime where the fixed masses belong to  $\mathcal{M}_\eta$ , we have

$$(7) \quad T = C(m_0, m_1, m_2, m_3) \frac{N}{\delta^\kappa}$$

where  $C$  is a constant depending only on the masses, for some exponent  $\kappa$  which does not depend on  $N$  or the itinerary. To be more precise, call  $\alpha_i = a_i/a_{i+1}$ ,  $i = 1, 2$ , the semimajor axis ratios. As  $\delta$  tends to zero, the  $\alpha_i$ 's will be chosen polynomially smaller, and the drifting time itself depends polynomially on the  $\alpha_i$ 's.

In the planetary regime where  $m_j = \rho \tilde{m}_j$  with  $\rho$  small, the drifting time satisfies

$$(8) \quad T = C(m_0, \tilde{m}_1, \tilde{m}_2, \tilde{m}_3) \frac{N}{\delta^\kappa \rho^2}$$

for some exponent  $\kappa$  which does not depend on  $N$  or the itinerary.

- (5) One may consider the setting in the present paper as a mix between a priori stable and a priori unstable. Indeed, this is a model with multiple time scales and degeneracies. This implies that some directions can be treated as a priori unstable, when one has to face a regular perturbation problem. These are the so-called secular variables. Other directions, which encapsulate the Keplerian mean motions, are much faster. In the present paper we manage to obtain drifting orbits along the a priori unstable directions, that is in the secular actions (angular momenta and inclinations), with a sufficiently robust mechanism, so that fast directions do not interfere with the slow ones. Drift in the actions conjugated to the mean anomalies, that is the semimajor axis, would require a deeper analysis since it would fall into an a priori stable regime. In particular, the drift in actions obtained in Theorem 4 requires time scales which are polynomial in the perturbative parameter. On the contrary, drift of the semimajor axis must take an exponentially long time since Nekhoroshev Theory applies along these directions (see [87]).
- (6) Note that the trajectories that we construct are only of finite length. This is a consequence of the *shadowing* techniques that we employ (see [26]). It may be possible to extend this argument to allow the shadowing of infinite pseudo-orbits, by using the method of cone conditions (see [14, 15]).

Here are some noteworthy consequences.

- (1) Why are inner planets not inclined? If the two inner planets have their mutual inclination close to  $\pi/2$ , Theorem 4 proves that the next planet might be subject to large instabilities. In reality, where semimajor axes are fixed, there is a competition between the semimajor axes ratios and  $e_2$ , which we do not control here. We conjecture that this mechanism leads  $e_2$  to become so large that planets 1 and 2 may collide.

In the solar System, indeed most planets have relatively small mutual inclinations. On the other hand, the two largest dwarf planets Pluto and Eris are trans-Neptunian (i.e. they play the role of our third planet) and have large inclinations to the ecliptic ( $17^\circ$  and  $44^\circ$  respectively).

- (2) Orbits described by Theorem 4 may flip from prograde to retrograde.

In the Solar System, the orbits around the Sun of all planets and most other objects, except many comets, are prograde.<sup>11</sup> However, protoplanetary disks can collide with or steal material from molecular clouds and this can lead to disks and their resulting planets having retrograde orbits around their stars. Retrograde motion may also result from the Lidov-Kozai mechanism, as already mentioned. Here we thus provide another mechanism possibly explaining the existence of some retrograde planets.

- (3) As far as the authors know, up to the present paper there has been no result (analytical or otherwise) on Arnold diffusion in a (non-restricted) N-body problem. As we have already explained, it relies on the existence of normally hyperbolic structure along secular resonances. We consider a purely elliptic regime, that is, at short time scales the bodies perform approximate ellipses, and therefore no hyperbolic invariant objects exist at first order. We rely on the analysis of the perturbed secular dynamics to detect such objects and use its invariant manifolds to obtain drifting orbits. This allows us to obtain Arnold diffusion in the classical planetary regime.

Note that our diffusion mechanism is robust in the following sense; if we consider the N-body problem and assume that the initial conditions of the first four bodies are as in Theorem 4 and the remaining  $N - 4$  bodies revolve sufficiently far away, the conclusion of Theorem 4 still holds. In fact, we expect that analogous diffusive behaviour could be observed in the remaining  $N - 4$  bodies in an appropriate regime. However this would require an inductive argument, the technical details of which would likely be very complicated, largely due to the nature of the coordinates we use (see Section 2.2 below for a description of the Deprit coordinates). A potential inductive step in this process would be to check how these results could be adapted to the 5-body problem [20].

---

<sup>11</sup> They orbit around the Sun in the same direction as the sun rotates about its axis, consistently with the most probable scenarios of formation of the Solar System.

- (4) In a sequel [27] to this paper, we consider a slightly different regime of the 4-body problem (namely, the regime where  $a_3 \gg a_2^3$ ), and we prove that, in addition to the chaotic motions described by Theorem 4, the semimajor axis  $a_3$  can also be made to follow any finite predetermined itinerary. In contrast, in the present paper we assume that  $a_2^{11/6} \ll a_3 \ll a_2^2$  (see (11) below), which implies that the phenomenon of exponentially small splitting of separatrices can be observed in the  $a_3$  direction, and consequently the semimajor axis  $a_3$  is stable in the current regime.

We now consider implications of our main theorem regarding Herman's local conjecture 3, in the case of the spatial 4-body problem. We need the following complement to the main theorem, which follows from our construction and [26, Remark 2.13], because in the secular directions aligned windows are bounded.

*Proposition 5.* — *Transition chains obtained for the secular system (thus disregarding the fast, Keplerian dynamics) have infinite length and the subsequent shadowing orbits are defined for all future times.*

We can thus conclude the following.

*Theorem 6 (Weak local Herman conjecture).* — *Under the hypotheses of Theorem 4, circular orbits are in the closure of the non-recurrent set (as well as of the recurrent set) of the secular dynamics at any finite order.*

The main theorem above (Theorem 4) will be reworded more precisely as two theorems (Theorems 7 and 8 below) which imply Theorem 4 above. These new theorems are more detailed and analyse the dynamics in a different set of coordinates called Deprit variables. The key point of these coordinates is that they are symplectic and moreover adapted to the symplectic reduction the 4 body problem with respect to all its symmetries (translation and rotation).

**1.3. Main ideas of the proof of Theorem 4.** — We first prove the theorem for the hierarchical regime (and fixed values of the masses), and then give a continuation argument which allows to vary the masses within  $\mathcal{M}_\eta$  (see (2)).

The starting point of our construction is a now classical scheme:

- (1) Establish the existence of a normally hyperbolic invariant manifold (see Appendix B for the definition).
- (2) Using Poincaré-Melnikov theory, prove that its stable and unstable manifolds have a transverse homoclinic intersection.
- (3) Compute the first order of the scattering maps in certain variables.
- (4) Using a shadowing theorem, find true orbits that shadow those of the scattering maps.

However there are many caveats. Since we assume that the semimajor axes of the Keplerian ellipses are of different orders, there are numerous time scales. In particular the mean anomalies revolve much faster than the other angles (see below for a description of the angles), and so the splitting of separatrices in the directions of their symplectic conjugate variables is exponentially small. This makes the computation of the scattering maps in those directions a significant challenge. The shadowing results of [26] are well equipped to dealing with problems of this nature; actually, this paper was precisely written for the purpose of its application to the present problem. Therefore we proceed by checking that the assumptions of that paper (summarised in Appendix C) are satisfied by the secular system (defined in Section 3), and thus the full four-body problem.

Beginning with the usual Hamiltonian (see (9) in Section 2) of the four-body problem, we have a conservative system with 12 degrees of freedom. It is well known that this system possesses many integrals of motion. We first pass to Jacobi coordinates to perform the symplectic reduction by translational invariance, which removes 3 degrees of freedom. For a three-body problem, the next step would usually be to pass to orbital elements (i.e. Delaunay variables) in order to perform Jacobi’s classical reduction of the nodes. It is known however that this approach does not extend to the N-body problem when  $N \geq 4$ . Instead, we use the so-called *Deprit variables*. These were introduced by Deprit in the paper [36], but he stated “Whether the new phase variables... are practical in the General Theory of Perturbations is an open question. At least, for planetary theories, the answer is likely to be in the negative”. The variables were subsequently forgotten for a number of years, until they were rediscovered by Chierchia and Pinzari [22, 23], who pointed out that they are in fact *very* useful, by defying Deprit’s advice and nonetheless implementing the coordinates in the planetary problem. The use of these coordinates reduces a further 2 degrees of freedom from our system, resulting in a system with 7 degrees of freedom. To our knowledge, this paper represents the first use of the Deprit variables since the papers of Chierchia and Pinzari.

Denote by  $C_j$  the angular momentum of the  $j^{\text{th}}$  fictitious Keplerian body. The (Deprit) coordinates we are left with after the symplectic reduction are as follows: the mean anomalies  $\ell_1, \ell_2, \ell_3$  and their symplectic conjugates  $L_1, L_2, L_3$  which are proportional to the square root of the semimajor axes; the arguments of the perihelia  $\gamma_1, \gamma_2, \gamma_3$  (*not* with respect to the ascending node, but rather some different nodes; see Section 2 for a precise definition) and their symplectic conjugates the absolute angular momenta  $\Gamma_1 = |C_1|$ ,  $\Gamma_2 = |C_2|$ ,  $\Gamma_3 = |C_3|$ ; an abstractly defined angle  $\psi_1$  and its symplectic conjugate  $\Psi_1 = |C_1 + C_2|$ .

In these coordinates the Hamiltonian can be written  $H = F_{\text{Kep}} + F_{\text{per}}$  where the *Keplerian function*  $F_{\text{Kep}}$  describes the motion of three uncoupled two-body problems, and the *perturbing function*  $F_{\text{per}}$  describes the gravitational forces between the planets. The assumptions we make on the semimajor axes of the Keplerian ellipses imply that the mean anomalies evolve much faster than the other angles, and so we can use the averaging theory to make a near-to-the-identity symplectic transformation that averages the angles

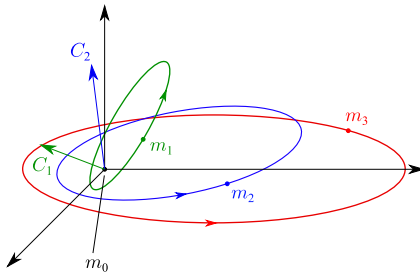


FIG. 1. — We assume that the semimajor axes of the Keplerian ellipses are of different orders. Moreover, assuming that the mutual inclination of bodies 1 and 2 (i.e. the angle between the angular momentum vectors  $C_1$  and  $C_2$  in the figure) is sufficiently large, the first order term in the expansion of the secular Hamiltonian has a saddle periodic orbit. Our main theorems say that, given any predetermined itinerary of mutual inclinations between bodies 2 and 3, and eccentricities of the 2<sup>nd</sup> ellipse, there exist trajectories of the four-body problem shadowing that itinerary

$\ell_j$  out of  $F_{\text{per}}$  up to arbitrarily high order. The Hamiltonian resulting from averaging  $F_{\text{per}}$  over the angles  $\ell_j$  is called the *secular* Hamiltonian and is denoted by  $F_{\text{sec}}$ . Since we assume that the semimajor axes are of different orders we have  $L_1 \ll L_2 \ll L_3$  and so, using the Legendre polynomials, we can expand  $F_{\text{sec}}$  in powers of  $\frac{L_1}{L_2}$  and  $\frac{L_2}{L_3}$ . The first order terms in the expansion of the secular Hamiltonian are referred to as the *quadrupolar* and *octupolar* Hamiltonians of the interactions between bodies 1 and 2 ( $F_{\text{quad}}^{12}$  and  $F_{\text{oct}}^{12}$ ), and the interactions between bodies 2 and 3 ( $F_{\text{quad}}^{23}$  and  $F_{\text{oct}}^{23}$ ); all other terms are of higher order, and are not required for our analysis. All parts of the proof described so far are contained in Section 3.

In order to continue, we observe that the actions are of different orders; indeed,  $\Gamma_1 = O(L_1) = O(1)$ ,  $\Gamma_2 = O(L_2)$ ,  $\Gamma_3 = O(L_3)$ ,  $\Psi_1 = O(L_2)$ . As it is more convenient to deal with actions of order 1, we perform a linear symplectic coordinate transformation and we “chop” the new action space in rectangles with sizes of order 1. We perform our analysis in each of these rectangles. Since the assumptions in [26] are local, one can verify them in each rectangle separately. This coordinate transformation leaves the variables  $\gamma_1, \Gamma_1$  unchanged, and we denote by  $\tilde{\gamma}_j, \tilde{\Gamma}_j, \tilde{\psi}_1, \tilde{\Psi}_1$  the new variables. Substituting these variables into the Hamiltonians  $F_{\text{quad}}^{12}, F_{\text{oct}}^{12}, F_{\text{quad}}^{23}$  and  $F_{\text{oct}}^{23}$  illuminates the time scales of our problem, and allows us to perform further Taylor expansions of each of these Hamiltonians, keeping a careful account of the term in which each variable appears for the first time (see Proposition 11). This linear change, localization and expansion of the secular Hamiltonian are also performed in Section 3.

Part of the beauty of the Deprit variables is that the Hamiltonians  $F_{\text{quad}}^{12}$  and  $F_{\text{oct}}^{12}$  are identical to the quadrupolar and octupolar Hamiltonians (respectively) in the three-body problem expressed in Delaunay coordinates (see [41], for example). Moreover, we choose the coordinate transformation in Section 3 carefully to ensure that the first-order term in the expansion of  $F_{\text{quad}}^{12}$  is identical to the first-order term in the expansion of the quadrupolar Hamiltonian in [41] (modulo some errata; see Appendix F). This Hamiltonian, which we denote by  $H_0^{12}$ , possesses a well-known saddle with a separatrix (see

for example [63]) whenever the mutual inclination between bodies 1 and 2 is sufficiently large (see Figure 1); moreover, the saddle is present for an interval of values of the energy  $H_0^{12}$ . Collecting these saddles on this interval of energy levels, we obtain a normally hyperbolic invariant manifold  $\Lambda$  (see Appendix B for definitions), whose stable and unstable manifolds coincide. Furthermore, Fenichel theory implies that the secular Hamiltonian  $F_{\text{sec}}$  inherits the normally hyperbolic cylinder  $\Lambda$  [44–46]. The results of [41] imply that the octupolar Hamiltonian splits the separatrix in the three-body problem; as the equations are identical,  $F_{\text{oct}}^{12}$  splits the stable and unstable manifolds of  $\Lambda$  in our case too. The manifold  $\Lambda$  for the secular Hamiltonian is written as a graph over the variables  $\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3$ , and so has the structure of a cylinder  $\mathbf{T}^3 \times [0, 1]^3$ . This is described in detail in Section 4.

In Section 5 we analyse the inner dynamics: the restriction of the secular flow to the normally hyperbolic cylinder  $\Lambda$ . The primary goal of this section is to find a new system of coordinates on the cylinder  $\Lambda$ , denoted by  $(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\Gamma}_3, \hat{\Gamma}_3)$ , in which we can continue our analysis. This system is constructed in two steps. The first step produces a coordinate transformation that brings the restriction to  $\Lambda$  of the symplectic structure into its canonical form, using Moser’s trick from his proof of Darboux’s theorem. The second step uses standard methods from averaging theory to push the dependence of the inner Hamiltonian  $F_{\text{sec}}|_{\Lambda}$  on the angles  $\hat{\gamma}_2, \hat{\psi}_1, \hat{\gamma}_3$  to higher order terms.

In Section 6 we prove that there are two homoclinic channels relative to  $\Lambda$  that give rise to two globally defined scattering maps (see Appendix B for definitions). Moreover we compute the first order of the jumps in the scattering maps in the variables  $\hat{\Psi}_1, \hat{\Gamma}_3$ . This section relies on the computation of three Poincaré-Melnikov integrals: the first of these is identical to the computation in [41] (see also Appendix F), whereas the computation of the remaining two integrals is performed in Section 6.2 by considering complex values of time and integrating over certain contours using the residue theorem. Since these computations are conducted in ‘tilde’ variables, we then have to do further computations to determine the jumps in the scattering maps in the ‘hat’ variables (Lemma 30).

In Section 7 we consider the return map to the Poincaré section  $\{\hat{\gamma}_2 = 0\}$  in an energy level of the secular Hamiltonian. The restriction to an energy level eliminates  $\hat{\Gamma}_2$ , and so we are left with a map of a four-dimensional cylinder. We prove that the map satisfies a twist condition, and we show that the jumps in the scattering maps in the  $\hat{\Psi}_1, \hat{\Gamma}_3$  directions are the same as those for the scattering maps corresponding to the flow. We deduce from the formulas for the scattering maps that there are pseudo-orbits (i.e. orbits of the iterated function system consisting of the Poincaré map and the two scattering maps) that follow any itinerary in the actions  $\hat{\Psi}_1, \hat{\Gamma}_3$ , and that the scattering maps map tori corresponding to constant values of  $\hat{\Psi}_1, \hat{\Gamma}_3$  transversely across other such tori. These are the assumptions of the first main theorem of [26] (these results have been summarised in Appendix C – see Theorem 47), and so we obtain orbits of the secular Hamiltonian that follow any predetermined itinerary of the variables  $\hat{\Psi}_1, \hat{\Gamma}_3$  (thus proving Proposition 5). As a consequence, we can then show that the full four-body problem therefore satisfies



the assumptions of the second main theorem in [26] (see Theorem 48 in Appendix C), and so we obtain an analogous result for the full four-body problem. Furthermore, the shadowing methods provide us with (non-optimal) time estimates, which are contained in Section 7.3.

In Section 8 we explain how we can pass from the hierarchical to the planetary case, where the masses of the planets are arbitrarily small of order  $0 < \rho \ll 1$ . We show that the separation of the semimajor axes of the planets can be taken independently of the smallness of the masses of the planets. This effectively amounts to a demonstration that we can rescale the planetary Hamiltonian by  $\rho^{-1}$  and time by a factor of  $\rho^{-2}$  in order to obtain the hierarchical Hamiltonian, where  $\rho$  is the small mass parameter.

## 2. Main results in Deprit coordinates

In Section 2.1 we introduce the Hamiltonian of the four-body problem, perform the reduction by translational symmetry, state our assumptions precisely, and expand the perturbing function using the Legendre polynomials. In Section 2.2 we recall the definition of the Deprit coordinates. Then, in Section 2.3 we give more detailed versions of Theorem 4 in terms of Deprit coordinates.

**2.1. Setting up the problem.** — Consider four point masses in space, interacting via gravitational attraction in the sense of Newton. Denote by  $m_j$  the mass, by  $x_j \in \mathbf{R}^3$  the position, and by  $y_j \in \mathbf{R}^3$  the linear momentum of body  $j$  for each  $j = 0, 1, 2, 3$ . This system is described by the flow of the Hamiltonian function

$$(9) \quad H = \sum_{0 \leq j \leq 3} \frac{y_j^2}{2m_j} - \sum_{0 \leq i < j \leq 3} \frac{m_i m_j}{\|x_j - x_i\|}.$$

Hamilton's equations of motion give a system of 24 differential equations. It is well-known that this system has (at least) 10 integrals of motion: the integral  $H$  corresponds to conservation of energy; 6 integrals correspond to the motion of the barycenter (translational symmetry); and 3 integrals correspond to conservation of angular momentum (rotational symmetry). By making suitable changes of coordinates we can make these integrals visible, thus reducing the equations of motion. First, we pass to Jacobi coordinates to perform the symplectic reduction by translational symmetry, and then use Deprit coordinates to reduce by rotational symmetry.

*Reduction by translational symmetry.* — Let

$$M_j = \sum_{i=0}^j m_i, \quad \sigma_{ij} = \frac{m_i}{M_j},$$

# WHY ARE INNER PLANETS NOT INCLINED?

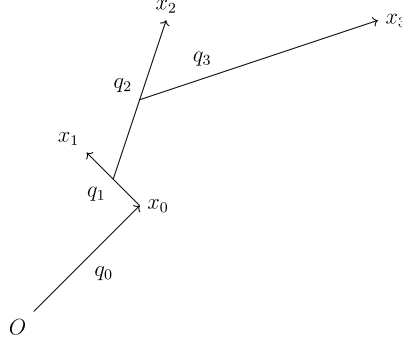


FIG. 2. — Jacobi coordinates

and define the Jacobi coordinates  $(q_j, p_j) \in \mathbf{R}^3 \times \mathbf{R}^3$  for  $j = 0, 1, 2, 3$  by

$$\begin{cases} q_0 = x_0 \\ q_1 = x_1 - x_0 \\ q_2 = x_2 - \sigma_{01} x_0 - \sigma_{11} x_1 \\ q_3 = x_3 - \sigma_{02} x_0 - \sigma_{12} x_1 - \sigma_{22} x_2 \end{cases} \quad \begin{cases} p_0 = y_0 + y_1 + y_2 + y_3 \\ p_1 = y_1 + \sigma_{11} y_2 + \sigma_{11} y_3 \\ p_2 = y_2 + \sigma_{22} y_3 \\ p_3 = y_3. \end{cases}$$

This is a symplectic change of coordinates, and the variable  $q_0$  does not appear in the Hamiltonian (9). The reduced phase space has coordinates  $(q_j, p_j)_{j=1,2,3}$ , and, without loss of generality, we may restrict to  $p_0 = 0$ . Thus write

$$(10) \quad H = F_{\text{Kep}} + F_{\text{per}}$$

where

$$F_{\text{Kep}} = \sum_{j=1}^3 \left( \frac{p_j^2}{2\mu_j} - \frac{\mu_j M_j}{\|q_j\|} \right)$$

and

$$\begin{aligned} F_{\text{per}} = & \sum_{j=2}^3 \frac{\mu_j M_j}{\|q_j\|} - \frac{m_0 m_2}{\|q_2 + \sigma_{11} q_1\|} \\ & - \frac{m_0 m_3}{\|q_3 + \sigma_{22} q_2 + \sigma_{11} q_1\|} - \frac{m_1 m_2}{\|q_2 - \sigma_{01} q_1\|} \\ & - \frac{m_1 m_3}{\|q_3 + \sigma_{22} q_2 + (\sigma_{11} - 1) q_1\|} - \frac{m_2 m_3}{\|q_3 + (\sigma_{22} - 1) q_2\|} \end{aligned}$$

with the reduced masses  $\mu_j$  defined by

$$\mu_j^{-1} = M_{j-1}^{-1} + m_j^{-1} \quad (j = 1, 2, 3).$$

The function  $F_{\text{Kep}}$  is an integrable Hamiltonian describing the motion of 3 uncoupled *Kepler problems*, and  $F_{\text{per}}$  is the so-called *perturbing function*.

*Assumptions and expansion in Legendre polynomials.* — In this paper, we assume that each term of  $F_{\text{Kep}}$  is negative, so each of the 3 uncoupled Keplerian trajectories is elliptical. We assume that the orbit of body 2 is far away from the orbits of the first two bodies. We assume moreover that the orbit of body 3 lies far from the orbits of the first 2 bodies, in a range dictated by the position of body 2. More precisely, if we denote by  $a_j$  the semimajor axis of the  $j^{\text{th}}$  Keplerian ellipse, our assumptions are:

$$(11) \quad \text{O}(1) = a_1 \ll a_2 \rightarrow \infty \quad \text{and} \quad a_2^{\frac{11}{6}} \ll a_3 \ll a_2^2.$$

The purpose of these assumptions is twofold. First, the fact that  $a_1 \ll a_2 \ll a_3 \ll a_2^2$  implies that the three mean anomalies are the fastest three angles, and so we can average out these three angles (see below in this section). Second, the fact that  $a_2^{11/6} \ll a_3$  implies both that the first two terms in the expansion of the secular Hamiltonian come from the quadrupolar Hamiltonian of bodies 1 and 2, and that the first nontrivial term in the Melnikov potential comes from the octupolar Hamiltonian of bodies 1 and 2 (see later).

We are in a near-integrable setting, as  $F_{\text{per}}$  is smaller than  $F_{\text{Kep}}$ . Indeed, denote by  $\zeta_j$  the angle between  $q_j$  and  $q_{j+1}$  for  $j = 1, 2$ , and denote by  $P_n$  the Legendre polynomial of degree  $n$ . Our assumptions imply that  $\|q_1\| \ll \|q_2\| \ll \|q_3\|$ . Therefore we can write the perturbing function as

$$(12) \quad F_{\text{per}} = F_{\text{per}}^{12} + F_{\text{per}}^{23} + \text{O}\left(\frac{\|q_1\|^2}{\|q_3\|^3}\right)$$

where

$$(13) \quad F_{\text{per}}^{12} = \frac{\mu_2 M_2}{\|q_2\|} - \frac{m_0 m_2}{\|q_2 + \sigma_{11} q_1\|} - \frac{m_1 m_2}{\|q_2 - \sigma_{01} q_1\|} = -\frac{\mu_1 m_2}{\|q_2\|} \sum_{n=2}^{\infty} \tilde{\sigma}_{1,n} P_n(\cos \zeta_1) \left(\frac{\|q_1\|}{\|q_2\|}\right)^n$$

is the perturbing function from the 3-body problem (see [41], for example), and where

$$(14) \quad F_{\text{per}}^{23} = -\frac{\mu_2 m_3}{\|q_3\|} \sum_{n=2}^{\infty} \tilde{\sigma}_{2,n} P_n(\cos \zeta_2) \left(\frac{\|q_2\|}{\|q_3\|}\right)^n$$

with

$$\tilde{\sigma}_{1,n} = \sigma_{01}^{n-1} + (-1)^n \sigma_{11}^{n-1}, \quad \tilde{\sigma}_{2,n} = (\sigma_{02} + \sigma_{12})^{n-1} + (-1)^n \sigma_{22}^{n-1}.$$

**2.2. The Deprit coordinates and reduction by rotational symmetry.** — To prove Theorem 4 and to take advantage of the rotational symmetry, we rely on Deprit coordinates. These

coordinates were discovered originally by Deprit [36] and rediscovered recently by Chierchia and Pinzari [22].

Denote by

$$\mathbf{C}_j = q_j \times p_j$$

the angular momentum of the  $j^{\text{th}}$  fictitious Keplerian body (Keplerian referring to  $F_{\text{Kep}}$ ), and let  $k_j$  be the  $j^{\text{th}}$  element of the standard orthonormal basis of  $\mathbf{R}^3$ . Define the nodes  $v_j$  by

$$v_1 = v_2 = \mathbf{C}_1 \times \mathbf{C}_2, \quad v_3 = (\mathbf{C}_1 + \mathbf{C}_2) \times \mathbf{C}_3, \quad v_4 = k_3 \times \mathbf{C}$$

where

$$\mathbf{C} = \mathbf{C}_1 + \mathbf{C}_2 + \mathbf{C}_3$$

is the total angular momentum vector. For a non-zero vector  $z \in \mathbf{R}^3$  and two non-zero vectors  $u, v$  lying in the plane orthogonal to  $z$ , denote by  $\alpha_z(u, v)$  the oriented angle between  $u, v$ , with orientation defined by the right hand rule with respect to  $z$ . Denote by  $\Pi_j$  the pericenter of  $q_j$  on its Keplerian ellipse. The Deprit variables  $(\ell_j, L_j, \gamma_j, \Gamma_j, \psi_j, \Psi_j)_{j=1,2,3}$  are defined as follows:

- $\ell_j$  is the mean anomaly of  $q_j$  on its Keplerian ellipse;
- $L_j = \mu_j \sqrt{M_j a_j}$ ;
- $\gamma_j = \alpha_{\mathbf{C}_j}(v_j, \Pi_j)$  (see Figure 3);
- $\Gamma_j = \|\mathbf{C}_j\|$ ;
- $\psi_1 = \alpha_{(\mathbf{C}_1 + \mathbf{C}_2)}(v_3, v_2)$ ,  $\psi_2 = \alpha_{\mathbf{C}}(v_4, v_3)$ ,  $\psi_3 = \alpha_{k_3}(k_1, v_4)$ ;
- $\Psi_1 = \|\mathbf{C}_1 + \mathbf{C}_2\|$ ,  $\Psi_2 = \|\mathbf{C}_1 + \mathbf{C}_2 + \mathbf{C}_3\| = \|\mathbf{C}\|$ ,  $\Psi_3 = \mathbf{C} \cdot k_3$ .

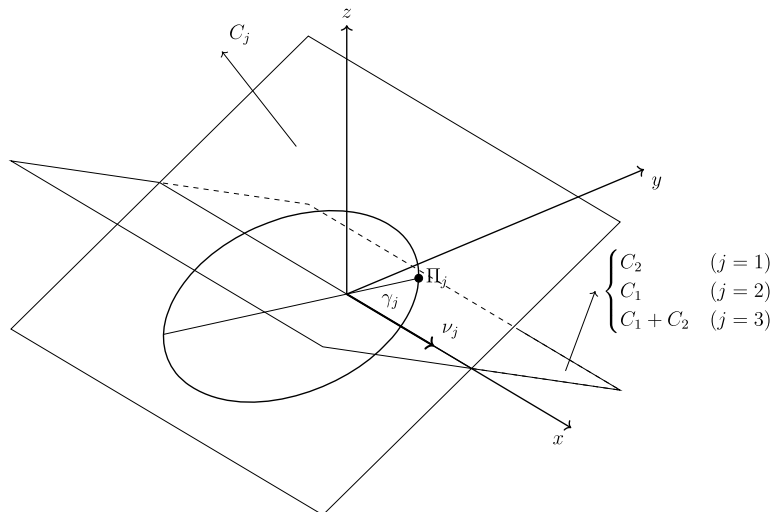
The Deprit variables are analytic over the open subset  $\mathcal{D}$  over which the 3 terms of  $F_{\text{Kep}}$  are negative, the eccentricities of the Keplerian ellipses lie strictly between 0 and 1, and the nodes  $v_j$  are nonzero.

The Deprit variables form a symplectic coordinate system over  $\mathcal{D}$ . In Appendix A we give an alternative proof of its symplecticity from that in [22, 36].

In these coordinates the Keplerian Hamiltonian becomes

$$F_{\text{Kep}} = - \sum_{j=1}^3 \frac{\mu_j^3 M_j^2}{2L_j^2}$$

and so the mean anomalies  $\ell_j$  undergo a rigid rotation with frequency  $\sqrt{\frac{M_j}{a_j^3}}$ , while all other variables remain fixed. Moreover the variables  $\psi_3, \Psi_2, \Psi_3$  are integrals of motion for the full 4-body problem (due to conservation of angular momentum), and so the Hamiltonian does not depend on  $\Psi_3, \psi_2, \psi_3$ .


 FIG. 3. — The node  $\nu_j$  and the argument  $\gamma_j$  of pericenter  $\Pi_j$ 

Note that one can express the orbital elements in terms of Deprit as follows. The osculating eccentricities are defined by

$$(15) \quad e_j = \sqrt{1 - \frac{\Gamma_j^2}{L_j^2}}, \quad j = 1, 2, 3.$$

The mutual inclination between the bodies 1 and 2 is given by  $i_{12}$ , which is defined by

$$(16) \quad \cos i_{12} = \frac{\Gamma_1^2 + \Gamma_2^2 - \Psi_1^2}{2\Gamma_1\Gamma_2}.$$

For the third planet one can measure its inclination with respect to the plane orthogonal to the vector  $S_1 = C_1 + C_2$ , that is with respect to the plane orthogonal to the angular momentum given by planet 1 and 2. If one denotes this angle by  $i_{23}$ , one has

$$(17) \quad \cos i_{23} = \frac{\Gamma_3^2 + \Psi_1^2 - \Psi_2^2}{2\Gamma_3\Psi_1}.$$

Note that we are in a regime where  $L_2 \gg L_1$  and therefore the angular momentum of the first two planets is essentially carried by planet 2.

**2.3. Arnold diffusion in Deprit coordinates.** — In this paper we consider both the hierarchical and planetary regimes. We first state the results in the hierarchical regime, which imply Theorem 4 for fixed values of masses. Later, we consider the planetary regime, that is we take masses of the planets arbitrarily small.

To this end, let us specify what is the hierarchical regime in Deprit coordinates. For now we assume that the masses  $m_0, m_1, m_2, m_3 > 0$  are fixed and belong to  $\mathcal{M}_\eta$  (see (2)).

Then, we consider that semimajor axes of the bodies are well separated

$$a_1 \ll a_2 \ll a_3$$

and equivalently

$$L_1 \ll L_2 \ll L_3.$$

Note that we assume that the eccentricities of the bodies are uniformly bounded away from 0 and 1 and therefore

$$(18) \quad \Gamma_i \sim L_i \quad \text{for } i = 1, 2, 3 \quad \text{and} \quad \Psi_i \sim \Gamma_{i+1} \quad \text{for } i = 1, 2$$

(recall that  $\Psi_2$  is the total angular momentum which is a conserved quantity).

In this regime the semimajor axes are very stable. The same happens for  $\Gamma_3$  due to the conservation of angular momentum vector. However a “small” transfer of angular momentum from the third planet to the second can have a big effect on the eccentricity and inclination of the second planet.

In particular,  $\Gamma_3$  may transition from

$$(19) \quad \Gamma_3 \sim \Psi_2 - \Psi_1 \quad \text{to} \quad \Gamma_3 \sim \Psi_2 + \Psi_1.$$

This corresponds to having  $C_3$  and  $C_1 + C_2$  close to parallel and either with the same sense or opposite sense. Such transition is equivalent to make a transition in the inclination  $\theta_{23}$  (see Section 9 below).

On the other hand, note that  $0 < \Gamma_2 < L_2$ . So the maximal transition  $\Gamma_2$  can make is

$$(20) \quad \Gamma_2 \sim L_2 \quad \text{to} \quad \Gamma_2 \sim 0.$$

By (15), this corresponds to the second planet transitioning from a close to circular orbit ( $e_2 \sim 0$ ) to a highly eccentric one ( $e_2 \sim 1$ ).

The next theorem shows that such transitions are possible and that one can freely vary  $\Gamma_3$  and  $\Gamma_2$  within their admitted ranges. In the constructed orbits the changes in the other actions are determined by those two.

*Theorem 7 (Hierarchical regime). — Fix masses  $m_0, m_1, m_2, m_3 > 0$  such that*

$$m_0 \neq m_1 \quad \text{and} \quad m_0 + m_1 \neq m_2.$$

*There exists  $0 < \kappa \ll 1$ ,  $\alpha > 0$ ,  $\beta > 0$ , such that the following is satisfied.*

*Fix  $N \geq 1$  any  $\{v^k\}_{k=0}^N \subset (0, 1)$ ,  $\{\eta^k\}_{k=0}^N \subset (-1, 1)$  and constants  $\Gamma_3^0 \in [\kappa L_3^0, (1 - \kappa)L_3^0]$ ,  $\Psi_1^0, \Gamma_2^0 \in [\kappa L_2^0, (1 - \kappa)L_2^0]$  such that  $|\Psi_1^0 - \Gamma_2^0| \leq \kappa$ ,  $\Psi_2^0$  such that  $\Psi_2^0 \in [\Gamma_3^0 - (1 - \kappa)\Psi_1^0, \Gamma_3^0 + (1 - \kappa)\Psi_1^0]$  and  $L^0 = (L_1^0, L_2^0, L_3^0)$  satisfying  $L_1^0 \in [1/2, 2]$  and*

$$(21) \quad L_1^0 \ll L_2^0 \quad \text{and} \quad (L_2^0)^{\frac{11}{6}} \ll L_3^0 \ll (L_2^0)^2.$$

Then, there exists an orbit of the Hamiltonian  $\mathbf{H}$  in (10) expressed in Deprit coordinates and times  $\{t_k\}_{k=0}^N$  satisfying

$$t_0 = 0 \quad \text{and} \quad |t_k| \leq (\mathbf{L}_2^0)^\alpha, \quad k \geq 1$$

such that

$$|\Gamma_2(t_k) - \nu_k \mathbf{L}_2^0| \leq (\mathbf{L}_2^0)^{-\beta}, \quad |\Psi_2^0 - \Gamma_3(t_k) - \eta_k \Psi_1(t_k)| \leq (\mathbf{L}_2^0)^{-\beta}$$

and

$$|\Gamma_1(t_k) - \mathbf{L}_1^0| \leq (\mathbf{L}_2^0)^{-\beta}, \quad |\Psi_1(t_k) - \Gamma_2(t_k) - \mathbf{M}_k| \leq (\mathbf{L}_2^0)^{-\beta}$$

where  $\mathbf{M}_k \in (0, \kappa)$  is determined by

$$\frac{\mathbf{M}_k^2}{(\mathbf{L}_2^0 - \Gamma_2(t_k))^{3/2}} = \frac{\mathbf{M}_0^2}{(\mathbf{L}_2^0 - \Gamma_2^0)^{3/2}} \quad \text{and} \quad \mathbf{M}_0 = \Psi_1^0 - \Gamma_2^0$$

whereas for all  $t \in [0, t_N]$ ,

$$|\Gamma_3(t) - \Gamma_3^0| \leq 2\mathbf{L}_2^0, \quad \text{and} \quad |\mathbf{L}_j(t) - \mathbf{L}_j^0| \leq (\mathbf{L}_2^0)^{-\beta} \quad \text{for } j = 1, 2, 3.$$

One can obtain an analogous statement in the planetary regime, that is when one considers arbitrarily small masses for the bodies 1,2,3. That is,  $m_0 \sim 1$  whereas  $m_i = \rho \tilde{m}_i$ ,  $i = 1, 2, 3$ , with  $\tilde{m}_i \sim 1$  and  $0 < \rho \ll 1$ . Note that when  $\rho$  tends to 0, the actions  $(\mathbf{L}, \Gamma, \Psi)$  all satisfy

$$\mathbf{L}, \Gamma, \Psi \rightarrow 0.$$

Therefore, to be able to capture the drift in actions it is convenient to perform a symplectic scaling

$$\mathbf{L} = \rho \tilde{\mathbf{L}}, \quad \Gamma = \rho \tilde{\Gamma}, \quad \Psi = \rho \tilde{\Psi},$$

Note that all the orbital elements are homogeneous functions of degree 0 of the Deprit coordinates (see (15), (16), (17)). Therefore, the drift in the scaled coordinates will determine the behavior described in Theorem 4 for the planetary regime.

**Theorem 8 (Planetary regime).** — Fix  $m_0, \tilde{m}_1, \tilde{m}_2, \tilde{m}_3 > 0$  and consider the Hamiltonian  $\mathbf{H}$  in (10) expressed in Deprit coordinates with masses  $m_0, m_j = \rho \tilde{m}_j$  with  $j = 1, 2, 3$ . There exists  $0 < \kappa_0 \ll 1, \alpha > 0, \beta > 0$ , such that the following is satisfied for any  $\kappa \in (0, \kappa_0)$ .



Fix  $N \geq 1$  any  $\{\nu^k\}_{k=0}^N \subset (0, 1)$ ,  $\{\eta^k\}_{k=0}^N \subset (-1, 1)$  and constants  $\tilde{\Psi}_1^0, \tilde{\Psi}_2^0, \tilde{\Gamma}_2^0, \tilde{\Gamma}_3^0, \tilde{L}_1^0, \tilde{L}_2^0, \tilde{L}_3^0$  such that

$$\begin{cases} \tilde{\Psi}_1^0, \tilde{\Gamma}_2^0 \in [\kappa \tilde{L}_2, (1 - \kappa) \tilde{L}_2] \\ |\tilde{\Psi}_1^0 - \tilde{\Gamma}_2^0| \leq \kappa \\ \tilde{\Gamma}_3^0 \in (\kappa \tilde{L}_3, (1 - \kappa) \tilde{L}_3) \\ \tilde{\Psi}_2^0 \in [\tilde{\Gamma}_3^0 - (1 - \kappa) \tilde{\Psi}_1^0, \tilde{\Gamma}_3^0 + (1 - \kappa) \tilde{\Psi}_1^0] \end{cases}$$

and

$$\tilde{L}_1^0 \in [1/2, 2], \quad \tilde{L}_1^0 \ll \tilde{L}_2^0 \quad \text{and} \quad (\tilde{L}_2^0)^{\frac{11}{6}} \ll \tilde{L}_3^0 \ll (\tilde{L}_2^0)^2.$$

Then, there exists  $\rho_0$  such that for any  $\rho \in (0, \rho_0)$ , there exists an orbit of the Hamiltonian  $H$  in (10) expressed in scaled Deprit coordinates and times  $\{t_k\}_{k=0}^N$  satisfying  $t_0 = 0$  and

$$|t_k| \leq C \frac{(\tilde{L}_2^0)^\alpha}{\rho^2}$$

with  $C > 0$  independent of  $\rho$  and  $\tilde{L}_2^0$ , such that

$$|\tilde{\Gamma}_2(t_k) - \nu_k \tilde{L}_2^0| \leq (\tilde{L}_2^0)^{-\beta}, \quad |\tilde{\Psi}_2^0 - \tilde{\Gamma}_3(t_k) - \eta_k \tilde{\Psi}_1(t_k)| \leq (\tilde{L}_2^0)^{-\beta}$$

and

$$|\tilde{\Gamma}_1(t_k) - \tilde{L}_1^0| \leq (\tilde{L}_2^0)^{-\beta}, \quad |\tilde{\Psi}_1(t_k) - \tilde{\Gamma}_2(t_k) - M_k| \leq (\tilde{L}_2^0)^{-\beta}$$

where  $M_k \in (0, \kappa)$ ,  $k = 1 \dots N$ , is determined by

$$M_k^2 = \frac{(\tilde{\Psi}_1^0 - \tilde{\Gamma}_2^0)^2}{(\tilde{L}_2^0 - \tilde{\Gamma}_2^0)^{3/2}} (\tilde{L}_2^0 - \tilde{\Gamma}_2(t_k)^2)^{3/2}$$

whereas for all  $t \in [0, t_N]$ ,

$$|\tilde{\Gamma}_3(t) - \tilde{\Gamma}_3^0| \leq 2\tilde{L}_2^0 \quad \text{and} \quad |\tilde{L}_j(t) - \tilde{L}_j^0| \leq (\tilde{L}_2^0)^{-\beta} \quad \text{for } j = 1, 2, 3.$$

Sections 3 - 7 are devoted to prove Theorem 7. Then, in Section 8 we explain how to extend the proof in the planetary regime.

### 3. The averaging procedure and the secular Hamiltonian

The purpose of this section is to define the secular Hamiltonian, and compute all terms relevant for our later analysis. This involves averaging the perturbing function (12) in Deprit coordinates, and performing a Taylor expansion of the resulting expressions in

powers of the small parameters  $\frac{1}{L_2}$  and  $\frac{L_2}{L_3}$ . In Section 3.1 we define the secular Hamiltonian, and compute the first two terms in its expansion. In Section 3.2, we introduce affine coordinate transformations to split the action space in substrips such that the new action variables in each strip are of order 1. This makes the multiple time scales of the problem visible, and allows us to perform further Taylor expansions.

**3.1. The averaging procedure and the secular Hamiltonian.** — It can be seen in equations (13), (14) that  $F_{\text{per}}^{12}, F_{\text{per}}^{23}$  are of order  $O(\frac{1}{a_2^3}), O(\frac{a_2^2}{a_3^3})$  respectively. Therefore our assumption (11) implies that the mean anomalies  $\ell_j$  have faster frequencies than the other angles, and moreover that the frequencies of the mean anomalies are of different orders. As a result we can apply the normal form theory (see [37, 96]): we may construct an arbitrary number  $k$  of non-resonant normal forms. One can achieve this by constructing successive changes of coordinates as follows. At the step  $j = 1 \dots k - 2$  the averaging transformation is obtained as the composition of two changes of coordinates.

- First, the terms of order  $a_2^{-j-1}$  in the Hamiltonian  $F_{\text{per}}^{12}$  in (13) depending on  $\ell_1$  and  $\ell_2$  are removed (recall that  $F_{\text{per}}^{12}$  is independent of  $\ell_3$ ). Since the small divisors have a lower bound of order  $a_2^{-3/2}$  (the size of the  $\ell_2$ -frequency), this can be achieved by an  $a_2^{-j+1/2}$ -close to the identity symplectic coordinate transformation.
- Then, we apply a second transformation to remove the terms in  $F_{\text{per}} - F_{\text{per}}^{12}$  (see (12)) of size  $a_2^{j+1}/a_3^{j+2}$  depending on the mean anomalies. Since these terms may depend on  $\ell_3$ , the small divisors have size  $a_3^{3/2}$  and therefore these terms can be removed by an  $a_2^{j+1}/a_3^{j-1/2}$ -close to the identity symplectic coordinate transformation.

This leads to the Hamiltonian

$$(22) \quad F = F_{\text{Kep}} + F_{\text{sec},k} + O\left(\left(\frac{1}{a_2}\right)^{k+1}, \frac{a_2^k}{a_3^{k+1}}\right)$$

where  $F_{\text{sec},k}$  is the *secular Hamiltonian* of order  $k$ , defined by

$$(23) \quad F_{\text{sec},k} = F_{\text{sec},k}^{12} + F_{\text{sec},k}^{23} + O\left(\frac{1}{a_3^3}\right)$$

with

$$F_{\text{sec},k}^{12} = \frac{1}{(2\pi)^2} \int_{\mathbf{T}^2} F_{\text{per}}^{12} d\ell_1 d\ell_2 + O\left(\frac{1}{a_2^{9/2}}\right),$$

and

$$F_{\text{sec},k}^{23} = \frac{1}{(2\pi)^2} \int_{\mathbf{T}^2} F_{\text{per}}^{23} d\ell_2 d\ell_3 + \mathcal{O}\left(\frac{a_2^4}{a_3^{9/2}}\right).$$

In what follows we drop the  $k$  subscript and simply write  $F_{\text{sec}}^{12}$  and  $F_{\text{sec}}^{23}$  to simplify notation.

Now, the first term in the expansions (13), (14) after averaging is called the *quadrupolar* Hamiltonian (with respect to bodies 1 and 2, and with respect to bodies 2 and 3, respectively), and the second is called the *octupolar* Hamiltonian (following the terminology of electrostatic multipoles). Since  $\tilde{\sigma}_{j,2} = 1$  for  $j = 1, 2$ , we write

$$(24) \quad F_{\text{sec}}^{j,j+1} = -\frac{\mu_j m_{j+1}}{(2\pi)^2} \left( F_{\text{quad}}^{j,j+1} + \tilde{\sigma}_{j,3} F_{\text{oct}}^{j,j+1} + \mathcal{O}\left(\frac{a_j^4}{a_{j+1}^5}\right) \right) \quad \text{for } j = 1, 2.$$

The eccentricity of the  $j^{\text{th}}$  Keplerian ellipse is given in terms of Deprit coordinates by

$$(25) \quad e_j = \sqrt{1 - \frac{\Gamma_j^2}{L_j^2}}.$$

**Lemma 9.** — *The quadrupolar and octupolar Hamiltonians of bodies 1 and 2 are given by*

$$(26) \quad F_{\text{quad}}^{12} = \frac{a_1^2}{8 a_2^3 (1 - e_2^2)^{\frac{3}{2}}} \left( (15 e_1^2 \cos^2 \gamma_1 - 12 e_1^2 - 3) \sin^2 i_{12} + 3 e_1^2 + 2 \right)$$

and

$$(27) \quad F_{\text{oct}}^{12} = -\frac{15 a_1^3}{64 a_2^4 (1 - e_2^2)^{\frac{5}{2}}} \frac{e_1 e_2}{\cos \gamma_1 \cos \gamma_2 \left[ \frac{\Gamma_1^2}{L_1^2} (5 \sin^2 i_{12} (6 - 7 \cos^2 \gamma_1) - 3) \right] - 35 \sin^2 \gamma_1 \sin^2 i_{12} + 7} \\ \times \left\{ \begin{array}{l} \cos \gamma_1 \cos \gamma_2 \left[ \frac{\Gamma_1^2}{L_1^2} (5 \sin^2 i_{12} (6 - 7 \cos^2 \gamma_1) - 3) \right] \\ - 35 \sin^2 \gamma_1 \sin^2 i_{12} + 7 \\ + \sin \gamma_1 \sin \gamma_2 \cos i_{12} \left[ \sin^2 i_{12} (4 - 7 \cos^2 \gamma_1) - 3 \right] \\ - 35 \sin^2 \gamma_1 \sin^2 i_{12} + 7 \end{array} \right\}$$

respectively, where  $e_j$  is the eccentricity of the  $j^{\text{th}}$  Keplerian ellipse, and  $i_{12}$  is the mutual inclination of Keplerian bodies 1 and 2, defined by

$$(28) \quad \cos i_{12} = \frac{\Gamma_1^2 + \Gamma_2^2 - \Psi_1^2}{2 \Gamma_1 \Gamma_2}.$$

*Proof.* — The quadrupolar and octupolar Hamiltonians are defined by

$$F_{\text{quad}}^{12} = \int_{\mathbf{T}^2} P_2(\cos \zeta_1) \frac{\|q_1\|^2}{\|q_2\|^3} d\ell_1 d\ell_2, \quad F_{\text{oct}}^{12} = \int_{\mathbf{T}^2} P_3(\cos \zeta_1) \frac{\|q_1\|^3}{\|q_2\|^4} d\ell_1 d\ell_2.$$

In order to measure  $\cos \zeta_1$ , it is sufficient to derive expressions for  $q_1, q_2$  in Deprit coordinates with respect to any orthonormal basis of  $\mathbf{R}^3$ . Recall the definition of the nodes  $v_i$ . As in Claim 1 of [22] we consider the *orbital basis*  $B_i = (k_{i,1}, k_{i,2}, k_{i,3})$  for  $i = 1, 2$ , where

$$k_{i,1} = \frac{v_i}{\|v_i\|}, \quad k_{i,3} = \frac{C_i}{\|C_i\|},$$

and where  $k_{i,2}$  is chosen to make the basis orthonormal. We assume that  $q_i \in \mathbf{R}^3 \setminus \{0\}$ . Let  $\bar{q}_i = \|q_i\|^{-1} q_i$ . Then, by the definition of the angle  $\gamma_i$ , the point  $\bar{q}_i$  is written with respect to  $B_i$  as  $\bar{Q}_i = (\cos(\gamma_i + v_i), \sin(\gamma_i + v_i), 0)$ , where  $v_i$  is the true anomaly corresponding to the mean anomaly  $\ell_i$ . Define the standard rotation matrix by an angle  $\theta \in \mathbf{T}$  around the  $x$ -axis by

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

Since  $v_1 = v_2$ , the change of basis matrix from  $B_1$  to  $B_2$  is  $\mathcal{R}_1(i_{12})$ , and so

$$\cos \zeta_1 = \mathcal{R}_1(i_{12}) \bar{Q}_1 \cdot \bar{Q}_2.$$

A standard computation (see Appendix C of [37]; compare also equations (6), (7) of [41]) completes the proof.  $\square$

*Remark 10.* — Analogous expressions exist for  $F_{\text{quad}}^{23}$  and  $F_{\text{oct}}^{23}$ . However these expressions are long, and difficult to interpret. In order to transform these expressions into a more easily understandable form, we first perform a coordinate transformation that allows us to perform a further Taylor expansion. This coordinate transformation and expansion is carried out in Section 3.2.

**3.2. Expansion of the secular Hamiltonian.** — In what follows, we perform a perturbative analysis with respect to the small parameters  $\frac{1}{L_2}, \frac{L_2}{L_3}, \frac{1}{L_3}$ . In order to do this, first notice that the actions  $\Gamma_1, \Gamma_2, \Psi_1, \Gamma_3$  are of different order; we therefore make an affine symplectic change of variables that results in actions of order 1. Indeed, our assumptions regarding the semi-major axes imply that  $\Gamma_2, \Psi_1$  are of order  $L_2$ , while  $\Gamma_3, \Psi_2$  are of order  $L_3$ . Recall that  $\Psi_2$  is the norm of the total angular momentum, and therefore it is a first integral. We fix it as

$$\Psi_2 = \delta_2 L_3$$

for some fixed  $\delta_2 \in (0, 1)$  independent of  $L_2$  and  $L_3$ . Note that different values of  $\delta_2$  give different (approximate) values for  $\Gamma_3$  (of order  $L_3$ ) and therefore different (approximate) values for the osculating eccentricity of the ellipse of the third planet.

We make the symplectic change of variables (with respect to the symplectic form  $\Omega = \sum_{j=1}^3 d\Gamma_j \wedge d\gamma_j + d\Psi_1 \wedge d\psi_1$ ):

$$(29) \quad \begin{cases} \tilde{\Psi}_1 = \Psi_1 - \delta_1 L_2, & \tilde{\psi}_1 = \psi_1 + \gamma_2 \\ \tilde{\Gamma}_2 = \Psi_1 - \Gamma_2, & \tilde{\gamma}_2 = -\gamma_2 \\ \tilde{\Gamma}_3 = \Psi_2 - \Gamma_3 - \delta_3 L_2, & \tilde{\gamma}_3 = -\gamma_3 \end{cases}$$

where  $\delta_1 \in (0, 1)$  and  $\delta_3 \in (-1, 1)$  are constant with respect to the secular Hamiltonian. Note that this symplectic transformation does not modify the variables  $\gamma_1, \Gamma_1$ .

We assume that

$$(30) \quad \tilde{\Gamma}_2 > 0$$

as the case where  $\tilde{\Gamma}_2$  is negative can be treated analogously. Moreover, we assume that the new actions  $\Gamma_1, \tilde{\Gamma}_2, \tilde{\Gamma}_3, \tilde{\Psi}_1$  live in a compact set away from the origin which is independent of  $L_2$  and  $L_3$ . Indeed, we can choose  $(\tilde{\Gamma}_3, \tilde{\Psi}_1) \in [-1, 1]^2$ . Then, choosing a discrete set of pairs  $(\delta_1, \delta_3)$  appropriately, we can cover the whole domain that we want to analyse (see (18), (19), (20)). Since all the analysis we have to perform can be done locally, it is then enough to do it in each of these rectangles to achieve Arnold diffusion in the whole range of actions.

The rest of this section is dedicated to a Taylor expansion of the secular Hamiltonian in powers of  $\frac{L_2}{L_3}$  and  $\frac{1}{L_2}$ ; observe that these quantities are both small as a result of our assumption (11). The following proposition summarises the results of the rest of this section, and its proof is effectively contained in the subsequent Lemmas 13, 14, 15, and 16 in which we expand  $F_{\text{quad}}^{12}$ ,  $F_{\text{oct}}^{12}$ ,  $F_{\text{quad}}^{23}$ , and  $F_{\text{oct}}^{23}$  respectively. The proposition is important as it tells us the order of the speed of each variable  $\gamma_1, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \Gamma_1, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3$ . It also tells us the first order terms containing products of trigonometric functions of  $\tilde{\psi}_1$  (resp.  $\tilde{\gamma}_3$ ) with functions of  $\gamma_1, \Gamma_1, \tilde{\gamma}_2$ ; this will be of great significance later on, as these will be the lowest order terms that give a nontrivial Poincaré-Melnikov potential in  $\tilde{\psi}_1$  (resp.  $\tilde{\gamma}_3$ ; see Proposition 27 in Section 6, as well as Section 6.2).

**Proposition 11.** — *The secular Hamiltonian (23) has the form*

$$(31) \quad F_{\text{sec}} = c + \sum_{i,j=0}^{\infty} \epsilon^i \mu^j F_{ij}$$

where  $\epsilon = \frac{1}{L_2}$  and  $\mu = \frac{L_2}{L_3}$ . Moreover the terms in this expansion satisfy the following properties.

- (1) *The first two nontrivial terms in the expansion are  $F_{6,0} = \alpha_0^{12} H_0^{12}$ ,  $F_{7,0} = \alpha_1^{12} H_1^{12}$  where  $\alpha_i^{12}$  are nontrivial constants, and where the Hamiltonians  $H_0^{12}$ ,  $H_1^{12}$  are defined by (33) and (34) respectively, are integrable, and do not depend on the masses. The Hamiltonians  $H_0^{12}$ ,*

- $H_1^{12}$  are the first order terms from  $F_{\text{quad}}^{12}$  (see (26)). The variables  $\gamma_1, \Gamma_1, \tilde{\Gamma}_2$  appear in  $H_0^{12}$ , and the action  $\tilde{\Psi}_1$  first appears in  $H_1^{12}$ .
- (2) The angle  $\tilde{\gamma}_2$  first appears in  $H_2^{12}$ , which is contained in  $F_{8,0}$ . The Hamiltonian  $H_2^{12}$  is defined by (38), and is the first order term in the expansion of  $F_{\text{oct}}^{12}$  (see (27)).
- (3) The angle  $\tilde{\psi}_1$  first appears in  $F_{2,6} = H_0^{23}$  where the Hamiltonian  $H_0^{23}$  is defined by (39), and is the first order term coming from  $F_{\text{quad}}^{23}$  (see (24)).
- (4) The term  $H_2^{23}$ , contained in  $F_{3,6}$ , is the first order term containing products of trigonometric functions of  $\psi_1$  with functions of  $\gamma_1, \Gamma_1, \tilde{\gamma}_2$ . The Hamiltonian  $H_2^{23}$  is defined by (40), and comes from  $F_{\text{quad}}^{23}$ .
- (5) The action  $\tilde{\Gamma}_3$  first appears in  $\tilde{H}_3$ , which is contained in  $F_{3,6}$ . The Hamiltonian  $\tilde{H}_3$  is defined by (42), and comes from  $F_{\text{quad}}^{23}$ .
- (6) The angle  $\tilde{\gamma}_3$  first appears in  $H_3^{23}$ , which is contained in  $F_{2,8}$ . The Hamiltonian  $H_3^{23}$  is defined by (50), and is the first order term from  $F_{\text{oct}}^{23}$  (see (24)).
- (7) The term  $H_5^{23}$ , contained in  $F_{3,8}$ , is the first order term containing products of trigonometric functions of  $\tilde{\gamma}_3$  with functions of  $\gamma_1, \Gamma_1, \tilde{\gamma}_2$ . The Hamiltonian  $H_5^{23}$  is defined by (53) and comes from  $F_{\text{oct}}^{23}$ .

*Proof.* — Notice that, as a result of our assumption (11) on the semimajor axes, we have the inequalities

$$(32) \quad \frac{1}{L_2^6} \gg \frac{1}{L_2^7} \gg \frac{L_2^4}{L_3^6} \gg \frac{L_2^5}{L_3^7} \gg \frac{L_2^3}{L_3^6}, \quad \frac{1}{L_2^8} \gg \frac{L_2^3}{L_3^6} \gg \frac{L_2^6}{L_3^8},$$

and, equivalently

$$\epsilon^6 \gg \epsilon^7 \gg \mu^6 \epsilon^2 \gg \mu^7 \epsilon^2 \gg \mu^6 \epsilon^3, \quad \epsilon^8 \gg \mu^6 \epsilon^3 \gg \mu^8 \epsilon^3.$$

Combining these inequalities with the contents of Lemmas 13, 14, 15, and 16 completes the proof of the proposition.  $\square$

**Notation 12.** — Throughout this paper, in order to simplify notation, we use ellipsis to mean the following. Fix some sufficiently large integer  $r \in \mathbf{N}$ . The notation  $F = \epsilon^i \mu^j G + \dots$  means that there are  $\eta_1, \eta_2 \in \mathbf{N}_0$ , not both 0, and a positive constant  $C$  such that

$$\|F - \epsilon^i \mu^j G\|_{C^r} \leq C \epsilon^{i+\eta_1} \mu^{j+\eta_2}.$$

Moreover, we use the expression nontrivial constant to mean a constant depending only on the masses and the parameters  $\delta_j$  that is nonzero for all  $m_0, m_1, m_2, m_3 > 0$ , all  $\delta_1, \delta_2 \in (0, 1)$ , and all  $\delta_3 \in (-1, 1)$ .

**Lemma 13.** — The Hamiltonian  $F_{\text{quad}}^{12}$  can be written in the variables (29) as

$$F_{\text{quad}}^{12} = \tilde{c}_0^{12} + \frac{1}{L_2^6} \alpha_0^{12} H_0^{12} + \frac{1}{L_2^7} \alpha_1^{12} H_1^{12} + \frac{1}{L_2^8} \tilde{\alpha}_2 \tilde{H}_2 + \dots$$

where

$$(33) \quad H_0^{12} = \left(1 - \frac{\Gamma_1^2}{L_1^2}\right) \left[2 - 5 \left(1 - \frac{\tilde{\Gamma}_2^2}{\Gamma_1^2}\right) \sin^2 \gamma_1\right] + \frac{\tilde{\Gamma}_2^2}{L_1^2}$$

$$(34) \quad H_1^{12} = (3H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) - 1) \tilde{\Psi}_1 - 4\tilde{\Gamma}_2 H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) + 3\tilde{\Gamma}_2 - \frac{\Gamma_1^2 \tilde{\Gamma}_2^2}{L_1^2}$$

$$(35) \quad \begin{aligned} \tilde{H}_2 = & (3H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) - 1) \tilde{\Psi}_1^2 + \left(6 - 8H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) - 2\frac{\Gamma_1^2}{L_1^2}\right) \tilde{\Gamma}_2 \tilde{\Psi}_1 \\ & + \frac{1}{8} \left[ \sin^2 \gamma_1 \left(5\Gamma_1^2 - \frac{5\Gamma_1^4}{L_1^2} + \frac{210\Gamma_1^2 \tilde{\Gamma}_2^2}{L_1^2} - \frac{205\tilde{\Gamma}_2^4}{L_1^2} + \frac{205\tilde{\Gamma}_2^4}{\Gamma_1^2}\right) \right. \\ & \left. + \frac{\Gamma_1^4}{L_1^2} - \frac{66\Gamma_1^2 \tilde{\Gamma}_2^2}{L_1^2} + \frac{41\tilde{\Gamma}_2^4}{L_1^2} + 40\tilde{\Gamma}_2^2 \right] \end{aligned}$$

and

$$\alpha_0^{12} = \frac{3L_1^4 M_2^3 \mu_2^6}{8M_1^2 \delta_1^3 \mu_1^4}, \quad \alpha_1^{12} = -\frac{3L_1^4 M_2^3 \mu_2^6}{8M_1^2 \delta_1^4 \mu_1^4}, \quad \tilde{\alpha}_2 = \frac{3L_1^4 M_2^3 \mu_2^6}{4M_1^2 \delta_1^5 \mu_1^4}.$$

Moreover  $F_{\text{quad}}^{12}$  is integrable.

*Proof.* — Recall from the definition of the Deprit variables in Section 2 that  $q_j = \frac{L_j^2}{\mu_j^2 M_j}$ , the eccentricity  $e_j$  is defined by (25), and the inclination  $i_{12}$  is defined via its cosine in (28). It follows from (29) that

$$\Gamma_2 = \delta_1 L_2 + \tilde{\Psi}_1 - \tilde{\Gamma}_2, \quad \Psi_1 = \delta_1 L_2 + \tilde{\Psi}_1.$$

Therefore

$$(36) \quad 1 - e_2^2 = \frac{\Gamma_2^2}{L_2^2} = \delta_1^2 + \frac{1}{L_2} 2\delta_1 (\tilde{\Psi}_1 - \tilde{\Gamma}_2) + \frac{1}{L_2^2} (\tilde{\Psi}_1 - \tilde{\Gamma}_2)^2$$

and

$$(37) \quad \begin{aligned} \cos i_{12} = & \frac{\Gamma_1^2 + \Gamma_2^2 - \Psi_1^2}{2\Gamma_1 \Gamma_2} \\ = & -\frac{\tilde{\Gamma}_2}{\Gamma_1} + \frac{1}{L_2} \frac{\Gamma_1^2 - \tilde{\Gamma}_2^2}{2\delta_1 \Gamma_1} - \frac{1}{L_2^2} \frac{(\Gamma_1^2 - \tilde{\Gamma}_2^2) \tilde{\Psi}_1 - \Gamma_1^2 \tilde{\Gamma}_2 + \tilde{\Gamma}_2^3}{2\delta_1^2 \Gamma_1} \\ & + O\left(\frac{1}{L_2^3}\right). \end{aligned}$$



Combining these formulas with (26) and expanding in powers of  $\frac{1}{L_2}$  yields the formulas (33), (34), and (35). Finally, the integrability of  $F_{\text{quad}}^{12}$  is due to the fact that it does not depend on  $\gamma_2$ , and therefore  $\Gamma_2$  is a constant of motion. Indeed, it is thus a Hamiltonian system with two degrees of freedom and two integrals of motion.  $\square$

**Lemma 14.** — *The Hamiltonian  $F_{\text{oct}}^{12}$  can be written in the rescaled variables (29) as*

$$F_{\text{oct}}^{12} = \frac{1}{L_2^8} \alpha_2^{12} H_2^{12} + \dots$$

where

$$(38) \quad H_2^{12} = \sqrt{1 - \frac{\Gamma_1^2}{L_1^2}} \left\{ \begin{array}{l} \cos \gamma_1 \cos \tilde{\gamma}_2 \left[ \frac{\Gamma_1^2}{L_1^2} \left( 5 \left( 1 - \frac{\tilde{\Gamma}_2^2}{\Gamma_1^2} \right) (6 - 7 \cos^2 \gamma_1) - 3 \right) \right. \\ \quad \left. - 35 \sin^2 \gamma_1 \left( 1 - \frac{\tilde{\Gamma}_2^2}{\Gamma_1^2} \right) + 7 \right] \\ + \frac{\tilde{\Gamma}_2}{\Gamma_1} \sin \gamma_1 \sin \tilde{\gamma}_2 \left[ \frac{\Gamma_1^2}{L_1^2} \left( 5 \left( 1 - \frac{\tilde{\Gamma}_2^2}{\Gamma_1^2} \right) (4 - 7 \cos^2 \gamma_1) - 3 \right) \right. \\ \quad \left. - 35 \sin^2 \gamma_1 \left( 1 - \frac{\tilde{\Gamma}_2^2}{\Gamma_1^2} \right) + 7 \right] \end{array} \right\}$$

and

$$\alpha_2^{12} = -\frac{15}{64} \frac{L_1^6 \mu_2^8 M_2^4}{\mu_1^6 M_1^3} \frac{\sqrt{1 - \delta_1^2}}{\delta_1^5}.$$

*Proof.* — Similarly to the proof of Lemma 13, we combine the formulas for  $a_j$ , the eccentricity  $e_j$ , as well as the expansions (36) and (37) with the formula (27) for  $F_{\text{oct}}^{12}$  and expand in powers of  $\frac{1}{L_2}$  to obtain (38).  $\square$

**Lemma 15.** — *The Hamiltonian  $F_{\text{quad}}^{23}$  can be written in the rescaled variables (29) as*

$$F_{\text{quad}}^{23} = \tilde{c}_0^{23} + \frac{L_2^4}{L_3^6} \alpha_0^{23} H_0^{23} + \frac{L_2^5}{L_3^7} \alpha_1^{23} H_1^{23} + \frac{L_2^3}{L_3^6} \alpha_2^{23} H_2^{23} + \dots$$

with

$$(39) \quad H_0^{23} = \cos^2 \tilde{\psi}_1 + O\left(\frac{L_2}{L_3}, \frac{1}{L_2}\right), \quad H_1^{23} = \sin^2 \tilde{\psi}_1 + O\left(\frac{L_2}{L_3}, \frac{1}{L_2}\right)$$

where the higher order terms depend only on  $\tilde{\psi}_1$ ,  $\tilde{\Psi}_1$ ,  $\tilde{\Gamma}_2$ ,  $\tilde{\Gamma}_3$ , and

$$(40) \quad H_2^{23} = \sqrt{\Gamma_1^2 - \tilde{\Gamma}_2^2} (c_1 \cos \tilde{\psi}_1 \cos \tilde{\gamma}_2 + c_2 \sin \tilde{\psi}_1 \sin \tilde{\gamma}_2) + O\left(\frac{L_2}{L_3}, \frac{1}{L_2}\right)$$

where

$$(41) \quad c_1 = \delta_1^2, \quad c_2 = -(5 - 4\delta_1^2).$$

Moreover, the term of order  $L_2^{-1}$  in the expansion of  $H_0^{23}$  is  $\frac{1}{L_2} \tilde{c}_3 \tilde{H}_3$  where  $\tilde{c}_3$  is a nontrivial constant, and

$$(42) \quad \begin{aligned} \tilde{H}_3 = & \tilde{\Psi}_1 [5(\delta_3^2 - \delta_1^4) \cos^2 \tilde{\psi}_1 - 5\delta_3^2 + 3\delta_1^4] \\ & + \tilde{\Gamma}_2 [5(\delta_1^4 - \delta_1^2 \delta_3^2) \cos^2 \tilde{\psi}_1 + 4\delta_1^2 \delta_3^2 - 3\delta_1^4] \\ & + \tilde{\Gamma}_3 [5(\delta_1^3 \delta_3 - \delta_1 \delta_3) \cos^2 \tilde{\psi}_1 + 5\delta_1 \delta_3 - 4\delta_1^3 \delta_3]. \end{aligned}$$

*Proof.* — By definition, we have

$$(43) \quad F_{\text{quad}}^{23} = \int_{\mathbf{T}^2} P_2(\cos \zeta_2) \frac{\|q_2\|^2}{\|q_3\|^3} d\ell_2 d\ell_3.$$

Denote by  $\mathcal{R}_1(\theta)$ ,  $\mathcal{R}_3(\theta)$  the rotation matrix by an angle  $\theta$  around the  $x$ ,  $z$ -axis respectively, and let  $I_3 = \mathcal{R}_3(\pi)$ . Write  $\bar{q}_j = \|q_j\|^{-1} q_j$ , and  $\bar{Q}_j = (\cos(\gamma_j + v_j), \sin(\gamma_j + v_j), 0)$  where  $v_j$  is the true anomaly corresponding to the mean anomaly  $\ell_j$ . By Proposition 4.1 of [88], we have

$$\bar{q}_2 = \mathcal{R}_3(\psi_3) \mathcal{R}_1(i) \mathcal{R}_3(\psi_2) \mathcal{R}_1(\tilde{i}_2) \mathcal{R}_3(\psi_1) I_3 \mathcal{R}_1(i_2) \bar{Q}_2$$

and

$$\bar{q}_3 = \mathcal{R}_3(\psi_3) \mathcal{R}_1(i) \mathcal{R}_3(\psi_2) I_3 \mathcal{R}_1(i_3) \bar{Q}_3$$

where

$$(44) \quad \cos i = \frac{\Psi_3}{\Psi_2}, \quad \cos \tilde{i}_2 = \frac{\Psi_2^2 + \Psi_1^2 - \Gamma_3^2}{2\Psi_1\Psi_2},$$

$$(45) \quad \cos i_2 = \frac{\Gamma_2^2 + \Psi_1^2 - \Gamma_1^2}{2\Psi_1\Gamma_2}, \quad \cos i_3 = \frac{\Gamma_3^2 + \Psi_2^2 - \Psi_1^2}{2\Psi_2\Gamma_3}.$$

Since the last 3 rotations performed in each expression  $\bar{q}_2, \bar{q}_3$  are the same, they can be ignored in the computation of  $\cos \zeta_2 = \bar{q}_2 \cdot \bar{q}_3$ .

First, we focus on the rotations by the angles  $i_2, i_3$ . Observe that, in our rescaled variables,

$$\cos i_2 = 1 - \frac{1}{L_2^2} \frac{\Gamma_1^2 - \tilde{\Gamma}_2^2}{2\delta_1^2} + O\left(\frac{1}{L_2^3}\right), \quad \sin i_2 = \frac{1}{L_2} \frac{\sqrt{\Gamma_1^2 - \tilde{\Gamma}_2^2}}{\delta_1} + O\left(\frac{1}{L_2^2}\right),$$

$$\begin{aligned}\cos i_3 &= 1 - \left(\frac{L_2}{L_3}\right)^2 \frac{\delta_1^2 - \delta_3^2}{2\delta_2^2} + O\left(\left(\frac{L_2}{L_3}\right)^3, \frac{L_2}{L_3^2}\right), \\ \sin i_3 &= \frac{L_2}{L_3} \frac{\sqrt{\delta_1^2 - \delta_3^2}}{\delta_2} + O\left(\left(\frac{L_2}{L_3}\right)^2, \frac{L_2}{L_3^2}\right).\end{aligned}$$

The square roots in the above terms are real-valued: indeed, the triangle inequality implies that

$$\Psi_1 \leq \Gamma_1 + \Gamma_2, \quad \Psi_2 \leq \Psi_1 + \Gamma_3.$$

Inserting the rescaled variables into these inequalities and using (30) yields

$$\Gamma_1 \geq \tilde{\Gamma}_2 > 0, \quad \delta_1 > \delta_3 > 0.$$

Therefore we can write

$$\begin{aligned}\mathcal{R}_1(i_2) &= \text{Id} + L_2^{-1} M_2 + O(L_2^{-2}), \\ \mathcal{R}_1(i_3) &= \text{Id} + \frac{L_2}{L_3} M_3 + O\left(\left(\frac{L_2}{L_3}\right)^2, \frac{L_2}{L_3^2}\right)\end{aligned}$$

where

$$M_j = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -b_j \\ 0 & b_j & 0 \end{pmatrix} \quad \text{with} \quad \begin{cases} b_2 = \frac{\sqrt{\Gamma_1^2 - \tilde{\Gamma}_2^2}}{\delta_1} \\ b_3 = \frac{\sqrt{\delta_1^2 - \delta_3^2}}{\delta_2}. \end{cases}$$

We thus obtain the expression

$$(46) \quad \cos \zeta_2 = \bar{q}_2 \cdot \bar{q}_3 = W_0 + \frac{L_2}{L_3} W_1 + \frac{1}{L_2} W_2 + O\left(\left(\frac{L_2}{L_3}\right)^2, \frac{1}{L_3}, \frac{1}{L_2}\right)$$

where

$$(47) \quad W_0 = \mathcal{R}_1(\tilde{i}_2) \mathcal{R}_3(\psi_1) I_3 \bar{Q}_2 \cdot I_3 \bar{Q}_3,$$

$$(48) \quad W_1 = \mathcal{R}_1(\tilde{i}_2) \mathcal{R}_3(\psi_1) I_3 \bar{Q}_2 \cdot I_3 M_3 \bar{Q}_3,$$

$$(49) \quad W_2 = \mathcal{R}_1(\tilde{i}_2) \mathcal{R}_3(\psi_1) I_3 M_2 \bar{Q}_2 \cdot I_3 \bar{Q}_3.$$

Recall the Legendre polynomial of degree 2 is  $P_2(x) = \frac{1}{2}(3x^2 - 1)$ . Thus

$$P_2(\cos \zeta_2) = P_2(W_0) + 3 \frac{L_2}{L_3} W_0 W_1 + 3 \frac{1}{L_2} W_0 W_2 + O\left(\left(\frac{L_2}{L_3}\right)^2\right).$$

Integrating these 3 terms separately (using again the technique introduced in Appendix C of [37]) and applying the change of coordinates (29), we obtain

$$F_{\text{quad}}^{23} = \frac{a_2^2}{a_3^3 (1 - e_3^2)^{\frac{3}{2}}} \left( K_0 + \frac{L_2}{L_3} K_1 + \frac{1}{L_2} K_2 + O\left(\left(\frac{L_2}{L_3}\right)^2\right) \right)$$

where

$$\begin{aligned} K_0 &= \frac{1}{8} \left( (15 e_2^2 \cos^2 \tilde{\psi}_1 - 12 e_2^2 - 3) \sin^2 \tilde{i}_2 + 3 e_2^2 + 2 \right), \\ K_1 &= -\frac{3 b_3}{4} \cos \tilde{i}_2 \sin \tilde{i}_2 \left( (1 - e_2^2) + 5 e_2^2 \sin^2 \tilde{\psi}_1 \right), \\ K_2 &= \frac{3 b_2}{4} \cos \tilde{i}_2 \sin \tilde{i}_2 \left( (1 - e_2^2) \cos \tilde{\psi}_1 \cos \tilde{\gamma}_2 - (1 + 4 e_2^2) \sin \tilde{\psi}_1 \sin \tilde{\gamma}_2 \right). \end{aligned}$$

Finally, combining the above equations with (24) and observing that, due to (25),

$$\begin{aligned} e_2^2 &= (1 - \delta_1^2) + \frac{1}{L_2} 2\delta_1 (\tilde{\Gamma}_2 - \tilde{\Psi}_1) + O(L_2^{-2}), \quad e_3^2 = 1 - \delta_2^2 + O\left(\frac{L_2}{L_3}\right), \\ \cos \tilde{i}_2 &= \frac{\delta_3}{\delta_1} + \frac{L_2}{L_3} \frac{\delta_1^2 - \delta_3^2}{2\delta_1 \delta_2} + \frac{1}{L_2} \frac{\delta_1 \tilde{\Gamma}_3 - \delta_3 \tilde{\Psi}_1}{\delta_1^2} + O\left(\frac{1}{L_3}\right), \end{aligned}$$

and

$$\begin{aligned} \sin^2 \tilde{i}_2 &= \left( 1 - \frac{\delta_3^2}{\delta_1^2} \right) - \frac{L_2}{L_3} \frac{2\delta_3}{\delta_1} \frac{\delta_1^2 - \delta_3^2}{2\delta_1 \delta_2} - \frac{1}{L_2} \frac{2\delta_3}{\delta_1} \frac{\delta_1 \tilde{\Gamma}_3 - \delta_3 \tilde{\Psi}_1}{\delta_1^2} \\ &\quad + O\left(\left(\frac{L_2}{L_3}\right)^2, \frac{1}{L_3}\right) \end{aligned}$$

where the higher-order terms depend only on  $\tilde{\Psi}_1, \tilde{\Gamma}_2, \tilde{\Gamma}_3$ , we obtain the expressions (39) and (40). Moreover, expanding  $K_0$  up to terms of order  $L_2^{-1}$  and using the above formulas, we see that the lowest-order term containing  $\tilde{\Gamma}_3$  is the expression  $\tilde{H}_3$  defined by (42).  $\square$

**Lemma 16.** — *The Hamiltonian  $F_{\text{oct}}^{23}$  can be written in the rescaled variables (29) as*

$$F_{\text{oct}}^{23} = \tilde{c}_1^{23} + \frac{L_2^6}{L_3^8} \alpha_3^{23} H_3^{23} + \frac{L_2^7}{L_3^9} \alpha_4^{23} H_4^{23} + \frac{L_2^5}{L_3^8} \alpha_5^{23} H_5^{23} + \dots$$

where:

– The Hamiltonian  $H_3^{23}$  is defined by

$$(50) \quad H_3^{23} = \nu_0 \cos(\tilde{\gamma}_3 + 3\tilde{\psi}_1) + \nu_1 \cos(\tilde{\gamma}_3 + \tilde{\psi}_1)$$

$$+ \nu_2 \cos(\tilde{\gamma}_3 - \tilde{\psi}_1) + \nu_3 \cos(\tilde{\gamma}_3 - 3\tilde{\psi}_1)$$

where

$$(51) \quad \begin{aligned} \nu_0 &= \frac{35}{8\delta_1^3} (\delta_1^2 - 1)(\delta_3 - \delta_1)(\delta_3 + \delta_1)^2, \\ \nu_1 &= -\frac{1}{8\delta_1^3} (3\delta_1^2 - 7)(\delta_3 + \delta_1)(15\delta_3^2 - 10\delta_1\delta_3 - \delta_1^2), \end{aligned}$$

$$(52) \quad \begin{aligned} \nu_2 &= \frac{1}{8\delta_1^3} (3\delta_1^2 - 7)(\delta_3 - \delta_1)(15\delta_3^2 + 10\delta_1\delta_3 - \delta_1^2), \\ \nu_3 &= -\frac{35}{8\delta_1^3} (\delta_1^2 - 1)(\delta_3 + \delta_1)(\delta_3 - \delta_1)^2. \end{aligned}$$

Moreover  $\mathbf{H}_3^{23}$  is the lowest-order term containing  $\tilde{\gamma}_3$ .

- The Hamiltonian  $\mathbf{H}_4^{23}$  does not depend on  $\gamma_1, \Gamma_1, \tilde{\gamma}_2$ . Moreover, the average of  $\mathbf{H}_4^{23}$  with respect to  $\tilde{\psi}_1$  does not depend on  $\tilde{\gamma}_3$ .
- The Hamiltonian  $\mathbf{H}_5^{23}$  is given by

$$(53) \quad \mathbf{H}_5^{23} = \sqrt{\Gamma_1^2 - \tilde{\Gamma}_2^2} (\mathbf{J}_1(\tilde{\psi}_1, \tilde{\gamma}_3) \cos \tilde{\gamma}_2 + \mathbf{J}_2(\tilde{\psi}_1, \tilde{\gamma}_3) \sin \tilde{\gamma}_2)$$

where

$$\begin{aligned} &\mathbf{J}_1(\tilde{\psi}_1, \tilde{\gamma}_3) \\ &= 30\delta_3^2 \sin \tilde{\gamma}_3 \cos \tilde{\psi}_1 \sin \tilde{\psi}_1 - 10\delta_1^2 \sin \tilde{\gamma}_3 \cos \tilde{\psi}_1 \sin \tilde{\psi}_1 \\ &\quad - 20\delta_1\delta_3 \cos \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 + 10\delta_1\delta_3 \cos \tilde{\gamma}_3 \\ &\mathbf{J}_2(\tilde{\psi}_1, \tilde{\gamma}_3) \\ &= -50\delta_1\delta_3 \cos \tilde{\gamma}_3 \cos \tilde{\psi}_1 \sin \tilde{\psi}_1 + \frac{70\delta_3}{\delta_1} \cos \tilde{\gamma}_3 \cos \tilde{\psi}_1 \sin \tilde{\psi}_1 \\ &\quad + \frac{105\delta_3^2}{\delta_1^2} \sin \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 \\ &\quad - 75\delta_3^2 \sin \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 + 25\delta_1^2 \sin \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 \\ &\quad - 35 \sin \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 - \frac{105\delta_3^2}{\delta_1^2} \sin \tilde{\gamma}_3 \\ &\quad + 60\delta_3^2 \sin \tilde{\gamma}_3 - 17\delta_1^2 \sin \tilde{\gamma}_3 + 28 \sin \tilde{\gamma}_3. \end{aligned}$$

*Proof.* — Similarly to the proof of Lemma 15, we use formula (46) for  $\cos \zeta_2$  and the Legendre polynomial  $P_3(x) = \frac{1}{2}(5x^3 - 3x)$  to obtain

$$\begin{aligned} P_3(\cos \zeta_2) &= P_3(W_0) + \frac{L_2}{L_3} \frac{1}{2} (15 W_0^2 - 3) W_1 \\ &\quad + \frac{1}{L_2} \frac{1}{2} (15 W_0^2 - 3) W_2 + O\left(\left(\frac{L_2}{L_3}\right)^2\right). \end{aligned}$$

Computing this using formulas (47), (48), and (49) for  $W_0, W_1, W_2$ , substituting the result into

$$F_{\text{oct}}^{23} = \int_{\mathbf{T}^2} P_3(\cos \zeta_2) \frac{\|q_2\|^3}{\|q_3\|^4} d\ell_2 d\ell_3,$$

integrating these 3 terms separately using the technique introduced in Appendix C of [37], and applying the change of coordinates (29) completes the proof of the lemma.  $\square$

#### 4. Analysis of the first-order Hamiltonian

The purpose of this section is to analyse the first order term  $H_0^{12}$  in the expansion of the secular Hamiltonian. We establish the existence of a saddle periodic orbit (in an interval of energy levels), the stable and unstable manifolds of which coincide. Collecting the saddle periodic orbits in this interval of energy levels yields a normally hyperbolic invariant manifold with a separatrix. In addition, we obtain an explicit time parametrisation of the separatrix.

It follows from Proposition 11 that the first term in the expansion of the secular Hamiltonian  $F_{\text{sec}}$  is  $H_0^{12}$ . A convenient property of the Deprit coordinates is that the quadrupolar Hamiltonian of the interaction between planets 1 and 2 in the 4-body problem (and indeed the N-body problem) coincides with the quadrupolar Hamiltonian from the 3-body problem, expressed in Delaunay coordinates. Furthermore,  $H_0^{12}$  corresponds precisely to its counterpart from the 3-body problem [41] (see also Appendix F). In this section we recall results from [41] regarding the existence of 2 hyperbolic periodic orbits of  $H_0^{12}$ , connected by a separatrix.

Since  $H_0^{12}$  does not depend on  $\tilde{\gamma}_2$ , we may consider  $\tilde{\Gamma}_2$  as a parameter. Differentiating (33), we see that Hamilton's equations of motion are

$$\begin{cases} \dot{\gamma}_1 = \frac{\partial H_0^{12}}{\partial \Gamma_1} = \frac{2\Gamma_1}{L_1^2} [5(1 - \frac{\tilde{\Gamma}_2^2}{\Gamma_1^2}) \sin^2 \gamma_1 - 2] - 10(1 - \frac{\Gamma_1^2}{L_1^2}) \frac{\tilde{\Gamma}_2^2}{\Gamma_1^3} \sin^2 \gamma_1 \\ \dot{\Gamma}_1 = -\frac{\partial H_0^{12}}{\partial \gamma_1} = 5(1 - \frac{\Gamma_1^2}{L_1^2}) (1 - \frac{\tilde{\Gamma}_2^2}{\Gamma_1^2}) \sin 2\gamma_1. \end{cases}$$

We seek equilibria of the Hamiltonian vector field. Although we have assumed that the eccentricities  $e_j$  satisfy  $0 < e_j < 1$ , which implies that  $\Gamma_1 \in (0, L_1)$ , the Hamiltonian  $H_0^{12}$

is analytic on a neighbourhood of the cylinder  $(\gamma_1, \Gamma_1) \in \mathbf{T} \times (0, L_1)$  in  $\mathbf{T} \times \mathbf{R}$ . Observe that  $\dot{\Gamma}_1 = 0$  if  $\Gamma_1 = L_1$ . In this case we have  $\dot{\gamma}_1 = 0$  if and only if

$$(54) \quad \sin^2 \gamma_1 = \frac{2}{5 \left(1 - \frac{\tilde{\Gamma}_2^2}{L_1^2}\right)}.$$

Assuming that

$$(55) \quad |\tilde{\Gamma}_2| < L_1 \sqrt{\frac{3}{5}},$$

equation (54) has two solutions in the interval  $\gamma_1 \in (0, \pi)$  (and two more for  $\gamma_1 \in (\pi, 2\pi)$ ). One of these solutions, which we denote by  $\gamma_1^{\min}$  lies in the interval  $(0, \frac{\pi}{2})$ , and the other solution is  $\gamma_1^{\max} = \pi - \gamma_1^{\min}$ .

We have thus found two equilibria  $(\gamma_1, \Gamma_1) = (\gamma_1^{\min, \max}, L_1)$ . These equilibria correspond to circular ellipses, and it can be shown by making a suitable change of coordinates that they are hyperbolic [41]. Lifting the equilibria to the full phase space  $(\gamma_1, \Gamma_1, \tilde{\gamma}_2, \tilde{\Gamma}_2)$  of  $H_0^{12}$ , we obtain the two hyperbolic periodic orbits

$$Z_{\min, \max}^0(t, \tilde{\gamma}_2^0) = (\gamma_1^{\min, \max}, L_1, \tilde{\gamma}_2^0 + \tilde{\gamma}_2^1(t), \tilde{\Gamma}_2)$$

where  $\tilde{\gamma}_2^0 \in \mathbf{T}$  is the initial condition, and

$$(56) \quad \tilde{\gamma}_2^1(t) = \frac{2 \tilde{\Gamma}_2}{L_1^2} t$$

is determined by differentiating (33) with respect to  $\tilde{\Gamma}_2$  and setting  $\Gamma_1 = L_1$ .

Suppose (55) holds, and recall moreover we have assumed in (30) that  $\tilde{\Gamma}_2 > 0$ . Define the positive constants

$$(57) \quad \chi = \sqrt{\frac{2}{3}} \frac{\tilde{\Gamma}_2}{L_1} \frac{1}{\sqrt{1 - \frac{5}{3} \frac{\tilde{\Gamma}_2^2}{L_1^2}}}, \quad A_2 = \frac{6}{L_1} \sqrt{\frac{2}{3}} \sqrt{1 - \frac{5}{3} \frac{\tilde{\Gamma}_2^2}{L_1^2}}.$$

The proof of the following result is identical to the proof of Lemma 3.1 in [41].

**Lemma 17.** — *There is a heteroclinic orbit of  $H_0^{12}$  joining  $Z_{\max}^0$  and  $Z_{\min}^0$  backward and forward in time respectively. It is defined by the equation*

$$\left(1 - \frac{\tilde{\Gamma}_2^2}{\Gamma_1^2}\right) \sin^2 \gamma_1 = \frac{2}{5}$$

where  $\gamma_1 \in (\gamma_1^{\min}, \gamma_1^{\max}) \subset (0, \pi)$ , and its time parametrisation is given by

$$Z^0(t, \tilde{\gamma}_2^0) = (\gamma_1(t), \Gamma_1(t), \tilde{\gamma}_2(t), \tilde{\Gamma}_2)$$



where

$$(58) \quad \cos \gamma_1(t) = \sqrt{\frac{3}{5}} \frac{\sinh(A_2 t)}{\sqrt{\chi^2 + (1 + \chi^2) \sinh^2(A_2 t)}},$$

$$(59) \quad \Gamma_1(t) = \tilde{\Gamma}_2 \sqrt{\frac{5}{3}} \frac{\sqrt{1 + \frac{3}{5} \frac{L_1^2}{\tilde{\Gamma}_2^2} \sinh^2(A_2 t)}}{\cosh(A_2 t)},$$

and

$$(60) \quad \tilde{\gamma}_2(t) = \tilde{\gamma}_2^0 + \tilde{\gamma}_2^1(t) + \tilde{\gamma}_2^2(t), \quad \text{with } \tilde{\gamma}_2^2(t) = \arctan(\chi^{-1} \tanh(A_2 t)).$$

Even though the Hamiltonian function  $H_0^{12}$  is analytic near  $\{\Gamma_1 = L_1\}$ , the Deprit coordinates, as is the case with Delaunay coordinates, are singular on this hypersurface. Indeed, on the circular ellipse  $\Gamma_1 = L_1$ , the argument  $\gamma_1$  of the perihelion is without meaning. We therefore introduce the Poincaré variables

$$(61) \quad \xi = \sqrt{2(L_1 - \Gamma_1)} \cos \gamma_1, \quad \eta = -\sqrt{2(L_1 - \Gamma_1)} \sin \gamma_1.$$

This is a symplectic change of variables, in the sense that

$$(62) \quad d\xi \wedge d\eta = d\Gamma_1 \wedge d\gamma_1.$$

In these variables, the Hamiltonian  $H_0^{12}$  becomes

$$(63) \quad \tilde{H}_0^{12} = \frac{1}{L_1} \left[ 2\xi^2 - \left( 3 - 5 \frac{\tilde{\Gamma}_2^2}{L_1^2} \right) \eta^2 \right] + \frac{\tilde{\Gamma}_2^2}{L_1^2} + O_2(\xi^2 + \eta^2)$$

and the entire hypersurface  $\{\Gamma_1 = L_1\}$  becomes a single hyperbolic periodic orbit

$$(\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2) = (0, 0, \tilde{\gamma}_2^0 + \tilde{\gamma}_2^1(t), \tilde{\Gamma}_2).$$

Moreover, the heteroclinic connection established in Lemma 17 becomes a homoclinic connection to this hyperbolic periodic orbit.

On the hyperbolic periodic orbit and the separatrix, the energy is given by  $\frac{\tilde{\Gamma}_2^2}{L_1^2}$ . It follows that we have a hyperbolic periodic orbit and a homoclinic connection for each positive value of  $\tilde{\Gamma}_2$  satisfying (55). In other words, the Hamiltonian  $\tilde{H}_0^{12}$  has a normally hyperbolic invariant manifold given by

$$(64) \quad \Lambda_0 = \{(\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2) : (\xi, \eta) = (0, 0), \tilde{\gamma}_2 \in \mathbf{T}, \tilde{\Gamma}_2 \in [\zeta_1, \zeta_2]\}$$

where  $\zeta_1, \zeta_2$  satisfy

$$(65) \quad 0 < \zeta_1 < \zeta_2 < L_1 \sqrt{\frac{3}{5}}.$$

Moreover the stable and unstable manifolds of  $\Lambda_0$  coincide.

## 5. The inner dynamics

The goal of this section is to establish the existence of a normally hyperbolic invariant manifold  $\Lambda$  for the secular Hamiltonian, and to determine a set of coordinates in which the inner dynamics (i.e. the dynamics carried by  $\Lambda$ ) can be analysed.

In Section 5.1, we consider a graph parameterization of the normally hyperbolic invariant cylinder and analyse the pull back of the Hamiltonian  $F_{\text{sec}}$  defined in (31) into the cylinder. We also average this Hamiltonian and show that it is very close to integrable. Finally, in Section 5.2, we compute the first and second derivatives of this Hamiltonian with respect to the actions. Such analysis will be fundamental in Section 7 to prove that this Hamiltonian has torsion.

**5.1. The parametrisation of the cylinder and the inner Hamiltonian.** — The normally hyperbolic invariant manifold  $\Lambda_0$  defined by (64) can be lifted to the full secular phase space simply by including the remaining variables, to obtain

$$(66) \quad \tilde{\Lambda}_0 = \{(\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3) : \xi = \eta = 0, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3 \in \mathbf{T}, \tilde{\Gamma}_2 \in [\zeta_1, \zeta_2], \tilde{\Psi}_1 \in [-1, 1], \tilde{\Gamma}_3 \in [-1, 1]\}.$$

This set obviously remains a normally hyperbolic invariant manifold for  $H_0^{12}$ . It is diffeomorphic to  $\mathbf{T}^3 \times [0, 1]^3$  and its stable and unstable manifolds, each of dimension 7, coincide. In a neighbourhood of  $\tilde{\Lambda}_0$ , the symplectic form is

$$\Omega = d\xi \wedge d\eta + d\tilde{\Gamma}_2 \wedge d\tilde{\gamma}_2 + d\tilde{\Psi}_1 \wedge d\tilde{\psi}_1 + d\tilde{\Gamma}_3 \wedge d\tilde{\gamma}_3$$

due to (62). The restriction of  $\Omega$  to  $\tilde{\Lambda}_0$  is

$$(67) \quad \Omega_0 = \Omega|_{\tilde{\Lambda}_0} = d\tilde{\Gamma}_2 \wedge d\tilde{\gamma}_2 + d\tilde{\Psi}_1 \wedge d\tilde{\psi}_1 + d\tilde{\Gamma}_3 \wedge d\tilde{\gamma}_3.$$

Note that, as a result of our assumption (11) on the semimajor axes of the Keplerian ellipses, we have

$$\frac{L_2^{11}}{L_3^6} \ll 1.$$

The following is the main result of this section.

**Theorem 18.** — *For any  $r \geq 2$  there is  $L_2^* > 0$  such that for all  $L_2 \geq L_2^*$  and all  $L_3 \in [L_3^-, L_3^+]$  (where  $L_3^+ > L_3^- > 0$  depend on  $L_2$ ) the following holds.*

- (1) *There is a  $C^r$  smooth normally hyperbolic invariant manifold  $\Lambda$  of  $F_{\text{sec}}$  that is  $O(L_2^{-1})$  close to  $\tilde{\Lambda}_0$  in the  $C^r$  topology. Moreover the variables  $(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3)$  define coordinates on  $\Lambda$  with respect to which the restriction of the symplectic form  $\Omega$  to  $\Lambda$  is closed and non-degenerate (but not necessarily in Darboux form).*

(2) Fix any  $k_1, k_2 \in \mathbf{N}$ . Then there is a coordinate transformation

$$(68) \quad (\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3) \mapsto (\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3)$$

on  $\Lambda$  that is  $\mathcal{O}(\frac{L_2^{11}}{L_3^6})$  close to the identity in the  $C^r$  topology such that  $\Omega|_\Lambda$  becomes the standard symplectic form, and the secular Hamiltonian  $F_{\text{sec}}$ , when restricted to  $\Lambda$ , becomes

$$(69) \quad \hat{F} = \hat{F}_0(\hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3; \epsilon, \mu) + \epsilon^{k_1} \mu^{k_2} \hat{F}_1(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3; \epsilon, \mu)$$

where  $\hat{F}_0 = \epsilon^6 c_0 \hat{\Gamma}_2^2 + \epsilon^7 \hat{h}_0(\hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3; \epsilon, \mu)$ , where  $\epsilon = \frac{1}{L_2}$ ,  $\mu = \frac{L_2}{L_3}$ , and where the  $C^r$  norms of  $\hat{h}_0$  and  $\hat{F}_j$  are uniformly bounded in  $\epsilon, \mu$  for  $j = 0, 1$ .

**Remark 19.** — The existence of the normally hyperbolic invariant manifold  $\Lambda$  for  $F_{\text{sec}}$  follows from Fenichel theory. Indeed,  $F_{\text{sec}}$  is  $\mathcal{O}(L_2^{-7})$  close to its first-order term  $\alpha_0^{12} L_2^{-6} H_0^{12}$  (see Proposition 11). Since  $L_2^{-1}$  is the small parameter in this instance, and since Fenichel theory applies to vector fields of order  $\mathcal{O}(1)$ , we must scale  $F_{\text{sec}}$  by  $L_2^6$  (by scaling time). We thus obtain a Hamiltonian that is  $\mathcal{O}(L_2^{-1})$  close to its first-order term  $\alpha_0^{12} H_0^{12}$ . Since  $H_0^{12}$  has a normally hyperbolic invariant manifold  $\tilde{\Lambda}_0$  defined by (66), Fenichel theory implies that  $L_2^6 F_{\text{sec}}$  has a normally hyperbolic invariant manifold  $\Lambda$  that is  $\mathcal{O}(L_2^{-1})$  close to  $\tilde{\Lambda}_0$  in the  $C^r$  topology [44–46]. Clearly  $\Lambda$  is a normally hyperbolic invariant manifold for the secular Hamiltonian  $F_{\text{sec}}$  itself; in fact, the invariance is only local, in the sense that the vector field is tangent to  $\Lambda$ , but trajectories can escape through the boundary (see Appendix B). The smoothness  $r$  of  $\Lambda$  in Theorem 18 can be made as large as required by increasing  $L_2^*$  if necessary; indeed,  $\tilde{\Lambda}_0$  is  $C^\infty$ , and the Lyapunov exponents of the flow of  $H_0^{12}$  on  $\tilde{\Lambda}_0$  are 0 in the directions transverse to the flow. We fix some sufficiently large value of  $r$ , and choose  $L_2^*$  large enough so that  $\Lambda$  is  $C^r$  smooth.

**Remark 20.** — To be precise, whenever we make a coordinate transformation (see (68) for example; the same applies to Lemmas 23 and 24 below), we must shrink the range of the actions (see (66)) on the cylinder  $\Lambda$  in order to continue to use the results of Section 4: there is a small  $\varrho > 0$  such that for all sufficiently large  $L_2, L_3$ , the new cylinder is defined for  $\hat{\Psi}_1 \in [-1 + \varrho, 1 - \varrho]$ ,  $\hat{\Gamma}_2 \in [\zeta_1 + \varrho, \zeta_2 - \varrho]$ ,  $\hat{\Gamma}_3 \in [-1 + \varrho, 1 - \varrho]$ . To maintain simplicity of notation, we continue to refer to this cylinder as  $\Lambda$ .

The rest of the section is dedicated to the proof of Theorem 18. Then, in Section 5.2, we analyse the second derivatives  $\hat{h}_0$ . Such analysis will be used later on in Section 7 to show that certain Poincaré map associated to the flow  $\hat{F}_0$  has non-degenerate torsion.

By Remark 19, Fenichel theory guarantees the existence of a function  $\rho : \tilde{\Lambda}_0 \rightarrow \mathbf{R}^2$  such that

$$(70) \quad \Lambda = \text{graph}(\rho) = \{(\rho(y), y) : y = (\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3) \\ \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3 \in \mathbf{T}; \tilde{\Gamma}_2 \in [\zeta_1, \zeta_2]; \tilde{\Psi}_1, \tilde{\Gamma}_3 \in [-1, 1]\}$$

where  $\zeta_1, \zeta_2$  satisfy (65) and such that  $\rho$  is  $O(L_2^{-1})$  close to 0 in the  $C^r$  topology. Note that  $\rho$  represents the values of the Poincaré variables  $\xi, \eta$  on the cylinder  $\Lambda$ , expressed as a function of the variables  $(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3)$ , and so we can consider these variables as coordinates on  $\Lambda$ . The following lemma provides information regarding the orders at which each of these variables first appears in the Taylor expansion of  $\rho$ .

**Lemma 21.** — *The function  $\rho$  can be expanded in the form*

$$(71) \quad \rho = \frac{1}{L_2} \rho_0 + \frac{1}{L_2^2} \rho_1 + \frac{L_2^{10}}{L_3^6} \rho_2 + \hat{\epsilon} \rho_3$$

where  $\hat{\epsilon} \lesssim \frac{L_2^9}{L_3^6}$  and

$$(72) \quad \begin{cases} \rho_0 = \rho_0(\tilde{\Gamma}_2), \\ \rho_1 = \rho_1(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\Psi}_1), \\ \rho_2 = \rho_2(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1), \\ \rho_3 = \rho_3(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3), \end{cases}$$

where the  $C^r$  norms of  $\rho_i$ ,  $i = 0, 1, 2, 3$ , are uniformly bounded with respect to  $L_2$  and  $L_3$ .

*Proof.* — As already explained in Remark 19, the existence of the  $C^r$  normally hyperbolic invariant manifold, its regularity with respect to the parameters, and its graph parametrisation is a direct consequence of Fenichel Theory. It only remains to compute its expansion. We follow the approach considered in [31] by Delshams, de la Llave and Seara (see also [60]). That is why we consider the invariance equation for the graph parametrisation of the cylinder and expand the parametrisation and the equation with respect to the parameters.

Let us denote by

$$x = (\xi, \eta), \quad y = (\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3)$$

the inner and normal variables respectively. Since the unperturbed cylinder (66) is defined by the equation  $x = 0$ , Fenichel Theory ensures the existence of a graph parametrisation of the cylinder  $x = \rho(y) = \rho(y; \varepsilon, \mu)$ .

The invariance of the graph then implies the following. Denote by  $X$  the Hamiltonian vector field (which we will make precise below), and denote by  $X^N$  and  $X^T$  its  $x$  (normal) and  $y$  (tangential) components respectively. Then, the invariance of  $\rho$  implies that

$$X^N(\rho(y), y) = D\rho(y)X^T(\rho(y), y).$$

In order to prove the lemma, we expand this equation and observe the variables on which each term in the expansion of  $\rho$  depends.

Denote by  $\tilde{X}$  the vector field associated to the Hamiltonian  $F_{\text{sec}}$  in (31). Note that the first term is of order  $\varepsilon^6$ . Therefore, to normalise and have a regular perturbation problem we define  $X = \varepsilon^{-6}\tilde{X}$ , which can be written as

$$X(x, y) = \sum_{i,j \geq 1} \varepsilon^{i-6} \mu^j X_{ij}(x, y)$$

where  $X_{ij}$  is the Hamiltonian vector field associated to the Hamiltonian  $F_{ij}$  (see (31)). Note that, on one hand, when  $\mu = 0$ , the terms in the series start with  $i = 6$  and therefore this is a regular perturbation problem in  $\varepsilon$ . On the other hand, the first terms in  $\mu$  are order  $O(\mu^6)$  and therefore, since  $\varepsilon \ll \mu$  the terms in  $\mu$  are also small (in fact, one could define the new small parameter  $\tilde{\mu} = \mu/\varepsilon$  and expand in parameters  $\varepsilon$  and  $\tilde{\mu}$ ).

We analyse the terms in the expansion  $\rho(y; \varepsilon, \mu)$  in two steps. First the case  $\mu = 0$  and then we consider the  $\mu$ -dependent terms.

When  $\mu = 0$ , by Proposition 11, the vector field  $X$  does not depend on  $\tilde{\psi}_1$ ,  $\tilde{\gamma}_3$  and  $\tilde{\Gamma}_3$ . Therefore, from the invariance equation, it can be seen that

$$\rho(y; \varepsilon, 0) = \rho(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\Psi}_1; \varepsilon, 0).$$

To obtain the more precise information on this term stated in the lemma, we expand it in  $\varepsilon$ . For the first term, it is enough to point out that the vector field associated to the Hamiltonian  $H_0^{12} + L_2^{-1}H_1^{12} = H_0^{12} + \varepsilon H_1^{12}$  is integrable and is independent of  $\tilde{\gamma}_2$  and  $\tilde{\Psi}_1$ . Then, it can be seen as a 1 degree of freedom Hamiltonian depending on the parameter  $\tilde{\Gamma}_2$ . If one writes this vector field in Poincaré coordinates  $(\xi, \eta)$ , it has a saddle close to  $(\xi, \eta) = (0, 0)$  which is of the form

$$(\xi, \eta) = \varepsilon \rho_0(\tilde{\Gamma}_2).$$

This implies that  $\rho(y; \varepsilon, 0)$  is of the form

$$\rho(y; \varepsilon, 0) = \varepsilon \rho_0(\tilde{\Gamma}_2) + \varepsilon^2 \rho_1(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\Psi}_1; \varepsilon)$$

for some functions  $\rho_0, \rho_1$ . Now we perform a perturbative argument in  $\mu$ , and to this end we write the vector field  $X$  as

$$X(x, y; \varepsilon, \mu) = \hat{X}_0(x, y; \varepsilon) + \mu^6 \hat{X}_1(x, y; \varepsilon) + O(\mu^8 \varepsilon^{-4}).$$

Note that the  $O(\mu^8 \varepsilon^{-4})$  terms in the vector field expansion lead to terms of the same order in  $\rho$  and therefore, since

$$\mu^8 \varepsilon^{-4} = \frac{L_2^8}{L_3^6} \ll \hat{\varepsilon},$$

they are contained in  $\rho_3$ . Thus, we must analyse the order  $\mu^6$  terms.

For this reason, we write the parametrisation  $\rho$  as

$$\rho(y; \varepsilon, \mu) = \rho^0(y; \varepsilon) + \mu^6 \rho^1(y; \varepsilon) + O(\mu^8 \varepsilon^{-4}),$$

where  $\rho^0(y; \varepsilon) = \rho(y; \varepsilon, 0)$ . Then,  $\rho^1$  must satisfy

$$(73) \quad [D\widehat{X}_0^N(\rho^0) - D\rho_0 D\widehat{X}_0^T(\rho^0)]\rho^1 - D\rho^1 \widehat{X}_0^T(\rho_0) = D\rho^0 \widehat{X}_1^T(\rho^0) - \widehat{X}_1^N(\rho^0).$$

To compute  $\rho^1$  it is enough to invert the linear operator on the left hand side. However, it is not necessary to compute the inverse operator explicitly, because we already know that  $\rho^1$  exists, due to the regularity of  $\rho$  with respect to  $\mu$ . Therefore it remains only to analyse the dependence of  $\rho^1$  on each variable. To this end we expand  $\rho^1$  in powers of  $\varepsilon$  and we analyse each term. Note that, by Proposition 11, the vector field  $\widehat{X}_1$  can be expanded as

$$\widehat{X}_1 = \varepsilon^{-4} X_{10} + O(\varepsilon^{-3}),$$

where  $X_{10}$  is the vector field associated to the Hamiltonian  $H_0^{23}$  introduced in Proposition 11 (see also Lemma 15). Therefore, one can conclude

$$\rho^1 = \varepsilon^{-4} \rho^{10} + O(\varepsilon^{-3}).$$

In other words,

$$\rho(y; \varepsilon, \mu) = \rho^0(y; \varepsilon) + \mu^6 \varepsilon^{-4} \rho^{10} + O(\hat{\varepsilon}).$$

That is,  $\rho^{10}$  is just the function  $\rho_2$  in the statement of the lemma. To prove that it depends on the variables  $(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1)$ , we look for the equation it satisfies. To this end, it is enough to expand (73) in powers of  $\varepsilon$  to obtain

$$[D\widehat{X}_0^N(\rho^0) - D\rho_0 D\widehat{X}_0^T(\rho^0)]\rho^{10} - D\rho^{10} \widehat{X}_0^T(\rho_0) = D\rho^0 X_{10}^T(\rho^0) - X_{10}^N(\rho^0).$$

Then, it is enough to point out that all the terms in this equation are independent of  $\tilde{\gamma}_3$  and  $\tilde{\Gamma}_3$  and therefore  $\rho^{10}$  must be independent of these variables too.  $\square$

Denote by  $F_{\text{in}} = F_{\text{sec}}|_{\Lambda}$  the restriction of the secular Hamiltonian to the normally hyperbolic invariant manifold  $\Lambda$ .

**Corollary 22.** — *The expansion of the inner Hamiltonian  $F_{\text{in}}$  in powers of  $\epsilon = \frac{1}{L_2}$  and  $\mu = \frac{L_2}{L_3}$  satisfies the following properties:*

(1) *The lowest order term containing  $\tilde{\Gamma}_2$  is*

$$(74) \quad \epsilon^6 \tilde{H}_0^{12} = \epsilon^6 \tilde{\Gamma}_2^2,$$

*up to a nonzero multiplicative constant independent of  $\epsilon$  and  $\mu$ , which is the leading term in the expansion of  $F_{\text{in}}$ .*

- (2) *The lowest order terms containing the actions  $\tilde{\Psi}_1$  and  $\tilde{\Gamma}_3$  respectively, up to nonzero multiplicative constants independent of  $\epsilon$  and  $\mu$ , are  $\epsilon^7 \bar{H}_1^{12}$  and  $\epsilon^3 \mu^6 \tilde{H}_3$ , where*

$$(75) \quad \bar{H}_1^{12} = \left( 3 \frac{\tilde{\Gamma}_2^2}{L_1^2} - 1 \right) \tilde{\Psi}_1 + H'_1(\tilde{\Gamma}_2),$$

for some function  $H'_1$ , and where  $\tilde{H}_3$  is defined by (42).

- (3) *The lowest order term containing the angle  $\tilde{\gamma}_2$  is at least of order  $\epsilon^8$ .*  
 (4) *The lowest order terms containing the angles  $\tilde{\psi}_1$  and  $\tilde{\gamma}_3$  respectively, up to nonzero multiplicative constants, are  $\epsilon^2 \mu^6 H_0^{23}$  and  $\epsilon^2 \mu^8 H_3^{23}$ , where  $H_0^{23}$  is defined by (39) and  $H_3^{23}$  is defined by (50).*

*Proof.* — The Poincaré variables  $(\xi, \eta)$  first appear in  $H_0^{12}$ , where they always appear at least quadratically (see (63)). Squaring the formula (71) for the function  $\rho$ , using (72), and recalling that the order of  $H_0^{12}$  is  $\epsilon^6 = L_2^{-6}$ , we see that the function  $\rho$  satisfies the following properties:

- The first-order correction from  $\rho$  to the inner Hamiltonian that contains  $\tilde{\gamma}_2, \tilde{\Gamma}_2$  is of order  $O(\epsilon^8)$ ;
- The first-order correction from  $\rho$  to the inner Hamiltonian that contains  $\tilde{\Psi}_1$  is of order  $O(\epsilon^9)$ ;
- The first-order correction from  $\rho$  to the inner Hamiltonian that contains  $\tilde{\psi}_1$  is of order  $O(\epsilon^3 \mu^6)$ ;
- The first-order correction from  $\rho$  to the inner Hamiltonian that contains  $\tilde{\gamma}_3, \tilde{\Gamma}_3$  is of order  $O(\hat{\epsilon} \epsilon^7)$  where  $\hat{\epsilon} \lesssim \frac{L_2^9}{L_3^6}$ .

Comparing these orders with the first order of appearance of each variable

$$\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3$$

in the expansion of the Hamiltonian (see Proposition 11), we see that parts 1-4 of the corollary follow. The formulas (74) for  $\bar{H}_0^{12}$ , and (75) for  $\bar{H}_1^{12}$  come from restricting formula (33) for  $H_0^{12}$  and formula (34) for  $H_1^{12}$  to  $\tilde{\Lambda}_0 = \{\xi = \eta = 0\}$ .  $\square$

Due to Corollary 22 and Proposition 11, we have the expansion

$$F_{\text{in}} = F_{\text{in}}^{12} + F_{\text{in}}^{23}$$

with

$$(76) \quad \begin{cases} F_{\text{in}}^{12} = c_0^{12} + \frac{1}{L_2^6} \alpha_0^{12} \bar{H}_0^{12} + \frac{1}{L_2^7} \alpha_1^{12} \bar{H}_1^{12} + \dots \\ F_{\text{in}}^{23} = c_0^{23} + \frac{L_2^4}{L_3^6} \alpha_0^{23} \bar{H}_0^{23} + \frac{L_2^5}{L_3^7} \alpha_1^{23} \bar{H}_1^{23} + \dots \end{cases}$$

for some nontrivial constants  $\alpha_n^{jj+1}$ , where  $\tilde{H}_0^{12}$  is defined by (74), and  $\tilde{H}_1^{12}$  is defined by (75). The term  $F_{\text{in}}^{12}$  contains the  $L_3$ -independent terms and  $F_{\text{in}}^{23}$  contains the rest.

The following lemma allows us to find a coordinate transformation that straightens the symplectic form.

**Lemma 23.** — *There is a change of coordinates*

$$(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3) \mapsto (\gamma'_2, \Gamma'_2, \psi'_1, \Psi'_1, \gamma'_3, \Gamma'_3)$$

on  $\Lambda$  that is  $O(L_2^{-2})$  close to the identity satisfying

$$\begin{cases} \Gamma'_2 &= \tilde{\Gamma}_2 + \frac{1}{L_2^3} P'_2(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1) + \frac{\hat{\epsilon}}{L_2} N'_2(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3), \\ \Psi'_1 &= \tilde{\Psi}_1 + \frac{1}{L_2^3} P'_1(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1) + \frac{\hat{\epsilon}}{L_2} N'_1(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3), \\ \Gamma'_3 &= \tilde{\Gamma}_3 + \frac{\hat{\epsilon}}{L_2} N'_3(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3), \end{cases}$$

where  $\hat{\epsilon} \lesssim \frac{L_2^9}{L_3^6}$  comes from Lemma 21, such that

$$(77) \quad \Omega|_{\Lambda} = d\Gamma'_2 \wedge d\gamma'_2 + d\Psi'_1 \wedge d\psi'_1 + d\Gamma'_3 \wedge d\gamma'_3.$$

*Proof.* — Recall in Lemma 21 we obtained the expansion (71) for the function  $\rho$ . Denote by  $U$  the neighbourhood of the origin in  $\mathbf{R}^2$  in which the Poincaré variables  $(\xi, \eta)$  are valid, and by  $V$  the subset of  $\mathbf{R}^3$  to which the actions  $(\tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3)$  belong. Define the inclusion

$$P: \mathbf{T}^3 \times V \longrightarrow U \times (\mathbf{T}^3 \times V)$$

by  $(\xi, \eta, y) = P(y) = (\rho(y), y)$ . Then  $\Omega_1 = P^*\Omega$ , is the pullback of the canonical form in the tilde coordinates

We claim that  $\Omega_1$  is exact. Indeed, we have  $\Omega = d\lambda$  where

$$\lambda = \xi d\eta + \tilde{\Gamma}_2 d\tilde{\gamma}_2 + \tilde{\Psi}_1 d\tilde{\psi}_1 + \tilde{\Gamma}_3 d\tilde{\gamma}_3$$

is the Liouville 1-form. Let  $\lambda_1 = P^*\lambda$ . Then  $\Omega_1 = d\lambda_1$  since the pullback commutes with the differential, and so  $\Omega_1$  is exact.

Now, denote by  $\rho_\xi, \rho_\eta$  the  $\xi, \eta$  components of  $\rho$ , and by  $\rho_{j,\xi}, \rho_{j,\eta}$  the  $\xi, \eta$  components of  $\rho_j$  for each  $j = 0, 1, 2, 3$ . Recall  $\Omega_0$  is defined by (67). We have

$$\Omega_1 = P^*\Omega = d\rho_\xi \wedge d\rho_\eta + \Omega_0 = \Omega_0 + R_1 + R_2$$

where, since  $\rho_0$  only depends on  $\tilde{\Gamma}_2$ ,

$$R_1 = \frac{1}{L_2^3} \left[ d\rho_{0,\xi} \wedge \left( d\rho_{1,\eta} + \frac{L_2^8}{L_3^6} d\rho_{2,\eta} \right) + \left( d\rho_{1,\xi} + \frac{L_2^8}{L_3^6} d\rho_{2,\xi} \right) \wedge d\rho_{0,\eta} \right]$$



$$+ \frac{1}{L_2^4} \left( d\rho_{1,\xi} + \frac{L_2^8}{L_3^6} d\rho_{2,\xi} \right) \wedge \left( \frac{L_2^8}{L_3^6} d\rho_{2,\eta} \right)$$

$$R_2 = d\rho_\xi \wedge d\rho_\eta - R_1.$$

Then  $R_1$  is of order  $\frac{1}{L_2^3}$  and does not depend on  $\tilde{\gamma}_3, \tilde{\Gamma}_3$ , and  $R_2$  is of order  $\frac{\hat{\epsilon}}{L_2}$ . We first eliminate  $R_1$  by a change of coordinates which does not depend on  $\tilde{\gamma}_3, \tilde{\Gamma}_3$  and then we eliminate  $R_2$ .

Suppose there exists a coordinate transformation

$$h_0 : (\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3) \mapsto (\gamma_2^*, \Gamma_2^*, \psi_1^*, \Psi_1^*, \tilde{\gamma}_3, \tilde{\Gamma}_3)$$

so that

$$(78) \quad h_0^* \Omega' = \Omega_0 \quad \text{where } \Omega' = \Omega_0 + R_1.$$

Suppose moreover there is a (nonautonomous) vector field  $X_t$  such that  $h_0 = \phi_\epsilon$  where  $\phi_t$  is the time- $t$  map of  $X_t$ , and where  $\epsilon = L_2^{-3}$ . Differentiating (78) with respect to  $\epsilon$  and using Cartan's magic formula, we get

$$0 = \frac{d}{d\epsilon} [\phi_\epsilon^* \Omega'] = \phi_\epsilon^* \left[ \frac{d}{d\epsilon} \Omega' + \mathcal{L}_{X_t} \Omega' \right] = \phi_\epsilon^* \left[ \frac{d}{d\epsilon} \Omega' + i_{X_t} d\Omega' + di_{X_t} \Omega' \right]$$

where  $\mathcal{L}_{X_t}$  is the Lie derivative with respect to  $X_t$  and  $i_{X_t}$  is the contraction operator of  $X_t$ . Now,  $\Omega' = \Omega_0 + R_1$  is also exact, with  $\Omega' = d\lambda'$  for some 1-form  $\lambda'$ , by an argument analogous to the above proof of exactness of  $\Omega_1$ . Therefore

$$di_{X_t} \Omega' = -\frac{d}{d\epsilon} \Omega' = -\frac{d}{d\epsilon} d\lambda' = -d \left( \frac{d}{d\epsilon} \lambda' \right).$$

Then, to obtain a vector field  $X_t$  with this property, it is enough to look for  $X_t$  satisfying

$$i_{X_t} \Omega' = -\frac{d}{d\epsilon} \lambda'.$$

It is easy to check that a solution  $X_t$  to this equation exists; its flow exists at least for time  $\epsilon$ ; and its time- $\epsilon$  map, by construction, is the required coordinate transformation  $h_0$ .

In the new coordinates, the symplectic  $\Omega_1$  becomes of the form  $\hat{\Omega}_1 = h_0^* \Omega_1 = \Omega_0 + \hat{R}_2$ , where  $\hat{R}_2 = h_0^* R_2$ . Then, since  $\hat{R}_2$  is of order  $\frac{\hat{\epsilon}}{L_2}$ , we may repeat the procedure again to complete the proof.  $\square$

Since the coordinate transformation provided by Lemma 23 is  $O(L_2^{-2})$  close to the identity, the expansions (76) are unchanged (at least) up to the terms  $\bar{H}_2^{jj+1}$ . The

restriction of the secular Hamiltonian to  $\Lambda$  can be written as

$$(79) \quad F_{\text{in}} = F_{\text{in}}^{12} + F_{\text{in}}^{23} = c_0 + \sum_{n=0}^{\infty} \epsilon_n \alpha_n h_n$$

where

$$\epsilon_0 = \frac{1}{L_2^6} \gg \epsilon_1 = \frac{1}{L_2^7} \gg \epsilon_2 = \frac{L_2^4}{L_3^6} \gg \dots$$

and

$$(80) \quad h_0 = \bar{H}_0^{12}, \quad h_1 = \bar{H}_1^{12}, \quad h_2 = \bar{H}_0^{23}, \dots \quad \alpha_0 = \alpha_0^{12}, \quad \alpha_1 = \alpha_1^{12}, \quad \alpha_2 = \alpha_0^{23}, \dots$$

(see (32)). In the next lemma we perform an arbitrary number  $k$  of steps of symplectic averaging.

**Lemma 24.** — *Let  $(k_1, k_2) \in \mathbf{N}^2$ . There is a symplectic coordinate transformation*

$$\Phi : (\gamma'_2, \Gamma'_2, \psi'_1, \Psi'_1, \gamma'_3, \Gamma'_3) \mapsto (\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3)$$

on  $\Lambda$  that is  $O(\frac{L_2^{11}}{L_3^6})$  close to the identity such that the restriction  $\hat{F}$  of the secular Hamiltonian to  $\Lambda$ , defined in (79), becomes

$$\hat{F} = \hat{F}_0(\hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3; \epsilon, \mu) + \epsilon^{k_1} \mu^{k_2} \hat{F}_1(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3; \epsilon, \mu)$$

where  $\hat{F}_0 = \epsilon^6 c_0 \hat{\Gamma}_2^2 + \epsilon^7 \hat{h}_0(\hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3; \epsilon, \mu)$ , where  $\epsilon = \frac{1}{L_2}$ ,  $\mu = \frac{L_2}{L_3}$ , and where  $\hat{F}_j$  is uniformly bounded in  $\epsilon, \mu$  for  $j = 0, 1$ . Moreover, the transformations of the actions satisfy

$$(81) \quad \hat{\Gamma}_2 = \Gamma'_2 + O\left(\frac{L_2^9}{L_3^6}\right),$$

$$(82) \quad \hat{\Psi}_1 = \Psi'_1 - \frac{L_2^{11}}{L_3^6} \frac{\alpha_2}{\alpha_1} \frac{\cos(2\psi'_1)}{2(3\frac{(\Gamma'_2)^2}{L_1^2} - 1)} + \dots$$

$$(83) \quad \begin{aligned} \hat{\Gamma}_3 = \Gamma'_3 + \frac{L_2^{13}}{L_3^8} \tilde{\alpha} \left( 3 \frac{(\Gamma'_2)^2}{L_1^2} - 1 \right)^{-1} & \left[ \frac{1}{3} v_0 \cos(\gamma'_3 + 3\psi'_1) + v_1 \cos(\gamma'_3 + \psi'_1) \right. \\ & \left. - v_2 \cos(\gamma'_3 - \psi'_1) - \frac{1}{3} v_3 \cos(\gamma'_3 - 3\psi'_1) \right] + \dots \end{aligned}$$

where  $\tilde{\alpha}$  is a nontrivial constant, and the constants  $v_j$  are given by (51), (52).

*Proof.* — The fact that such a coordinate transformation exists is a standard result from averaging theory and the fact that the frequencies  $(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)$  are at different time

scales. We compute the first order. There is a Hamiltonian function  $K$  such that the corresponding flow  $\phi_K^t$  with respect to the standard symplectic form  $\Omega_1$  (given by (77)) satisfies  $\phi_K^1 = \Phi$ . We can write

$$K = \sum_{n=2}^k \hat{\epsilon}_n K_n$$

where the sum starts at  $n = 2$  because  $h_0, h_1$  do not depend on the angles  $\gamma'_2, \psi'_1, \gamma'_3$ . In order to obtain a first order approximation of  $\Phi$  we determine  $\hat{\epsilon}_2, K_2$ . The purpose of  $K_2$  is to eliminate the angle  $\psi'_1$  from  $h_2$  by adjusting the action  $\Psi'_1$ , which first appears in  $h_1$ . Set

$$\hat{\epsilon}_2 = \frac{\epsilon_2}{\epsilon_1} = \frac{L_2^{11}}{L_3^6}.$$

Suppose  $K_2$  does not depend on  $\gamma'_2$ . Then, from (75) and (80),

$$\begin{aligned} (84) \quad F_{\text{in}} \circ \phi_{K_2}^{-\hat{\epsilon}_2} &= c_0 + \epsilon_0 \alpha_0 h_0 \circ \phi_{K_2}^{-\hat{\epsilon}_2} + \epsilon_1 \alpha_1 h_1 \circ \phi_{K_2}^{-\hat{\epsilon}_2} + \epsilon_2 \alpha_2 h_2 \circ \phi_{K_2}^{-\hat{\epsilon}_2} + \dots \\ &= c_0 + \epsilon_0 \alpha_0 \hat{\Gamma}_2^2 + \epsilon_1 \alpha_1 h_1 + \epsilon_2 (\alpha_2 h_2 - \alpha_1 \{h_1, K_2\}) + \dots \end{aligned}$$

where  $\{\cdot, \cdot\}$  is the Poisson bracket. We therefore require that  $K_2$  solves the equation

$$(85) \quad \{h_1, K_2\} = \frac{\alpha_2}{\alpha_1} (h_2 - \langle h_2 \rangle) = \frac{\alpha_2}{\alpha_1} \left( \cos^2 \psi'_1 - \frac{1}{2} \right).$$

Let

$$K_2 = \frac{\alpha_2}{\alpha_1} \frac{\sin 2\psi'_1}{4(3\frac{(\Gamma'_2)^2}{L_1^2} - 1)}.$$

Since  $K_2$  does not depend on  $\gamma'_2$ , equation (84) holds; moreover  $K_2$  solves (85). From Hamilton's equations of motion we have

$$\hat{\Psi}_1 = \Psi'_1 - \hat{\epsilon}_2 \frac{\partial K_2}{\partial \psi'_1} + \dots$$

which implies (82).

Since  $F_{\text{oct}}^{12}$  has a factor of  $e_1$  (see (27)), it is zero on  $\tilde{\Lambda}_0$ , and thus  $O(L_2^{-1})$  on  $\Lambda$ . It follows that the angle  $\gamma'_2$  does not appear in the inner secular Hamiltonian until the term of order  $O(L_2^3 L_3^{-6})$ . Since the action  $\Gamma'_2$  appears in the term of order  $O(L_2^{-6})$ , we get the approximation (81).

Let us now prove formula (83). By part 6 of Proposition 11, the first-order term where we see  $\tilde{\gamma}_3$  is  $\frac{L_2^6}{L_3^8} \alpha_3^{23} H_3^{23}$  where  $H_3^{23}$  is defined by (50). Notice that

$$\langle H_3^{23} \rangle_{\tilde{\psi}_1} = \frac{1}{2\pi} \int_{\mathbf{T}} H_3^{23} d\tilde{\psi}_1 = 0.$$

So this term is completely eliminated in the averaging process if one considers as integrable first order  $L_2^{-6} \tilde{H}_0^{12} + L_2^{-7} \tilde{H}_1^{12}$  since then  $\tilde{\psi}_1$  has a non-trivial frequency (see (75)).

By similar logic to that used in the definition of  $\hat{\epsilon}_2$  and  $K_2$ , we define  $\hat{\epsilon} = \frac{L_2^{13}}{L_3^8}$ , and we search for a Hamiltonian function  $\hat{K} = \tilde{\alpha} \tilde{K}$ , with  $\tilde{\alpha} = \frac{\alpha_3^{23}}{\alpha_1^{12}}$ , such that  $H_3^{23} - \{H_1^{12}, \tilde{K}\} = 0$ . Notice that this equation is satisfied by the Hamiltonian function

$$\tilde{K} = - \left( 3 \frac{(\Gamma'_2)^2}{L_1^2} - 1 \right)^{-1} \left[ \frac{1}{3} \nu_0 \sin(\gamma'_3 + 3\psi'_1) + \nu_1 \sin(\gamma'_3 + \psi'_1) \right. \\ \left. - \nu_2 \sin(\gamma'_3 - \psi'_1) - \frac{1}{3} \nu_3 \sin(\gamma'_3 - 3\psi'_1) \right].$$

Differentiating this expression with respect to  $\gamma'_3$  and using the formula

$$\hat{\Gamma}_3 = \Gamma'_3 - \hat{\epsilon} \frac{\partial \hat{K}}{\partial \gamma'_3} + \dots = \Gamma'_3 - \hat{\epsilon} \tilde{\alpha} \frac{\partial \tilde{K}}{\partial \gamma'_3} + \dots$$

yields (83). □

This completes the proof of Theorem 18. Now it only remains to compute the second derivatives of the averaged Hamiltonian in the cylinder. This will be used to show that certain Poincaré map associated to this Hamiltonian has nonzero torsion (see Section 7).

**5.2. The Hessian of the averaged inner Hamiltonian.** — Now, consider the restriction of the secular Hamiltonian to  $\Lambda$  given by Lemma 24. We have denoted this Hamiltonian by  $\hat{F}$  and it is expressed as a function of the “hat” coordinates  $(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3)$ .

The following lemma gives first-order expressions for the partial derivatives of  $\hat{F}$  of orders 1 and 2 with respect to the actions  $\hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3$ . Note that the Hamiltonian  $\hat{F}$  given in Lemma 24 is just given by  $\hat{F}_0$  plus very small terms. Therefore, we only need to analyse the derivatives of  $\hat{F}_0$ .

The Hessian matrix of  $\hat{F}$  will be useful to determine the torsion of a certain Poincaré map of the inner dynamics. Indeed, the highest order term of the torsion will depend only on those derivatives.

**Lemma 25.** — *The first and second-order partial derivatives of  $\hat{F}$  with respect to the actions  $\hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3$  are*

$$\frac{\partial \hat{F}}{\partial \hat{\Gamma}_2} = \epsilon^6 C_{12} \frac{6 \hat{\Gamma}_2}{L_1^2 \delta_1^3} + \dots, \quad \frac{\partial \hat{F}}{\partial \hat{\Psi}_1} = 3\epsilon^7 C_{12} \frac{L_1^2 - 3 \hat{\Gamma}_2^2}{L_1^2 \delta_1^4} + \dots$$

$$\begin{aligned}
 \frac{\partial \hat{F}}{\partial \hat{\Gamma}_3} &= \epsilon^3 \mu^6 C_{23} \frac{(20 - 2\delta_1^2)\delta_3}{\delta_1^2 \delta_2^3} + \dots, & \frac{\partial^2 \hat{F}}{\partial \hat{\Gamma}_2^2} &= \epsilon^6 C_{12} \frac{6}{L_1^2 \delta_1^3} + \dots \\
 \frac{\partial^2 \hat{F}}{\partial \hat{\Psi}_1^2} &= 12\epsilon^8 C_{12} \left( \frac{3\hat{\Gamma}_2^2 - L_1^2}{L_1^2 \delta_1^5} \right) + \dots, \\
 \frac{\partial^2 \hat{F}}{\partial \hat{\Gamma}_3^2} &= \epsilon^4 \mu^6 C_{23} \frac{(20 - 12\delta_1^2)}{\delta_1^2 \delta_2^3} + \dots \\
 \frac{\partial^2 \hat{F}}{\partial \hat{\Gamma}_2 \partial \hat{\Psi}_1} &= -\epsilon^7 C_{12} \frac{18\hat{\Gamma}_2}{L_1^2 \delta_1^4} + \dots, & \frac{\partial^2 \hat{F}}{\partial \hat{\Gamma}_2 \partial \hat{\Gamma}_3} &= \epsilon^4 \mu^6 C_{23} \frac{24\delta_3}{\delta_1 \delta_2^3} + \dots \\
 \frac{\partial^2 \hat{F}}{\partial \hat{\Psi}_1 \partial \hat{\Gamma}_3} &= -\epsilon^4 \mu^6 C_{23} \frac{40\delta_3}{\delta_1^3 \delta_2^3} + \dots
 \end{aligned}$$

where  $\epsilon = \frac{1}{L_2}$  and  $\mu = \frac{L_2}{L_3}$ , and where  $C_{12}, C_{23}$  are nonzero constants independent of  $L_2$  and  $L_3$  coming from  $F_{\text{quad}}^{12}, F_{\text{quad}}^{23}$  respectively.

To prove this lemma we proceed in several steps. First we analyse the averaged Hamiltonian

$$\langle F_{\text{in}} \rangle_{(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)}$$

where  $F_{\text{in}}$  is the Hamiltonian given by Corollary 22. In particular, we analyse the linear and quadratic terms with respect to the actions of the averaged Hamiltonian.

Then, we prove that these terms remain unchanged (at first order) when one applies the change of coordinates given by Lemmas 23 and 24. This implies that these terms give the first order of the first and second derivatives given in Lemma 25.

*Proof of Lemma 25.* — The formulas for the first derivatives are straightforward using the expansion of the Hamiltonian  $F_{\text{sec}}$  given by Proposition 11 and the expansion of the normally hyperbolic invariant cylinder given by Lemma 21. Indeed, the expansion of  $F_{\text{sec}}$  given in Proposition 11 gives explicitly the first order of the linear terms in the actions  $\tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3$  of the Hamiltonian. One can easily compute the pullback of these terms in the cylinder and then show that the changes of coordinates given by Lemmas 23 and 24 do not alter these first orders.

Now we analyse the second derivatives. Note that Proposition 11 does not give terms which are quadratic in the actions and therefore it does not provide first orders of the second derivatives. To this end we have to do a deeper analysis of the Hamiltonian  $\langle F_{\text{in}} \rangle_{(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)}$ .

To analyse it, we first analyse the Hamiltonian  $\langle F_{\text{sec}} \rangle_{(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)}$ , where  $F_{\text{sec}}$  is the Hamiltonian analysed in Proposition 11, restricted to the normally hyperbolic invariant cylinder given by Lemma 21. We use the expansions of the Hamiltonian given by

the Proposition 11 and the expansion of the parameterization of the cylinder given in Lemma 21.

The first orders in the actions in the Hamiltonian  $\langle F_{\text{in}} \rangle_{(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)}$  come from  $F_{\text{quad}}^{12}$  and  $F_{\text{quad}}^{23}$ . In order to compute the derivatives, we go back to the original formulas (26) for  $F_{\text{quad}}^{12}$  and (43) for  $F_{\text{quad}}^{23}$ . We average out the angles  $\tilde{\gamma}_2, \tilde{\psi}_1$  from these Hamiltonians (recall  $F_{\text{quad}}^{12}$  and  $F_{\text{quad}}^{23}$  do not depend on  $\gamma_3$ ), and compute the derivatives with respect to the actions. We compute the expansions of these derivatives in powers of  $\epsilon = \frac{1}{L_2}$  and  $\mu = \frac{L_2}{L_3}$ .

First we compute the derivatives of  $F_{\text{quad}}^{12}$  restricted to  $\tilde{\Lambda}_0$ . Recall that  $\Gamma_1 = L_1$  on  $\tilde{\Lambda}_0$ ; therefore  $e_1|_{\tilde{\Lambda}_0} = 0$  and so

$$K_1 = F_{\text{quad}}^{12}|_{\tilde{\Lambda}_0} = \frac{a_1^2}{8a_2^3(1 - e_2^2)^{\frac{3}{2}}} (3 \cos^2 i_{12} - 1)$$

where  $e_2$  is defined by (25), and  $\cos i_{12}$  is defined by (28). Therefore, using again the fact that  $\Gamma_1|_{\tilde{\Lambda}_0} = L_1$ , we see that the first and second partial derivatives of  $K_1$  are

$$(86) \quad \begin{cases} \frac{\partial^2 K_1}{\partial \Gamma_2^2} &= \frac{a_1^2}{8a_2^3} \left( \frac{45L_2^3 \Psi_1^4}{2\Gamma_2^7 L_1^2} - \frac{18L_2^3 \Psi_1^2}{\Gamma_2^5 L_1^2} - \frac{45L_2^3 \Psi_1^2}{\Gamma_2^7} + \frac{45L_1^2 L_2^3}{2\Gamma_2^7} + \frac{3L_2^3}{2\Gamma_2^3 L_1^2} + \frac{6L_2^3}{\Gamma_2^5} \right) \\ \frac{\partial^2 K_1}{\partial \Psi_1^2} &= \frac{a_1^2}{8a_2^3} \left( \frac{9L_2^3 \Psi_1^2}{\Gamma_2^5 L_1^2} - \frac{3L_2^3}{\Gamma_2^3 L_1^2} - \frac{3L_2^3}{\Gamma_2^5} \right) \\ \frac{\partial^2 K_1}{\partial \Gamma_2 \partial \Psi_1} &= \frac{a_1^2}{8a_2^3} \left( -\frac{15L_2^3 \Psi_1^3}{\Gamma_2^6 L_1^2} + \frac{9L_2^3 \Psi_1}{\Gamma_2^4 L_1^2} + \frac{15L_2^3 \Psi_1}{\Gamma_2^6} \right) \\ \frac{\partial K_1}{\partial \Gamma_2} &= \frac{a_1^2}{8a_2^3} \left( -\frac{15L_2^3 \Psi_1^4}{4\Gamma_2^6 L_1^2} + \frac{9L_2^3 \Psi_1^2}{2\Gamma_2^4 L_1^2} + \frac{15L_2^3 \Psi_1^2}{2\Gamma_2^6} - \frac{15L_1^2 L_2^3}{4\Gamma_2^6} - \frac{3L_2^3}{4\Gamma_2^2 L_1^2} - \frac{3L_2^3}{2\Gamma_2^4} \right) \\ \frac{\partial K_1}{\partial \Psi_1} &= \frac{a_1^2}{8a_2^3} \left( \frac{3L_2^3 \Psi_1^3}{\Gamma_2^5 L_1^2} - \frac{3L_2^3 \Psi_1}{\Gamma_2^3 L_1^2} - \frac{3L_2^3 \Psi_1}{\Gamma_2^5} \right). \end{cases}$$

Due to (29) we have

$$(87) \quad \frac{\partial K_1}{\partial \tilde{\Gamma}_2} = -\frac{\partial K_1}{\partial \Gamma_2}, \quad \frac{\partial K_1}{\partial \tilde{\Psi}_1} = \frac{\partial K_1}{\partial \Psi_1} + \frac{\partial K_1}{\partial \Gamma_2}, \quad \frac{\partial^2 K_1}{\partial \tilde{\Gamma}_2^2} = \frac{\partial^2 K_1}{\partial \Gamma_2^2},$$

and

$$(88) \quad \frac{\partial^2 K_1}{\partial \tilde{\Psi}_1^2} = \frac{\partial^2 K_1}{\partial \Psi_1^2} + 2 \frac{\partial^2 K_1}{\partial \Gamma_2 \partial \Psi_1} + \frac{\partial^2 K_1}{\partial \Gamma_2^2}, \quad \frac{\partial^2 K_1}{\partial \tilde{\Gamma}_2 \partial \tilde{\Psi}_1} = -\frac{\partial^2 K_1}{\partial \Gamma_2 \partial \Psi_1} - \frac{\partial^2 K_1}{\partial \Gamma_2^2}.$$

Combining (86), (87), and (88), effecting the coordinate transformation (29), and expanding in powers of  $\epsilon = \frac{1}{L_2}$ , we see that

$$(89) \quad \begin{cases} \frac{\partial^2 K_1}{\partial \tilde{\Gamma}_2^2} &= \epsilon^6 C_{12} \frac{6}{L_1^2 \delta_1^3} + O(\epsilon^7) \\ \frac{\partial^2 K_1}{\partial \tilde{\Psi}_1^2} &= 12\epsilon^8 C_{12} \left( \frac{3\tilde{\Gamma}_2^2 - L_1^2}{L_1^2 \delta_1^5} \right) + O(\epsilon^9) \\ \frac{\partial^2 K_1}{\partial \tilde{\Gamma}_2 \partial \tilde{\Psi}_1} &= -\epsilon^7 C_{12} \frac{18\tilde{\Gamma}_2}{L_1^2 \delta_1^4} + O(\epsilon^8) \\ \frac{\partial K_1}{\partial \tilde{\Gamma}_2} &= \epsilon^6 C_{12} \frac{6\tilde{\Gamma}_2}{L_1^2 \delta_1^3} + O(\epsilon^7) \\ \frac{\partial K_1}{\partial \tilde{\Psi}_1} &= 3\epsilon^7 C_{12} \frac{L_1^2 - 3\tilde{\Gamma}_2^2}{L_1^2 \delta_1^4} + O(\epsilon^8) \end{cases}$$

where the coefficient  $C_{12}$  is a nonzero constant depending on  $L_1$  and  $\mu_j, M_j$  for  $j = 1, 2$  (but independent of  $L_2$  and  $L_3$ ).

Now, the restriction to  $\tilde{\Lambda}_0$  of  $F_{\text{quad}}^{23}$  depends on  $\gamma_2, \psi_1$ . Since we have shown in Lemma 24 that a  $C^r$  near-to-the-identity coordinate transformation transforms the Hamiltonian to its average over  $\tilde{\gamma}_2, \tilde{\psi}_2$ , we instead consider

$$\begin{aligned} K_2 &= \left[ \int_{\mathbf{T}^4} P_2(\cos \zeta_2) \frac{\|q_2\|^2}{\|q_3\|^3} d\ell_2 d\ell_3 d\tilde{\gamma}_2 d\tilde{\psi}_1 \right]_{\Gamma_1=L_1} \\ &= \left[ \int_{\mathbf{T}^2} \langle P_2(\cos \zeta_2) \rangle_{\tilde{\gamma}_2, \tilde{\psi}_1} \frac{\|q_2\|^2}{\|q_3\|^3} d\ell_2 d\ell_3 \right]_{\Gamma_1=L_1} \end{aligned}$$

where we have taken formula (43) for  $F_{\text{quad}}^{23}$ , restricted attention to  $\tilde{\Lambda}_0$  by setting  $\Gamma_1 = L_1$  (and noticing that  $F_{\text{quad}}^{23}$  does not depend on  $\gamma_1$ ), averaged over the variables  $\tilde{\gamma}_2, \tilde{\psi}_1$ , and used the fact that  $\|q_2\|, \|q_3\|$  do not depend on  $\tilde{\gamma}_2, \tilde{\psi}_1$ . Recall from the proof of Lemma 15 and the coordinate transformation (29) that

$$\cos \zeta_2 = (\mathcal{R}_1(\tilde{i}_2) \mathcal{R}_3(\tilde{\psi}_1 + \tilde{\gamma}_2) I_3 \mathcal{R}_1(i_2) \bar{Q}_2) \cdot (I_3 \mathcal{R}_1(i_3) \bar{Q}_3)$$

where  $\mathcal{R}_1, \mathcal{R}_3$  is the rotation around the  $x, z$ -axis respectively,  $I_3 = \mathcal{R}_3(\pi)$ , the angles  $\tilde{i}_2, i_2, i_3$  are defined by (44) and (45), and  $\bar{Q}_j = (\cos(v_j - \tilde{\gamma}_j), \sin(v_j - \tilde{\gamma}_j), 0)$  with  $v_j$  denoting the true anomaly of the  $j^{\text{th}}$  fictitious body. We can thus compute  $\langle P_2(\cos \zeta_2) \rangle_{\tilde{\gamma}_2, \tilde{\psi}_1} = \int_{\mathbf{T}_2} P_2(\cos \zeta_2) d\tilde{\gamma}_2 d\tilde{\psi}_1$ , and use the resulting expression to obtain

$$\begin{aligned} K_2 &= \frac{-4\pi^2 a_2^2}{32 a_3^3 (1 - e_3^2)^{\frac{3}{2}}} (3e_2^2 + 2) (3 \cos^2 i_2 - 1) \\ &\quad \times [6 \cos \tilde{i}_2 \sqrt{1 - \cos^2 \tilde{i}_2} \cos i_3 \sqrt{1 - \cos^2 i_3} \\ &\quad - 6 \cos^2 \tilde{i}_2 \cos^2 i_3 + 3 \cos^2 i_3 + 3 \cos^2 \tilde{i}_2 - 2]. \end{aligned}$$

TABLE 1. — The order at which the relevant products of the actions appear in the inner Hamiltonian, as a result of appearing explicitly in the averaged secular Hamiltonian  $\langle F_{\text{sec}} \rangle_{(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)}$ , and as a result of restricting the Hamiltonian  $\langle F_{\text{sec}} \rangle_{(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)}$  to the manifold  $\Lambda$  using the function  $\rho$

	$\tilde{\Gamma}_2 \tilde{\Psi}_1$	$\tilde{\Gamma}_2 \tilde{\Gamma}_3$	$\tilde{\Psi}_1 \tilde{\Gamma}_3$	$\tilde{\Gamma}_2^2$	$\tilde{\Psi}_1^2$	$\tilde{\Gamma}_3^2$
From $\langle F_{\text{sec}} \rangle_{(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)}$	$\epsilon^7$	$\epsilon^4 \mu^6$	$\epsilon^4 \mu^6$	$\epsilon^6$	$\epsilon^8$	$\epsilon^4 \mu^6$
From $\rho$	$\epsilon^9$	$\hat{\epsilon} \epsilon^7$	$\hat{\epsilon} \epsilon^7$	$\epsilon^8$	$\epsilon^9$	$\hat{\epsilon} \epsilon^7$

To compute this integral we have used again the technique described in Appendix C of [37] (used above in Lemmas 9, 15, and 16). We now effect the coordinate transformation (29), and expand  $K_2$  in powers of  $\mu = \frac{L_2}{L_3}$  and  $\epsilon = \frac{1}{L_2}$  to obtain

$$(90) \quad K_2 = \epsilon^2 \mu^6 C_{23} [(\bar{K}_{0,0} + \epsilon \bar{K}_{0,1} + \epsilon^2 \bar{K}_{0,2} + \dots) + \mu(\bar{K}_{1,0} + \epsilon \bar{K}_{1,1} + \epsilon^2 \bar{K}_{1,2} + \dots) + \mu^2(\bar{K}_{2,0} + \epsilon \bar{K}_{2,1} + \epsilon^2 \bar{K}_{2,2} + \dots)]$$

where the formulas for  $\bar{K}_{i,j}$  can be found in Appendix E. Therefore we can compute

$$(91) \quad \begin{cases} \frac{\partial^2 K_2}{\partial \tilde{\Gamma}_3^2} &= \epsilon^4 \mu^6 C_{23} \frac{(20-12\delta_1^2)}{\delta_1^2 \delta_2^3} + O(\epsilon^5 \mu^6, \epsilon^4 \mu^7) \\ \frac{\partial^2 K_2}{\partial \tilde{\Gamma}_3 \partial \tilde{\Psi}_1} &= -\epsilon^4 \mu^6 C_{23} \frac{40\delta_3}{\delta_1^3 \delta_2^3} + O(\epsilon^5 \mu^6, \epsilon^4 \mu^7) \\ \frac{\partial^2 K_2}{\partial \tilde{\Gamma}_3 \partial \tilde{\Gamma}_2} &= \epsilon^4 \mu^6 C_{23} \frac{24\delta_3}{\delta_1 \delta_2^3} + O(\epsilon^5 \mu^6, \epsilon^4 \mu^7) \\ \frac{\partial K_2}{\partial \tilde{\Gamma}_3} &= \epsilon^3 \mu^6 C_{23} \frac{(20-12\delta_1^2)\delta_3}{\delta_1^2 \delta_2^3} + O(\epsilon^4 \mu^6, \epsilon^3 \mu^7). \end{cases}$$

Note that these second derivatives are the derivatives of the pullback of the Hamiltonian to the unperturbed cylinder  $\Lambda_0$ . To ensure that the first order of the derivatives remain unchanged if one considers the pullback to the perturbed normally hyperbolic invariant cylinder given by Lemma 21 it is enough to use the expansion of its parameterization given by this lemma.

The sizes of each of the terms in the Hamiltonian and the cylinder parameterization are gathered in Table 1. In each case the dominant term is the term coming from the averaged secular Hamiltonian; this can be checked using the assumption (11) and the inequality  $\hat{\epsilon} \ll \frac{L_2^9}{L_3^6}$ .

To complete the proof of Lemma 25 it is enough to ensure that the change of coordinates given in (68) does not alter the first order of the second derivatives given in (89) and (91). This change of coordinates is the composition of the changes given by Lemmas 23 and 24. Then, it is enough to use the estimates given in these lemmas on how close to the identity are these changes.  $\square$



## 6. The outer dynamics

Recall from Section 5 that Theorem 18 ensures the existence of the normally hyperbolic (weakly) invariant cylinder  $\Lambda$  (see (70) and (65)) with stable and unstable invariant manifolds. The purpose of this section is to show that these invariant manifolds intersect along two homoclinic channels. Such channels allow us to establish the existence of two scattering maps. This is done in Section 6.1 where we provide asymptotic formulas for those maps. These formulas are given by an associated Poincaré-Melnikov potential. Its computation is contained in Section 6.2.

**6.1. The scattering map.** — Recall from Section 5 that Theorem 18 ensures the existence of the normally hyperbolic (weakly) invariant cylinder  $\Lambda$  (see (70) and (65)) with stable and unstable invariant manifolds. The purpose of this section is to show that these invariant manifolds intersect along two homoclinic channels. Such channels allow us to establish the existence of two scattering maps  $S_{\pm} : \Lambda \rightarrow \Lambda'$ , for some  $\Lambda'$  such that  $\Lambda \subset \Lambda' \subset \mathbf{T}^3 \times \mathbf{R}^3$  (see Appendix B for the definition). Then, we derive a first-order approximation of the changes in the actions  $\hat{\Psi}_1, \hat{\Gamma}_3$  in the ‘hat’ coordinates defined by Theorem 18. The following theorem is the main result of this section.

**Theorem 26.** — *The stable and unstable invariant manifolds of the normally hyperbolic invariant cylinder  $\Lambda$  introduced in Theorem 18 intersect transversally along (at least) two homoclinic channels. Associated to these channels, there exist two scattering maps  $S_{\pm} : \Lambda \rightarrow \Lambda'$  such that*

$$S_{\pm} : (\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3) \mapsto (\hat{\gamma}_2^*, \hat{\Gamma}_2^*, \hat{\psi}_1^*, \hat{\Psi}_1^*, \hat{\gamma}_3^*, \hat{\Gamma}_3^*)$$

with

$$\hat{\Psi}_1^* = \hat{\Psi}_1 + \frac{L_2^9}{L_3^6} \mathcal{S}_1^{\pm}(\hat{\psi}_1, \hat{\Gamma}_2) + \dots, \quad \hat{\Gamma}_3^* = \hat{\Gamma}_3 + \frac{L_2^{11}}{L_3^8} \mathcal{S}_3^{\pm}(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) + \dots$$

where

$$(92) \quad \begin{aligned} \mathcal{S}_1^{\pm}(\hat{\psi}_1, \hat{\Gamma}_2) = & \mp \alpha_2^{23} \kappa \left( \frac{\pi \hat{\Gamma}_2}{A_2 L_1^2} \right) c_2 \cos \hat{\psi}_1 \\ & + \frac{\alpha_0^{23}}{\alpha_1^{12}} \beta_2 \frac{L_1}{6} \sqrt{\frac{3}{2}} \hat{\Gamma}_2 \sqrt{1 - \frac{5}{3} \frac{\hat{\Gamma}_2^2}{L_1^2} \frac{\sin 2\hat{\psi}_1}{\frac{\hat{\Gamma}_2^2}{L_1^2} - 1}}, \end{aligned}$$

$$(93) \quad \begin{aligned} \mathcal{S}_3^{\pm}(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) \\ = & \mp \alpha_5^{23} \kappa \left( \frac{\pi \hat{\Gamma}_2}{A_2 L_1^2} \right) \left[ 50 \delta_1 \delta_3 \sin \hat{\gamma}_3 \cos \hat{\psi}_1 \sin \hat{\psi}_1 \right. \end{aligned}$$

$$\begin{aligned}
 & - \frac{70 \delta_3}{\delta_1} \sin \hat{\gamma}_3 \cos \hat{\psi}_1 \sin \hat{\psi}_1 \\
 & + \frac{105 \delta_3^2}{\delta_1^2} \cos \hat{\gamma}_3 \cos^2 \hat{\psi}_1 - 75 \delta_3^2 \cos \hat{\gamma}_3 \cos^2 \hat{\psi}_1 \\
 & + 25 \delta_1^2 \cos \hat{\gamma}_3 \cos^2 \hat{\psi}_1 - 35 \cos \hat{\gamma}_3 \cos^2 \hat{\psi}_1 \\
 & - \frac{105 \delta_3^2}{\delta_1^2} \cos \hat{\gamma}_3 + 60 \delta_3^2 \cos \hat{\gamma}_3 - 17 \delta_1^2 \cos \hat{\gamma}_3 + 28 \cos \hat{\gamma}_3 \Big] \\
 & + \frac{L_1}{6} \sqrt{\frac{3}{2}} \beta_2 \hat{\Gamma}_2 \sqrt{1 - \frac{5}{3} \frac{\hat{\Gamma}_2^2}{L_1^2}} \tilde{\alpha} \left( 3 \frac{\hat{\Gamma}_2^2}{L_1^2} - 1 \right)^{-1} \left[ v_0 \cos(\hat{\gamma}_3 + 3 \hat{\psi}_1) \right. \\
 & \left. + v_1 \cos(\hat{\gamma}_3 + \hat{\psi}_1) + v_2 \cos(\hat{\gamma}_3 - \hat{\psi}_1) + v_3 \cos(\hat{\gamma}_3 - 3 \hat{\psi}_1) \right].
 \end{aligned}$$

where the function  $\kappa$  is defined by

$$(94) \quad \kappa(x) = \sqrt{\frac{2}{3}} \frac{L_1^2}{\chi} \left[ 1 - \frac{x}{\sinh x} \right],$$

the constant  $c_2$  is defined in (41), and  $\chi$ ,  $A_2$  are the constants defined in (57), the constants  $\alpha_{ij}^k$  come from the expansion of the secular Hamiltonian, and are defined in Section 3, the constant  $\beta_2$  is defined in Appendix D, the nontrivial constant  $\tilde{\alpha}$  comes from Lemma 24, and the constants  $v_j$  are defined by (51) and (52).

The proof of Theorem 26 is done in several steps. The first step is to prove that the invariant manifolds of the cylinder intersect. This analysis is performed in the ‘tilde’ coordinates introduced in (29). The analysis of the transversality of these invariant manifolds allows us to derive first-order expressions for the associated scattering maps in ‘tilde’ coordinates. Then, the second step is to express the scattering maps in the ‘hat’ variables provided by Theorem 18.

The first step is achieved by means of a suitable Poincaré-Melnikov Theory. We follow the approach by Delshams, de la Llave and Seara in [31] which deals with the transverse intersection of the invariant manifolds of normally hyperbolic invariant manifolds in nearly integrable regimes.

Recall that the Hamiltonian  $F_{\text{quad}}^{12}$  analysed in Lemma 13 is integrable. We show that the higher order terms in the Hamiltonian make the invariant manifolds of the cylinder split.

Recall that in Section 4 we established the existence of two hyperbolic periodic orbits  $Z_{\min}^0, Z_{\max}^0$  for the Hamiltonian  $H_0^{12}$ ; moreover we found that the stable and unstable manifolds of these saddles coincide along a heteroclinic trajectory  $Z^0$ . Furthermore, in Poincaré variables (61) the two saddles are reduced to a single hyperbolic periodic orbit, which we denote by  $Z_*^0$ , and this hyperbolic periodic orbit possesses a homoclinic

connection, which for convenience we continue to denote by  $Z^0$ . Write  $\tilde{F}_{\text{quad}}^{12} = L_2^6 F_{\text{quad}}^{12}$ . Then  $\tilde{F}_{\text{quad}}^{12}$  possesses a hyperbolic periodic orbit  $Z_*^{\text{quad}}$  that is  $O(L_2^{-1})$  close to  $Z_*^0$ ; moreover there is a homoclinic orbit  $Z^{\text{quad}}$  to  $Z_*^{\text{quad}}$  that is  $O(L_2^{-1})$  close to  $Z^0$ . Since  $\tilde{F}_{\text{quad}}^{12}$  is integrable, the homoclinic trajectory  $Z^{\text{quad}}$  corresponds to a non-transverse homoclinic manifold intersection of the stable and unstable manifolds of  $Z_*^{\text{quad}}$ . Note that, as happened in Section 4, such objects are referred to the Hamiltonian  $\tilde{F}_{\text{quad}}^{12}$  seen as a 1 degree of freedom Hamiltonian in the  $(\tilde{\gamma}_1, \tilde{\Gamma}_1)$  variables (or equivalently in  $(\eta, \xi)$  variables). One can consider the same objects in the full phase space. Then, the hyperbolic periodic orbit becomes the normally hyperbolic invariant cylinder and the homoclinic becomes a 7 dimensional homoclinic manifold. Note that, by the particular form of  $F_{\text{quad}}^{12}$  provided by Lemma 13, the invariant cylinder of  $F_{\text{quad}}^{12}$  is just foliated by invariant quasiperiodic tori. Abusing notation, we use the same notation for the objects in the reduced and the full phase space.

Now, write  $\bar{H} = L_2^6 F_{\text{sec}} - \tilde{F}_{\text{quad}}^{12}$ , and define the Poincaré-Melnikov potential by

$$(95) \quad \mathcal{L}(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3) = \int_{-\infty}^{\infty} (\bar{H}(Z^{\text{quad}}(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3)) - \bar{H}(Z_*^{\text{quad}}(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3))) dt.$$

As with  $\bar{H}$  itself, the Poincaré-Melnikov potential  $\mathcal{L}$  can be expanded in ratios of powers of  $L_2$  and  $L_3$ . The following result gives an expression for the first-order term at which each angle  $\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3$  appears in the expansion of  $\mathcal{L}$ . The proposition is proved in Section 6.2.

**Proposition 27.** — *The expansion of the Poincaré-Melnikov potential  $\mathcal{L}$  satisfies the following properties, where the notation  $\alpha_k^{ij}, H_k^{ij}$  is as in Proposition 11.*

- (1) *The first nontrivial term in the expansion of  $\mathcal{L}$  is  $\frac{1}{L_2^2} \alpha_2^{12} \mathcal{L}_2^{12}$  where*

$$\begin{aligned} \mathcal{L}_2^{12}(\tilde{\gamma}_2, \tilde{\Gamma}_2) &= \int_{-\infty}^{\infty} (H_2^{12}(Z^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3)) \\ &\quad - H_2^{12}(Z_*^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3))) dt \\ &= \tilde{\mathcal{L}}_2^{12}(\tilde{\Gamma}_2) \sin \tilde{\gamma}_2 \end{aligned}$$

*and where  $\tilde{\mathcal{L}}_2^{12}$  is an analytic function of  $\tilde{\Gamma}_2$  that is nonvanishing for  $\tilde{\Gamma}_2 \in [\zeta_1, \zeta_2]$  (see (70), (65), and Appendix F).*

- (2) *The first term in the expansion of  $\mathcal{L}$  that depends on  $\tilde{\psi}_1$  is  $\frac{L_2^9}{L_3^6} \alpha_2^{23} \mathcal{L}_2^{23}$  where*

$$\mathcal{L}_2^{23}(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\Gamma}_2) = \int_{-\infty}^{\infty} (H_2^{23}(Z^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3))$$

$$\begin{aligned}
 & - H_2^{23}(Z_*^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3)) \Big) dt \\
 & = \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 I_1^2} \right) [c_1 \cos \tilde{\gamma}_2 \cos \tilde{\psi}_1 + c_2 \sin \tilde{\gamma}_2 \sin \tilde{\psi}_1]
 \end{aligned}$$

where the function  $\kappa$  is defined in (94), the constants  $c_1, c_2$  are defined in (41), and  $\chi, A_2$  are the constants defined in (57).

(3) The first term in the expansion of  $\mathcal{L}$  that depends on  $\tilde{\gamma}_3$  is  $\frac{L_2^{11}}{I_3^8} \alpha_5^{23} \mathcal{L}_5^{23}$  where

$$\begin{aligned}
 \mathcal{L}_5^{23}(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2) & = \int_{-\infty}^{\infty} (H_5^{23}(Z^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3)) \\
 & \quad - H_5^{23}(Z_*^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3))) dt \\
 & = \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 I_1^2} \right) (J_1(\tilde{\psi}_1, \tilde{\gamma}_3) \cos \tilde{\gamma}_2 + J_2(\tilde{\psi}_1, \tilde{\gamma}_3) \sin \tilde{\gamma}_2)
 \end{aligned}$$

where again  $\kappa$  is the function defined by (94), and where

$$\begin{aligned}
 J_1(\tilde{\psi}_1, \tilde{\gamma}_3) & = 30 \delta_3^2 \sin \tilde{\gamma}_3 \cos \tilde{\psi}_1 \sin \tilde{\psi}_1 - 10 \delta_1^2 \sin \tilde{\gamma}_3 \cos \tilde{\psi}_1 \sin \tilde{\psi}_1 \\
 & \quad - 20 \delta_1 \delta_3 \cos \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 + 10 \delta_1 \delta_3 \cos \tilde{\gamma}_3 \\
 J_2(\tilde{\psi}_1, \tilde{\gamma}_3) & = -50 \delta_1 \delta_3 \cos \tilde{\gamma}_3 \cos \tilde{\psi}_1 \sin \tilde{\psi}_1 + \frac{70 \delta_3}{\delta_1} \cos \tilde{\gamma}_3 \cos \tilde{\psi}_1 \sin \tilde{\psi}_1 \\
 & \quad + \frac{105 \delta_3^2}{\delta_1^2} \sin \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 \\
 & \quad - 75 \delta_3^2 \sin \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 + 25 \delta_1^2 \sin \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 \\
 & \quad - 35 \sin \tilde{\gamma}_3 \cos^2 \tilde{\psi}_1 - \frac{105 \delta_3^2}{\delta_1^2} \sin \tilde{\gamma}_3 \\
 & \quad + 60 \delta_3^2 \sin \tilde{\gamma}_3 - 17 \delta_1^2 \sin \tilde{\gamma}_3 + 28 \sin \tilde{\gamma}_3.
 \end{aligned}$$

The following result guarantees the existence of two homoclinic channels, and thus two scattering maps; furthermore it provides first-order approximations of the scattering maps in ‘tilde’ variables.

**Lemma 28.** — *The secular Hamiltonian has two homoclinic channels  $\Gamma_{\pm}$  corresponding to the normally hyperbolic invariant manifold  $\Lambda$ , and there are two scattering maps defined globally on  $\Lambda$  (see (70) and (65)) in the variables (29) by*

$$\tilde{S}_{\pm} : (\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3) \longmapsto (\tilde{\gamma}_2^*, \tilde{\psi}_1^*, \tilde{\gamma}_3^*, \tilde{\Gamma}_2^*, \tilde{\Psi}_1^*, \tilde{\Gamma}_3^*)$$

with

$$\tilde{\gamma}_2^* = \tilde{\gamma}_2 + \Delta_2^\pm(\tilde{\Gamma}_2; \dots), \quad \tilde{\psi}_1^* = \tilde{\psi}_1 + \frac{1}{L_2^2} \Delta_1^\pm(\tilde{\Gamma}_2; \dots),$$

$$\tilde{\gamma}_3^* = \tilde{\gamma}_3 + \frac{L_2^{10}}{L_3^7} \Delta_3^\pm(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3)$$

and

$$\tilde{\Gamma}_2^* = \tilde{\Gamma}_2 + \frac{L_2^8}{L_3^6} \Theta_2^\pm(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots), \quad \tilde{\Psi}_1^* = \tilde{\Psi}_1 + \frac{L_2^9}{L_3^6} \Theta_1^\pm(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots),$$

$$\tilde{\Gamma}_3^* = \tilde{\Gamma}_3 + \frac{L_2^{11}}{L_3^8} \Theta_3^\pm(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2; \dots)$$

where the ellipsis after the semicolon denotes dependence on the remaining variables at higher order, and where

$$\begin{aligned} \Delta_2^\pm(\tilde{\Gamma}_2; \dots) &= 2 \arctan \chi^{-1} + \dots, \\ \Delta_1^\pm(\tilde{\Gamma}_2; \dots) &= \frac{L_1}{6} \sqrt{\frac{3}{2}} \beta_2 \tilde{\Gamma}_2 \sqrt{1 - \frac{5}{3} \frac{\tilde{\Gamma}_2^2}{L_1^2}} + \dots, \quad \Delta_3^\pm = O(1) \end{aligned}$$

and

$$\begin{aligned} \Theta_2^\pm(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots) &= \pm \alpha_1^{12} \alpha_2^{23} \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 L_1^2} \right) c_2 \left( 3 \frac{\tilde{\Gamma}_2^2}{L_1^2} - 1 \right) \frac{L_1^2}{2 \tilde{\Gamma}_2} \cos \tilde{\psi}_1 + \dots \\ \Theta_1^\pm(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots) &= \mp \alpha_2^{23} \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 L_1^2} \right) c_2 \cos \tilde{\psi}_1 + \dots \\ \Theta_3^\pm(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2; \dots) &= \mp \alpha_5^{23} \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 L_1^2} \right) \frac{\partial J_2}{\partial \tilde{\gamma}_3}(\tilde{\psi}_1, \tilde{\gamma}_3) + \dots \end{aligned}$$

where  $\kappa$  is the function defined by (94).

*Proof.* — Denote by  $(\omega_0, \omega_1, \omega_2)$  the frequency vector of the angles  $(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3)$  on a torus on  $\Lambda$  corresponding to fixed values of the actions  $\tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3$  for the Hamiltonian  $\tilde{F}_{\text{quad}}^{12}$ . Its particular form, given in Lemma 13, implies that

$$\omega_0 = 2\alpha_0^{12} \frac{\tilde{\Gamma}_2}{L_1^2} + O\left(\frac{1}{L_2}\right) \quad \omega_1 = O\left(\frac{1}{L_2}\right), \quad \omega_2 = 0.$$

Consider the function

$$(96) \quad \tau \longmapsto \mathcal{L}(\tilde{\gamma}_2 - \omega_0 \tau, \tilde{\psi}_1 - \omega_1 \tau, \tilde{\gamma}_3 - \omega_2 \tau, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3)$$

where  $\mathcal{L}$  is the Poincaré-Melnikov potential defined by (95). Results of [31] imply that nondegenerate critical points of (96) correspond to transverse homoclinic intersections of the stable and unstable manifolds of  $\Lambda$ . Equation (95), Proposition 11, and Proposition 27 imply that

$$\mathcal{L} = \frac{1}{L_2^2} \alpha_2^{12} \mathcal{L}_2^{12} + O\left(\frac{L_2^9}{L_3^6}\right).$$

The function  $\tau \mapsto \mathcal{L}_2^{12}(\tilde{\gamma}_2 - \omega_0 \tau, \tilde{\Gamma}_2)$  has nondegenerate critical points  $\tau_{\pm}$  where  $\omega_0 \tau_{\pm} = \tilde{\gamma}_2 \pm \frac{\pi}{2}$ . It follows that there are functions

$$\tau_{\pm}^*(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3) = \frac{1}{\omega_0} \left( \tilde{\gamma}_2 \pm \frac{\pi}{2} \right) + O\left(\frac{L_2^{11}}{L_3^6}\right)$$

such that

$$\frac{d}{d\tau} \Big|_{\tau=\tau_{\pm}^*} \mathcal{L}(\tilde{\gamma}_2 - \omega_0 \tau, \tilde{\psi}_1 - \omega_1 \tau, \tilde{\gamma}_3 - \omega_2 \tau, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3) = 0.$$

We now introduce the reduced Poincaré-Melnikov potentials

$$\begin{aligned} \mathcal{L}_{\pm}^*(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3) \\ = \mathcal{L}(\tilde{\gamma}_2 - \omega_0 \tau_{\pm}^*, \tilde{\psi}_1 - \omega_1 \tau_{\pm}^*, \tilde{\gamma}_3 - \omega_2 \tau_{\pm}^*, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3). \end{aligned}$$

Now, following again [31], the changes in the actions coming from the scattering maps  $\tilde{S}_{\pm}$  are defined using the functions  $\mathcal{L}_{\pm}^*$  via

$$\tilde{\Gamma}_2^* = \tilde{\Gamma}_2 + \frac{\partial \mathcal{L}_{\pm}^*}{\partial \tilde{\gamma}_2} + \dots, \quad \tilde{\Psi}_1^* = \tilde{\Psi}_1 + \frac{\partial \mathcal{L}_{\pm}^*}{\partial \tilde{\psi}_1} + \dots, \quad \tilde{\Gamma}_3^* = \tilde{\Gamma}_3 + \frac{\partial \mathcal{L}_{\pm}^*}{\partial \tilde{\gamma}_3} + \dots$$

Note that the cylinder frequencies in the model in [31] all have the same time scale and moreover the first order of the perturbation depends on all the angles. On the contrary, in our model all have different speeds and the angles  $\tilde{\psi}_1$  and  $\tilde{\gamma}_3$  appear only at higher order terms (see Proposition 11). Still, one can easily check that the statements in [31] are still valid in the present setting. The only difference is that the first order of the scattering maps in the actions  $\tilde{\Psi}_1$  and  $\tilde{\Gamma}_3$  come from higher orders of the Melnikov potential.

Indeed, we have

$$\begin{aligned} \frac{\partial \mathcal{L}_{\pm}^*}{\partial \tilde{\gamma}_2} &= \frac{L_2^9}{L_3^6} \alpha_2^{23} \frac{\partial \mathcal{L}_2^{23}}{\partial \tilde{\psi}_1} \frac{\partial}{\partial \tilde{\gamma}_2} (\tilde{\psi}_1 - \tau_{\pm}^* \omega_1) + \dots \\ &= -\frac{L_2^9}{L_3^6} \alpha_2^{23} \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 L_1^2} \right) c_2 \sin\left(\mp \frac{\pi}{2}\right) \cos(\tilde{\psi}_1) \omega_1 \frac{\partial \tau_{\pm}^*}{\partial \tilde{\gamma}_2} + \dots \end{aligned}$$

$$\begin{aligned}
 &= \pm \frac{L_2^8}{L_3^6} \alpha_1^{12} \alpha_2^{23} \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 L_1^2} \right) c_2 \left( 3 \frac{\tilde{\Gamma}_2^2}{L_1^2} - 1 \right) \frac{1}{\omega_0} \cos \tilde{\psi}_1 + \dots \\
 &= \pm \frac{L_2^8}{L_3^6} \alpha_1^{12} \alpha_2^{23} \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 L_1^2} \right) c_2 \left( 3 \frac{\tilde{\Gamma}_2^2}{L_1^2} - 1 \right) \frac{L_1^2}{2 \tilde{\Gamma}_2} \cos \tilde{\psi}_1 + \dots, \\
 \frac{\partial \mathcal{L}_\pm^*}{\partial \tilde{\psi}_1} &= \frac{L_2^9}{L_3^6} \alpha_2^{23} \frac{\partial \mathcal{L}_2^{23}}{\partial \tilde{\psi}_1} + \dots = \mp \frac{L_2^9}{L_3^6} \alpha_2^{23} \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 L_1^2} \right) c_2 \cos \tilde{\psi}_1 + \dots,
 \end{aligned}$$

and

$$\frac{\partial \mathcal{L}_\pm^*}{\partial \tilde{\gamma}_3} = \frac{L_2^{11}}{L_3^8} \alpha_5^{23} \frac{\partial \mathcal{L}_5^{23}}{\partial \tilde{\gamma}_3} + \dots = \mp \frac{L_2^{11}}{L_3^8} \alpha_5^{23} \kappa \left( \frac{\pi \tilde{\Gamma}_2}{A_2 L_1^2} \right) \frac{\partial J_2}{\partial \tilde{\gamma}_3}(\tilde{\psi}_1, \tilde{\gamma}_3) + \dots$$

For the angles  $\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3$ , the first-order term under application of the scattering map is a so-called *phase shift*. This is a change in the angle that comes from the integrable part of the Hamiltonian along the separatrix, and does not necessarily depend on the functions  $\mathcal{L}_\pm^*$  at first order. The phase shift in  $\tilde{\gamma}_2$  comes from (60) as follows:

$$\Delta_2(\tilde{\Gamma}_2; \dots) = \lim_{t \rightarrow +\infty} (\tilde{\gamma}_2^2(t) - \tilde{\gamma}_2^2(-t)) = 2 \arctan \chi^{-1}.$$

The phase shift in  $\tilde{\psi}_1$  is computed in Appendix D. As for the phase shift in  $\tilde{\gamma}_3$ , we simply estimate that it cannot be larger than  $O(\frac{L_2^{10}}{L_3^7})$  for the following reason: the first order term in the expansion of the secular Hamiltonian containing  $\tilde{\Gamma}_3$  is of order  $\frac{L_2^3}{L_3^6}$  (see Proposition 11). Since this term provides no phase shift in  $\tilde{\gamma}_3$  at first order, the largest possible term that can produce a phase shift has another factor of  $\frac{L_2}{L_3}$ . Since we normalise the entire secular Hamiltonian by  $L_2^6$ , we see that the phase shift in  $\tilde{\gamma}_3$  cannot be larger than terms of order  $\frac{L_2^3}{L_3^6} \frac{L_2}{L_3} L_2^6 = \frac{L_2^{10}}{L_3^7}$ .  $\square$

**Remark 29.** — *Note that we do not give an expression for  $\Delta_3$  in Lemma 28. Indeed we require only an estimate on its order; as we will see in the proof of Lemma 30, the phase shift in  $\tilde{\gamma}_3$  is small enough that we can ignore it.*

Next, we combine Lemma 28 and the coordinate transformation  $(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3) \mapsto (\hat{\gamma}_2, \hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3)$  provided by Theorem 18 (see also Remark 20) to produce an expression for the scattering maps  $S_\pm$  in ‘hat’ variables, thus completing the proof of Theorem 26.

**Lemma 30.** — *In the ‘hat’ coordinates, the scattering maps  $S_\pm : \Lambda \rightarrow \Lambda'$  introduced in Lemma 28 are given by*

$$S_\pm : (\hat{\gamma}_2, \hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3) \mapsto (\hat{\gamma}_2^*, \hat{\psi}_1^*, \hat{\gamma}_3^*, \hat{\Gamma}_2^*, \hat{\Psi}_1^*, \hat{\Gamma}_3^*)$$

with

$$\hat{\Psi}_1^* = \hat{\Psi}_1 + \frac{L_2^9}{L_3^6} \mathcal{S}_1^\pm(\hat{\psi}_1, \hat{\Gamma}_2) + \dots, \quad \hat{\Gamma}_3^* = \hat{\Gamma}_3 + \frac{L_2^{11}}{L_3^8} \mathcal{S}_3^\pm(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) + \dots$$

where  $\mathcal{S}_1^\pm$  and  $\mathcal{S}_3^\pm$  are given by (92) and (93) respectively.

*Proof.* — Denote by  $\hat{z}$  a point on  $\Lambda$  in ‘hat’ coordinates, by  $\tilde{z}$  the same point expressed in ‘tilde’ coordinates, and by  $\Phi : \tilde{z} \mapsto \hat{z}$  the coordinate transformation provided by Theorem 18. The maps  $\tilde{S}_\pm$ , defined in Lemma 28 send  $\tilde{z} \mapsto \tilde{z}^*$ . In ‘hat’ coordinates, the scattering maps  $S_\pm$  are therefore defined by

$$\hat{z}^* = S_\pm(\hat{z}) = \Phi \circ \tilde{S}_\pm \circ \Phi^{-1}(\hat{z}).$$

We compute the effect of these maps in the  $\hat{\Psi}_1, \hat{\Gamma}_3$  variables. Comparing Lemma 23 and Lemma 24 in the context of the coordinate transformation  $\Phi$ , notice that we can write

$$\hat{\Psi}_1 = \tilde{\Psi}_1 + \frac{L_2^{11}}{L_3^6} \Phi_1(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots), \quad \hat{\Gamma}_3 = \tilde{\Gamma}_3 + \frac{L_2^{13}}{L_3^8} \Phi_3(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2; \dots),$$

where (see equation (80))

$$\begin{aligned} \Phi_1(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots) &= -\frac{\alpha_0^{23}}{\alpha_1^{12}} \frac{\cos(2\tilde{\psi}_1)}{2(3\frac{\tilde{\Gamma}_2^2}{L_1^2} - 1)} + \dots \\ \Phi_3(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2; \dots) &= \tilde{\alpha} \left( 3\frac{\tilde{\Gamma}_2^2}{L_1^2} - 1 \right)^{-1} \left[ \frac{1}{3} \nu_0 \cos(\tilde{\gamma}_3 + 3\tilde{\psi}_1) + \nu_1 \cos(\tilde{\gamma}_3 + \tilde{\psi}_1) \right. \\ &\quad \left. - \nu_2 \cos(\tilde{\gamma}_3 - \tilde{\psi}_1) - \frac{1}{3} \nu_3 \cos(\tilde{\gamma}_3 - 3\tilde{\psi}_1) \right] + \dots \end{aligned}$$

where  $\tilde{\alpha}$  is a nontrivial constant. Thus

$$\begin{aligned} \hat{\Psi}_1^* &= \tilde{\Psi}_1^* + \frac{L_2^{11}}{L_3^6} \Phi_1(\tilde{\psi}_1^*, \tilde{\Gamma}_2^*; \dots) \\ &= \tilde{\Psi}_1 + \frac{L_2^9}{L_3^6} \Theta_1^\pm(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots) \\ &\quad + \frac{L_2^{11}}{L_3^6} \Phi_1\left(\tilde{\psi}_1 + \frac{1}{L_2^2} \Delta_1^\pm(\tilde{\Gamma}_2; \dots), \tilde{\Gamma}_2 + \frac{L_2^8}{L_3^6} \Theta_2^\pm(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots); \dots\right) \\ &= \hat{\Psi}_1 + \frac{L_2^9}{L_3^6} [\Theta_1^\pm(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots) + \partial_{\tilde{\psi}_1} \Phi_1(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots) \Delta_1^\pm(\tilde{\Gamma}_2; \dots)] + \dots \end{aligned}$$



$$= \hat{\psi}_1 + \frac{L_2^9}{L_3^6} \left[ \Theta_1^\pm(\hat{\psi}_1, \hat{\Gamma}_2; \dots) + \partial_{\hat{\psi}_1} \Phi_1(\hat{\psi}_1, \hat{\Gamma}_2; \dots) \Delta_1^\pm(\hat{\Gamma}_2; \dots) \right] + \dots$$

and

$$\begin{aligned} \hat{\Gamma}_3^* &= \tilde{\Gamma}_3^* + \frac{L_2^{13}}{L_3^8} \Phi_3(\tilde{\psi}_1^*, \tilde{\gamma}_3^*, \tilde{\Gamma}_2^*; \dots) \\ &= \tilde{\Gamma}_3 + \frac{L_2^{11}}{L_3^8} \Theta_3^\pm(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2; \dots) \\ &\quad + \frac{L_2^{13}}{L_3^8} \Phi_3 \left( \tilde{\psi}_1 + \frac{1}{L_2^2} \Delta_1^\pm(\tilde{\Gamma}_2; \dots), \tilde{\gamma}_3 + \frac{L_2^{10}}{L_3^7} \Delta_3^\pm(\dots), \right. \\ &\quad \left. \tilde{\Gamma}_2 + \frac{L_2^8}{L_3^6} \Theta_2^\pm(\tilde{\psi}_1, \tilde{\Gamma}_2; \dots); \dots \right) \\ &= \hat{\Gamma}_3 + \frac{L_2^{11}}{L_3^8} \left[ \Theta_3^\pm(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2; \dots) \right. \\ &\quad \left. + \partial_{\tilde{\psi}_1} \Phi_3(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2; \dots) \Delta_1^\pm(\tilde{\Gamma}_2; \dots) \right] + \dots \\ &= \hat{\Gamma}_3 + \frac{L_2^{11}}{L_3^8} \left[ \Theta_3^\pm(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2; \dots) \right. \\ &\quad \left. + \partial_{\hat{\psi}_1} \Phi_3(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2; \dots) \Delta_1^\pm(\hat{\Gamma}_2; \dots) \right] + \dots \end{aligned}$$

Therefore it remains only to compute

$$\begin{aligned} \mathcal{S}_1^\pm(\hat{\psi}_1, \hat{\Gamma}_2) &= \Theta_1^\pm(\hat{\psi}_1, \hat{\Gamma}_2) + \partial_{\hat{\psi}_1} \Phi_1(\hat{\psi}_1, \hat{\Gamma}_2) \Delta_1^\pm(\hat{\Gamma}_2; \dots) \\ &= \mp \alpha_2^{23} \kappa \left( \frac{\pi \hat{\Gamma}_2}{A_2 L_1^2} \right) c_2 \cos \hat{\psi}_1 \\ &\quad + \frac{\alpha_0^{23}}{\alpha_1^{12}} \beta_2 \frac{L_1}{6} \sqrt{\frac{3}{2}} \hat{\Gamma}_2 \sqrt{1 - \frac{5}{3} \frac{\hat{\Gamma}_2^2}{L_1^2} \frac{\sin 2\hat{\psi}_1}{3 \frac{\hat{\Gamma}_2^2}{L_1^2} - 1}} \end{aligned}$$

and

$$\begin{aligned} \mathcal{S}_3^\pm(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) &= \Theta_3^\pm(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) + \partial_{\hat{\psi}_1} \Phi_3(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) \Delta_1^\pm(\hat{\Gamma}_2) \\ &= \mp \alpha_5^{23} \kappa \left( \frac{\pi \hat{\Gamma}_2}{A_2 L_1^2} \right) \frac{\partial J_2}{\partial \hat{\gamma}_3}(\hat{\psi}_1, \hat{\gamma}_3) \end{aligned}$$

$$\begin{aligned}
 & + \frac{L_1}{6} \sqrt{\frac{3}{2}} \beta_2 \hat{\Gamma}_2 \sqrt{1 - \frac{5}{3} \frac{\hat{\Gamma}_2^2}{L_1^2}} \tilde{\alpha} \left( 3 \frac{\hat{\Gamma}_2^2}{L_1^2} - 1 \right)^{-1} \\
 & \times \left[ \nu_0 \cos(\hat{\gamma}_3 + 3\hat{\psi}_1) + \nu_1 \cos(\hat{\gamma}_3 + \hat{\psi}_1) + \nu_2 \cos(\hat{\gamma}_3 - \hat{\psi}_1) + \nu_3 \cos(\hat{\gamma}_3 - 3\hat{\psi}_1) \right] \\
 & = \mp \alpha_5^{23} \kappa \left( \frac{\pi \hat{\Gamma}_2}{A_2 L_1^2} \right) \\
 & \times \left[ \begin{aligned} & 50 \delta_1 \delta_3 \sin \hat{\gamma}_3 \cos \hat{\psi}_1 \sin \hat{\psi}_1 - \frac{70 \delta_3}{\delta_1} \sin \hat{\gamma}_3 \cos \hat{\psi}_1 \sin \hat{\psi}_1 \\ & + \frac{105 \delta_3^2}{\delta_1^2} \cos \hat{\gamma}_3 \cos^2 \hat{\psi}_1 - 75 \delta_3^2 \cos \hat{\gamma}_3 \cos^2 \hat{\psi}_1 \\ & + 25 \delta_1^2 \cos \hat{\gamma}_3 \cos^2 \hat{\psi}_1 - 35 \cos \hat{\gamma}_3 \cos^2 \hat{\psi}_1 \\ & - \frac{105 \delta_3^2}{\delta_1^2} \cos \hat{\gamma}_3 + 60 \delta_3^2 \cos \hat{\gamma}_3 - 17 \delta_1^2 \cos \hat{\gamma}_3 \\ & + 28 \cos \hat{\gamma}_3 \end{aligned} \right] \\
 & + \frac{L_1}{6} \sqrt{\frac{3}{2}} \beta_2 \hat{\Gamma}_2 \sqrt{1 - \frac{5}{3} \frac{\hat{\Gamma}_2^2}{L_1^2}} \tilde{\alpha} \left( 3 \frac{\hat{\Gamma}_2^2}{L_1^2} - 1 \right)^{-1} \left[ \nu_0 \cos(\hat{\gamma}_3 + 3\hat{\psi}_1) \right. \\
 & \left. + \nu_1 \cos(\hat{\gamma}_3 + \hat{\psi}_1) + \nu_2 \cos(\hat{\gamma}_3 - \hat{\psi}_1) + \nu_3 \cos(\hat{\gamma}_3 - 3\hat{\psi}_1) \right]. \quad \square
 \end{aligned}$$

**6.2. Computation of the Poincaré-Melnikov potential.** — In this section we address the proof of Proposition 27. First of all, note that part 1 of the proposition was proved in [41] (see also Appendix F), since  $F_{\text{sec}}^{12}$  coincides with the secular Hamiltonian from the three body problem. It remains to prove parts 2 and 3 of the proposition. By Proposition 11, the first terms that could potentially split the separatrices in the  $\tilde{\psi}_1, \tilde{\Gamma}_3$  directions are  $H_2^{23}, H_5^{23}$  respectively, since they are the first-order terms combining the separatrix variables  $\gamma_1, \Gamma_1, \tilde{\gamma}_2$  with the angles  $\tilde{\psi}_1, \tilde{\gamma}_3$  respectively. Since the variables  $\tilde{\psi}_1, \tilde{\gamma}_3$  are constant with respect to  $H_0^{12}$ , we can easily write the periodic orbits  $Z_{\min, \max}^0$  and the heteroclinic orbit  $Z^0$  as functions of  $(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2)$ . We must therefore compute

$$\begin{aligned}
 & \mathcal{L}_j^{23}(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2) \\
 & = \int_0^\infty (H_j^{23} \circ Z^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2) - H_j^{23} \circ Z_{\min}^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2)) dt + \\
 & + \int_{-\infty}^0 (H_j^{23} \circ Z^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2) - H_j^{23} \circ Z_{\max}^0(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2)) dt
 \end{aligned}$$

for  $j = 2, 5$  where  $H_2^{23}$  is defined by (40) and  $H_5^{23}$  is defined by (53). Now, notice that, since  $\tilde{\psi}_1, \tilde{\gamma}_3$  are constant with respect to  $H_0^{12}$ , we can write

$$(97) \quad \begin{aligned} \mathcal{L}_2^{23}(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\Gamma}_2) &= c_1 \cos \tilde{\psi}_1 \mathcal{L}_1(\tilde{\gamma}_2, \tilde{\Gamma}_2) + c_2 \sin \tilde{\psi}_1 \mathcal{L}_2(\tilde{\gamma}_2, \tilde{\Gamma}_2) \\ \mathcal{L}_5^{23}(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2) &= J_1(\tilde{\psi}_1, \tilde{\gamma}_3) \mathcal{L}_1(\tilde{\gamma}_2, \tilde{\Gamma}_2) + J_2(\tilde{\psi}_1, \tilde{\gamma}_3) \mathcal{L}_2(\tilde{\gamma}_2, \tilde{\Gamma}_2) \end{aligned}$$

(with the constants  $c_1, c_2$  defined in Lemma 15 and the expressions  $J_1, J_2$  defined in Lemma 16) where we have discarded the higher-order terms for convenience, where

$$\begin{aligned} \mathcal{L}_j(\tilde{\gamma}_2, \tilde{\Gamma}_2) &= \int_0^\infty (\mathcal{F}_j \circ Z^0(t, \tilde{\gamma}_2, 0, 0, \tilde{\Gamma}_2) - \mathcal{F}_j \circ Z_{\min}^0(t, \tilde{\gamma}_2, 0, 0, \tilde{\Gamma}_2)) dt \\ &\quad + \int_{-\infty}^0 (\mathcal{F}_j \circ Z^0(t, \tilde{\gamma}_2, 0, 0, \tilde{\Gamma}_2) - \mathcal{F}_j \circ Z_{\max}^0(t, \tilde{\gamma}_2, 0, 0, \tilde{\Gamma}_2)) dt, \end{aligned}$$

and where the functions

$$\mathcal{F}_1 = \sqrt{\Gamma_1^2 - \tilde{\Gamma}_2^2} \cos \tilde{\gamma}_2, \quad \mathcal{F}_2 = \sqrt{\Gamma_1^2 - \tilde{\Gamma}_2^2} \sin \tilde{\gamma}_2$$

do not depend on  $\tilde{\psi}_1, \tilde{\gamma}_3$ .

For the rest of this section, to simplify notation, we write  $\gamma, \Gamma$  instead of  $\tilde{\gamma}_2, \tilde{\Gamma}_2$ , respectively. Recall from Lemma 5 that  $\gamma = \gamma^0 + \gamma^1 + \gamma^2$  along the separatrix, where  $\gamma^2$  is defined by (60). Since  $\gamma^2$  does not tend to 0 as  $t$  goes to  $\pm\infty$ , we must take into account the phase shifts

$$(98) \quad \Delta_\pm = \lim_{t \rightarrow \pm\infty} \arctan(\chi^{-1} \tanh(A_2 t)) = \pm \arctan \chi^{-1}$$

along the periodic orbits  $Z_{\min, \max}^0$  in order to ensure that the integrals  $\mathcal{L}_j$  are well-defined. This is done by taking  $\gamma = \gamma^0 + \gamma^1 + \Delta_+$  on  $Z_{\min}^0$  and  $\gamma = \gamma^0 + \gamma^1 + \Delta_-$  on  $Z_{\max}^0$ .

Expanding the Fourier series of  $\mathcal{L}_j$  in  $\gamma^0$  for  $j = 1, 2$ , we have

$$(99) \quad \mathcal{L}_j(\gamma^0, \Gamma) = \mathcal{L}_j^* e^{i\gamma^0} + \overline{\mathcal{L}_j^*} e^{-i\gamma^0}$$

where

$$\mathcal{L}_j^* = \frac{1}{2} \int_0^\infty (\xi_j^+(t) + i\eta_j^+(t)) e^{i\gamma^1(t)} dt + \frac{1}{2} \int_{-\infty}^0 (\xi_j^-(t) + i\eta_j^-(t)) e^{i\gamma^1(t)} dt$$

with

$$(100) \quad \begin{cases} \xi_1^\pm(t) = -\eta_2^\pm(t) = \sqrt{\Gamma_1(t)^2 - \Gamma^2} \cos \gamma^2(t) - \sqrt{L_1^2 - \Gamma^2} \cos \Delta_\pm \\ \eta_1^\pm(t) = \xi_2^\pm(t) = \sqrt{\Gamma_1(t)^2 - \Gamma^2} \sin \gamma^2(t) - \sqrt{L_1^2 - \Gamma^2} \sin \Delta_\pm \end{cases}$$

and where  $\Gamma_1(t)$ , which describes the motion of  $\Gamma_1$  along the separatrix, is defined by (59).

*Lemma 31.* — *We can write*

$$(101) \quad \mathcal{L}_1^* = \frac{i}{2} \sqrt{\frac{2}{3}} \frac{\Gamma}{A_2 \chi} \left[ \int_0^\infty f^+(\tau) d\tau + \int_{-\infty}^0 f^-(\tau) d\tau \right]$$

where

$$(102) \quad f^\pm(\tau) = (\tanh \tau \mp 1) e^{i \frac{2\Gamma}{A_2 L_1^2} \tau}$$

and where  $\tau = A_2 t$ . Furthermore

$$(103) \quad \mathcal{L}_2^* = -i \mathcal{L}_1^*.$$

*Proof.* — As in [41] (see Lemma 5.1, 5.2, and 5.3 of that paper) we first write the integrands of  $\mathcal{L}_1^*$  in terms of  $\gamma_1$  and then in terms of  $\tau = A_2 t$ . It follows directly from (58), (59), and (60) (see also Lemma 5.1 of [41]) that, on the separatrix, we have

$$(104) \quad \cos \gamma^2 = \sqrt{1 - \frac{5}{3} \cos^2 \gamma_1}, \quad \sin \gamma^2 = \sqrt{\frac{5}{3}} \cos \gamma_1, \quad \Gamma_1 = \sqrt{\frac{5}{3}} \Gamma \frac{\sin \gamma_1}{\sqrt{1 - \frac{5}{3} \cos^2 \gamma_1}}.$$

It follows that

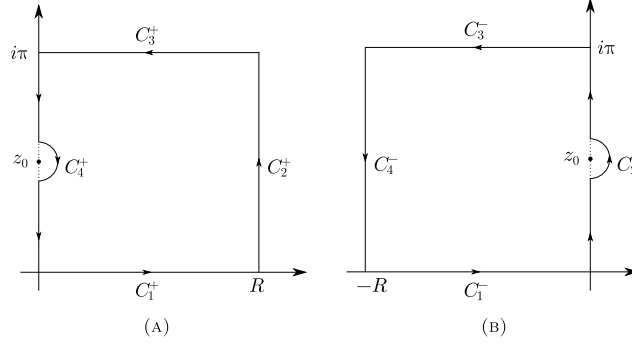
$$(105) \quad \sqrt{\Gamma_1^2 - \Gamma^2} = \sqrt{\frac{2}{3}} \Gamma \frac{1}{\sqrt{1 - \frac{5}{3} \cos^2 \gamma_1}}.$$

Moreover, from (57) and (98) we get

$$(106) \quad \begin{aligned} \cos \Delta_\pm &= \frac{\chi}{\sqrt{1 + \chi^2}} = \sqrt{\frac{2}{3}} \Gamma \frac{1}{\sqrt{L_1^2 - \Gamma^2}}, \\ \sin \Delta_\pm &= \pm \frac{1}{\sqrt{1 + \chi^2}} = \pm \sqrt{\frac{2}{3}} \frac{\Gamma}{\chi} \frac{1}{\sqrt{L_1^2 - \Gamma^2}}. \end{aligned}$$

Combining (100) with (104), (105), and (106) yields

$$(107) \quad \xi_1^\pm = 0, \quad \eta_1^\pm = \sqrt{\frac{2}{3}} \sqrt{\frac{5}{3}} \Gamma \frac{\cos \gamma_1}{\sqrt{1 - \frac{5}{3} \cos^2 \gamma_1}} \mp \sqrt{\frac{2}{3}} \frac{\Gamma}{\chi}.$$


 FIG. 4. — The contours  $C^\pm$  over which we integrate  $f^\pm$ 

Now that  $\eta_1^\pm$  is written in terms of  $\gamma_1$ , we proceed to write it in terms of  $\tau$ . Equation (58) implies that

$$\sqrt{1 - \frac{5}{3} \cos^2 \gamma_1} = \frac{\chi \cosh A_2 t}{\sqrt{\chi^2 + (1 + \chi^2) \sinh^2(A_2 t)}}.$$

This expression, together with (56), (58), and (107) proves (101) and (102). Finally, comparing (100) proves (103).  $\square$

In order to determine the integral of  $f^\pm$  over the positive and negative halves of the real line, we instead integrate  $f^\pm$  over appropriately chosen contours  $C^\pm$  in the complex plane using the residue theorem. Initially we assume the contour has “width”  $R > 0$ . Then, taking the limit as  $R \rightarrow \infty$  allows us to derive an expression for  $\mathcal{L}_1^*$ .

From (102), it can be seen that

$$(108) \quad f^\pm(\tau + i\pi) = e^{-\frac{2\pi\Gamma}{A_2 L_1^2}} f^\pm(\tau).$$

Moreover the only singularity of  $\tanh \tau$ , and thus of  $f^\pm$ , in the region  $\{0 < \text{Im } \tau < \pi\}$  is  $z_0 = i\frac{\pi}{2}$ , and the residue of  $f^\pm$  at  $z_0$  is

$$\text{Res}(f^\pm, z_0) = e^{-\frac{\pi\Gamma}{A_2 L_1^2}}.$$

We therefore define the contours  $C^\pm$  as follows. Fix some  $R > 0$ , and denote by  $C_1^+$  the segment of the real line from 0 to  $R$ ; denote by  $C_2^+$  the straight line from  $R$  to  $R + i\pi$ ; and denote by  $C_3^+$  the straight line from  $R + i\pi$  to  $i\pi$ . The last part  $C_4^+$  of the contour  $C^+$  is composed of two straight lines and a small semicircle centred at  $z_0$ , so as to exclude  $z_0$  from the interior of the region bounded by  $C^+$  (see Figure 4(a)). Similarly,  $C_1^-$  is the segment of the real line from  $-R$  to 0;  $C_2^-$  is  $C_4^+$  traversed in the opposite direction;  $C_3^-$  is the straight line from  $i\pi$  to  $-R + i\pi$ ; and  $C_4^-$  is the straight line from  $-R + i\pi$  to  $-R$ .

(see Figure 4(b)). Thus we define

$$C^\pm = C_1^\pm \cup C_2^\pm \cup C_3^\pm \cup C_4^\pm.$$

The residue theorem implies that

$$\int_{C^+} f^+(\tau) d\tau = 0, \quad \int_{C^-} f^-(\tau) d\tau = 2\pi i e^{-\frac{\pi \Gamma}{A_2 L_1^2}}.$$

In the limit as  $R \rightarrow \infty$  we have

$$\lim_{R \rightarrow \infty} \int_{C_2^+} f^+(\tau) d\tau = 0 = \lim_{R \rightarrow \infty} \int_{C_4^-} f^-(\tau) d\tau$$

and

$$\lim_{R \rightarrow \infty} \int_{C_1^\pm \cup C_3^\pm} f^\pm(\tau) d\tau = \pm \left(1 - e^{-\frac{2\pi \Gamma}{A_2 L_1^2}}\right) \int_0^{\pm\infty} f^\pm(\tau) d\tau$$

due to (108). The last part of the integral is

$$\int_{C_4^+} f^+(\tau) d\tau + \int_{C_2^-} f^-(\tau) d\tau = \int_{C_2^-} (f^-(\tau) - f^+(\tau)) d\tau = 2 \int_{C_2^-} e^{i\frac{2\pi \Gamma}{A_2 L_1^2} \tau} d\tau.$$

As the integrand of the last expression is holomorphic, this equals

$$2 \int_{i[0, \pi]} e^{i\frac{2\pi \Gamma}{A_2 L_1^2} \tau} d\tau = i \frac{A_2 L_1^2}{\Gamma} \left(1 - e^{-\frac{2\pi \Gamma}{A_2 L_1^2}}\right).$$

Therefore summing the integral of  $f^+$  over  $C^+$  and the integral of  $f^-$  over  $C^-$ , and taking the limit as  $R \rightarrow \infty$  yields

$$2\pi i e^{-\frac{\pi \Gamma}{A_2 L_1^2}} = \left(1 - e^{-\frac{2\pi \Gamma}{A_2 L_1^2}}\right) \left[ \int_0^\infty f^+(\tau) d\tau + \int_{-\infty}^0 f^-(\tau) d\tau + i \frac{A_2 L_1^2}{\Gamma} \right].$$

Rearranging terms and using (101) gives

$$\mathcal{L}_1^* = \frac{1}{2} \kappa \left( \frac{\pi \Gamma}{A_2 L_1^2} \right)$$

where  $\kappa$  is the function defined by (94). As this is real-valued, we find from (99) and (103) that

$$\mathcal{L}_1(\gamma^0) = \kappa \left( \frac{\pi \Gamma}{A_2 L_1^2} \right) \cos \gamma^0, \quad \mathcal{L}_2(\gamma^0) = \kappa \left( \frac{\pi \Gamma}{A_2 L_1^2} \right) \sin \gamma^0$$

which, together with (97), completes the proof of Proposition 27.

## 7. Reduction to a Poincaré map and the shadowing argument

The purpose of this section is to show that there are orbits of the four body problem along which the quantities  $\Psi_1$  and  $\Gamma_3$  drift along any itinerary we choose. The idea is to apply the shadowing results proved by the authors in [26] (see Appendix C of the present paper for a summary of those results). As the results of that paper are for maps (as is typically more convenient for shadowing results), we must first reduce the secular system to a Poincaré map; this is done in Section 7.1. In doing so, we show that the map satisfies the assumptions of Theorem 47. We then show in the subsequent sections that the Poincaré map of the full four body problem thus satisfies the assumptions of Theorem 48, and so, there are drifting orbits.

**7.1. The Poincaré map.** — In this section we show that we can extend the coordinates  $(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3)$  on  $\Lambda$  to (probably non-symplectic) coordinates on a subset of the secular phase space, and show that we can define a Poincaré map in the region in which these coordinates are defined, within an energy level of the secular Hamiltonian. We show that the inner map satisfies a twist condition, and deduce first-order expressions for the scattering maps in the action variables.

Consider the region  $\tilde{\mathcal{D}}$  of the phase space in which the variables  $(\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3)$  (introduced in Sections 3 and 4) satisfy  $\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3 \in \mathbf{T}$ ,  $\tilde{\Gamma}_2 \in [\zeta_1, \zeta_2]$ ,  $\tilde{\Psi}_1, \tilde{\Gamma}_3 \in [-1, 1]$ , and  $(\xi, \eta)$  belongs to the open ball of radius  $\sqrt{2L_1}$  centred at the origin in  $\mathbf{R}^2$  (see (61)). Recall the constants  $\zeta_1, \zeta_2$  were defined in (65). We now further restrict these constants by assuming that

$$(109) \quad 0 < \zeta_1 < \zeta_2 < \frac{L_1}{\sqrt{3}}.$$

The reason for this further restriction is to guarantee that the Poincaré map which we will construct in this section satisfies a twist condition; see Lemma 34.

By slightly shrinking this region  $\tilde{\mathcal{D}}$  in the sense of Remark 20, we obtain a domain  $\mathcal{D}$  in which there is a (non-symplectic) coordinate transformation

$$(110) \quad : (\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3) \mapsto (\xi, \eta, \hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3)$$

where  $(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3)$  are the coordinates on  $\Lambda$  constructed in Section 5. Define the subset  $\mathcal{E} \subset \mathbf{R}$  by  $\mathcal{E} = \{F_{\text{sec}}(z) : z \in \mathcal{D}\}$ . The following theorem is the main result of this section.

**Theorem 32.** — *Fix some energy level  $\{F_{\text{sec}} = E_0\}$  of the secular Hamiltonian where  $E_0 \in \mathcal{E}$ , and consider the Poincaré section  $\mathbf{M} = \mathcal{D} \cap \{F_{\text{sec}} = E_0\} \cap \{\hat{\gamma}_2 = 0\}$ .*

- (1) *There is a well-defined Poincaré map  $F : \mathbf{M} \rightarrow \mathbf{M}$ , and the set  $\hat{\Lambda} = \Lambda \cap \mathbf{M}$  is a normally hyperbolic invariant manifold for  $F$ .*

- (2) The variables  $(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3)$  define coordinates on  $\hat{\Lambda}$ , and the inner map  $f = F|_{\hat{\Lambda}}$  has the form

$$(111) \quad f : \begin{cases} (\hat{\psi}_1^*, \hat{\gamma}_3^*) = (\hat{\psi}_1, \hat{\gamma}_3) + g(\hat{\Psi}_1, \hat{\Gamma}_3) + O(\epsilon^{k_1-6} \mu^{k_2}) \\ (\hat{\Psi}_1^*, \hat{\Gamma}_3^*) = (\hat{\Psi}_1, \hat{\Gamma}_3) + O(\epsilon^{k_1-6} \mu^{k_2}) \end{cases}$$

where  $\epsilon = \frac{1}{L_2}$ ,  $\mu = \frac{L_2}{L_3}$ , where  $k_1, k_2 \in \mathbf{N}$  come from part 2 of Theorem 18, and where

$$\det \text{Dg}(\hat{\Psi}_1, \hat{\Gamma}_3) \neq 0.$$

Moreover the bottom eigenvalue of  $\text{Dg}(\hat{\Psi}_1, \hat{\Gamma}_3)$  is of order  $\frac{L_2^8}{L_3^6}$ .

- (3) There are two scattering maps

$$\hat{S}_{\pm} : \hat{\Lambda} \longrightarrow \hat{\Lambda}'$$

where  $\hat{\Lambda}'$  is an open cylinder in  $\mathbf{T}^2 \times \mathbf{R}^2$  containing  $\hat{\Lambda}$ . Moreover the actions  $(\hat{\Psi}_1^*, \hat{\Gamma}_3^*)$  of the image of a point  $(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3) \in \hat{\Lambda}$  under the scattering maps  $\hat{S}_{\pm}$  is given by

$$\hat{\Psi}_1^* = \hat{\Psi}_1 + \frac{L_2^9}{L_3^6} \mathcal{S}_1^{\pm}(\hat{\psi}_1, \hat{\Gamma}_2) + \dots, \quad \hat{\Gamma}_3^* = \hat{\Gamma}_3 + \frac{L_2^{11}}{L_3^8} \mathcal{S}_3^{\pm}(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) + \dots$$

where  $\mathcal{S}_1^{\pm}$  is defined by (92) and  $\mathcal{S}_3^{\pm}$  is defined by (93).

We divide the proof of Theorem 32 into three lemmas. The following lemma establishes the existence of coordinates in a neighbourhood of  $\Lambda$  involving the ‘hat’ variables, and the existence of a Poincaré map.

**Lemma 33.** — The variables  $(\hat{\xi}, \hat{\eta}, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3)$  define coordinates on the section  $\mathbf{M} = \mathcal{D} \cap \{F_{\text{sec}} = E_0\} \cap \{\hat{\gamma}_2 = 0\}$  for  $E_0 \in \mathcal{E}$ , and there is a well-defined Poincaré map  $F : \mathbf{M} \rightarrow \mathbf{M}$ . Furthermore the set  $\hat{\Lambda} = \Lambda \cap \mathbf{M}$  is a normally hyperbolic invariant manifold for  $F$ .

*Proof.* — Denote by  $\hat{\Phi}$  the coordinate transformation (110). The transformation  $\hat{\Phi}$  may not be symplectic, but the flow possesses an integral  $\hat{H} = L_2^6 F_{\text{sec}} \circ \hat{\Phi}$  nonetheless. Fix some level set  $\{\hat{H} = E_0\}$  of the integral. Since  $\hat{\Phi}$  is  $O(\frac{L_2^{11}}{L_3^6})$  close to the identity, we have (differentiating (33) with respect to  $\tilde{\Gamma}_2$  and effecting the coordinate transformation)

$$(112) \quad \frac{\partial \hat{H}}{\partial \hat{\Gamma}_2} = \hat{\Gamma}_2 \left[ \frac{2}{L_1^2} + \frac{10}{\Gamma_1^2} \left( 1 - \frac{\Gamma_1^2}{L_1^2} \right) \sin^2 \gamma_1 \right] + O\left(\frac{L_2^{11}}{L_3^6}\right).$$

Observe that this is nonzero as long as  $\hat{\Gamma}_2 \neq 0$ ; indeed the term  $1 - \frac{\Gamma_1^2}{L_1^2}$  is  $e_1^2$  and thus belongs to  $(0, 1)$ , and so the term inside the square brackets is strictly positive. Therefore, by the implicit function theorem, we may express  $\hat{\Gamma}_2$  as a function of the other variables



in the energy level  $\{\hat{H} = E_0\}$ . Moreover, notice that  $\frac{d\hat{\gamma}_2}{dt}$  is equal, up to higher order terms, to (112). Since the time derivative of  $\hat{\Gamma}_2$  is of higher order, it follows that the return map  $F$  to the section  $M = \mathcal{D} \cap \{F_{\text{sec}} = E_0\} \cap \{\hat{\gamma}_2 = 0\}$  is well-defined as long as  $\hat{\Gamma}_2 \neq 0$ , which is true on  $\mathcal{D}$ .

Finally, the fact that  $\hat{\Lambda} = \Lambda \cap M$  is a normally hyperbolic invariant manifold for the return map  $F$  follows immediately from the fact that  $\Lambda$  is a normally hyperbolic invariant manifold for the flow.  $\square$

**Lemma 34.** — *The inner map  $f = F|_{\hat{\Lambda}}$  has the form*

$$(113) \quad f : \begin{cases} (\hat{\psi}_1^*, \hat{\gamma}_3^*) = (\hat{\psi}_1, \hat{\gamma}_3) + g(\hat{\Psi}_1, \hat{\Gamma}_3) + O(\epsilon^{k_1-6} \mu^{k_2}) \\ (\hat{\psi}_1^*, \hat{\Gamma}_3^*) = (\hat{\psi}_1, \hat{\Gamma}_3) + O(\epsilon^{k_1-6} \mu^{k_2}) \end{cases}$$

where  $\epsilon = \frac{1}{L_2}$ ,  $\mu = \frac{L_2}{L_3}$ , where  $k_1, k_2 \in \mathbf{N}$  come from part 2 of Theorem 18, and where

$$(114) \quad \det Dg(\hat{\Psi}_1, \hat{\Gamma}_3) \neq 0.$$

Moreover the bottom eigenvalue of  $Dg(\hat{\Psi}_1, \hat{\Gamma}_3)$  is of order  $\frac{\mu^6}{\epsilon^2}$ .

*Proof.* — Consider the normalised inner Hamiltonian  $L_2^6 \hat{F}$  where  $\hat{F}$  is defined by (69). By part 2 of Theorem 18, the first order term of this Hamiltonian is  $c_0 \hat{\Gamma}_2^2$ , and so the rate of change of  $\hat{\gamma}_2$  is  $2c_0 \hat{\Gamma}_2 + O(\epsilon)$ . Since this is of order 1, the return time of the flow associated with the Hamiltonian function  $L_2^6 \hat{F}$  to the Poincaré section  $\{\hat{\gamma}_2 = 0\}$  is itself of order 1. Due to equation (69), the angles  $\hat{\psi}_1, \hat{\gamma}_3$  do not appear in the Hamiltonian  $L_2^6 \hat{F}$  before the terms of order  $\epsilon^{k_1-6} \mu^{k_2}$ , and so the return map  $f$  has the form (113). It remains to prove the formula (114), and to determine the order of the bottom eigenvalue of  $Dg$ .

Write  $P_0 = \hat{\Gamma}_2$ ,  $P_1 = \hat{\Psi}_1$ ,  $P_2 = \hat{\Gamma}_3$ , and  $P = (P_1, P_2)$ . Write

$$\omega_i(P_0, P) = \frac{\partial \hat{F}_0}{\partial P_i}$$

for  $i = 0, 1, 2$ , where  $\hat{F}_0$  is defined in (69). By the implicit function theorem, in the energy level  $\{\hat{F}_0 = E_0\}$  where  $E_0 \in \mathcal{E}$ , we can write  $P_0 = \alpha(P)$ , and the derivatives of  $\alpha$  are given by

$$\frac{\partial \alpha}{\partial P_j}(P) = -\frac{\hat{\omega}_j(P)}{\hat{\omega}_0(P)}$$

where we have defined  $\hat{\omega}_i(P) = \omega_i(\alpha(P), P)$ . Using this notation we have  $g_i(P) = \hat{\omega}_0(P)^{-1} \hat{\omega}_i(P)$ . In the following computation, for convenience, we suppress the depen-

dence of functions and their derivatives on  $\mathbf{P}$  and  $\mathbf{P}_0 = \alpha(\mathbf{P})$ .

$$\begin{aligned}
 \hat{\omega}_0^3 (\mathbf{D}g)_{\bar{j}} &= \hat{\omega}_0^3 \frac{\partial g_i}{\partial \mathbf{P}_j} = \hat{\omega}_0^2 \left( \frac{\partial \omega_i}{\partial \mathbf{P}_0} \frac{\partial \alpha}{\partial \mathbf{P}_j} + \frac{\partial \omega_i}{\partial \mathbf{P}_j} \right) - \hat{\omega}_0 \hat{\omega}_i \left( \frac{\partial \omega_0}{\partial \mathbf{P}_0} \frac{\partial \alpha}{\partial \mathbf{P}_j} + \frac{\partial \omega_0}{\partial \mathbf{P}_j} \right) \\
 &= -\hat{\omega}_0 \hat{\omega}_j \frac{\partial \omega_i}{\partial \mathbf{P}_0} + \hat{\omega}_0^2 \frac{\partial \omega_i}{\partial \mathbf{P}_j} + \hat{\omega}_i \hat{\omega}_j \frac{\partial \omega_0}{\partial \mathbf{P}_0} - \hat{\omega}_0 \hat{\omega}_i \frac{\partial \omega_0}{\partial \mathbf{P}_j} \\
 &= -\frac{\partial \hat{\mathbf{F}}_0}{\partial \mathbf{P}_0} \frac{\partial \hat{\mathbf{F}}_0}{\partial \mathbf{P}_j} \frac{\partial^2 \hat{\mathbf{F}}_0}{\partial \mathbf{P}_0 \partial \mathbf{P}_i} + \left( \frac{\partial \hat{\mathbf{F}}_0}{\partial \mathbf{P}_0} \right)^2 \frac{\partial^2 \hat{\mathbf{F}}_0}{\partial \mathbf{P}_j \partial \mathbf{P}_i} + \frac{\partial \hat{\mathbf{F}}_0}{\partial \mathbf{P}_i} \frac{\partial \hat{\mathbf{F}}_0}{\partial \mathbf{P}_j} \frac{\partial^2 \hat{\mathbf{F}}_0}{\partial \mathbf{P}_0^2} \\
 &\quad - \frac{\partial \hat{\mathbf{F}}_0}{\partial \mathbf{P}_0} \frac{\partial \hat{\mathbf{F}}_0}{\partial \mathbf{P}_i} \frac{\partial^2 \hat{\mathbf{F}}_0}{\partial \mathbf{P}_j \partial \mathbf{P}_0}.
 \end{aligned}$$

Combining this formula with the derivatives given in Lemma 25, we see that

$$\begin{aligned}
 \hat{\omega}_0^3 (\mathbf{D}g)_{11} &= \epsilon^{20} C_{12}^3 \frac{54 (L_1^2 - 3 \hat{\Gamma}_2^2) (L_1^2 + \hat{\Gamma}_2^2)}{L_1^6 \delta_1^{11}} + \dots \\
 \hat{\omega}_0^3 (\mathbf{D}g)_{12} &= -\epsilon^{16} \mu^6 C_{12}^2 C_{23} \frac{72 (3 L_1^2 \delta_1^2 + 9 \hat{\Gamma}_2^2 \delta_1^2 - 5 L_1^2 + 5 \hat{\Gamma}_2^2) \delta_3}{L_1^4 \delta_1^9 \delta_2^3} + \dots \\
 \hat{\omega}_0^3 (\mathbf{D}g)_{21} &= -\epsilon^{16} \mu^6 C_{12}^2 C_{23} \frac{72 (3 L_1^2 \delta_1^2 + 9 \hat{\Gamma}_2^2 \delta_1^2 - 5 L_1^2 + 5 \hat{\Gamma}_2^2) \delta_3}{L_1^4 \delta_1^9 \delta_2^3} + \dots \\
 \hat{\omega}_0^3 (\mathbf{D}g)_{22} &= -\epsilon^{16} \mu^6 C_{12}^2 C_{23} \frac{36 \hat{\Gamma}_2^2 (12 \delta_1^2 - 20)}{L_1^4 \delta_1^8 \delta_2^3} + \dots
 \end{aligned}$$

As a result of (11) we have  $\epsilon^{20} \gg \epsilon^{16} \mu^6$ . Since  $O(\hat{\omega}_0^3) = \epsilon^{18}$ , the top eigenvalue of  $\mathbf{D}g$  is of order  $O((\mathbf{D}g)_{11}) = \epsilon^2$ . Moreover, the determinant is equal, up to higher order terms, to the product of the entries on the diagonal of  $\mathbf{D}g$ ; therefore it is of order  $\mu^6$ , and it is nonzero as long as

$$\hat{\Gamma}_2^2 (L_1^2 - 3 \hat{\Gamma}_2^2) (L_1^2 + \hat{\Gamma}_2^2) (12 \delta_1^2 - 20) \neq 0.$$

This is true whenever  $\hat{\Gamma}_2 \in (\zeta_1, \zeta_2)$  where  $\zeta_1, \zeta_2$  satisfy condition (109) is satisfied since the second last factor is strictly positive, and since  $\delta_1^2 \in (0, 1)$  implies that the last factor never vanishes. The fact that the bottom eigenvalue is of order  $\epsilon^{-2} \mu^6$  follows from the fact that the top eigenvalue is of order  $\epsilon^2$  and the determinant is of order  $\mu^6$ .  $\square$

*Remark 35.* — The fact that condition (109) implies the twist property (114) can also be obtained via Arnold's isoenergetic nondegeneracy condition; that is, the condition that the so-called bordered Hessian of  $\hat{\mathbf{F}}_0$  (where  $\hat{\mathbf{F}}_0$  is defined by (69)) has nonzero determinant (see Appendix 8 of [5] and the references therein). The reason we have computed explicitly the derivative of the frequency vector

$g$  is that we need to know the order of the eigenvalues of  $\mathbf{D}g$  in order to obtain estimates on the diffusion time; the bordered Hessian and its determinant do not readily provide this data.

**Lemma 36.** — *There are two scattering maps*

$$\hat{S}_\pm : \hat{\Lambda} \longrightarrow \hat{\Lambda}'$$

where  $\hat{\Lambda}'$  is an open cylinder in  $\mathbf{T}^2 \times \mathbf{R}^2$  containing  $\hat{\Lambda}$ . Moreover the actions  $\hat{\Psi}_1^*, \hat{\Gamma}_3^*$  of the image of a point  $(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3) \in \Lambda$  under the scattering maps  $\hat{S}_\pm$  are given by

$$\hat{\Psi}_1^* = \hat{\Psi}_1 + \frac{L_2^9}{L_3^6} \mathcal{S}_1^\pm(\hat{\psi}_1, \hat{\Gamma}_2) + \dots, \quad \hat{\Gamma}_3^* = \hat{\Gamma}_3 + \frac{L_2^{11}}{L_3^8} \mathcal{S}_3^\pm(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) + \dots$$

where  $\hat{\Gamma}_2$  is a function of the coordinates  $(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3)$  and the energy  $E_0 \in \mathcal{E}$  on the cylinder  $\hat{\Lambda}$ , where  $\mathcal{S}_1^\pm$  is defined by (92) and  $\mathcal{S}_3^\pm$  is defined by (93).

*Proof.* — By Lemma 28, there are two homoclinic channels  $\Gamma_\pm$  corresponding to the normally hyperbolic invariant manifold  $\Lambda$ , which give rise to two globally defined scattering maps  $S_\pm : \Lambda \rightarrow \Lambda'$  for some cylinder  $\Lambda'$  such that  $\Lambda \subset \Lambda' \subset \mathbf{T}^3 \times \mathbf{R}^3$ . It follows that the sets  $\hat{\Gamma}_\pm = \Gamma_\pm \cap \{\hat{\gamma}_2 = 0\} \cap \{F_{\text{sec}} = E_0\}$  are homoclinic channels for the return map  $F$  to the section  $\{\hat{\gamma}_2 = 0\} \cap \{F_{\text{sec}} = E_0\}$ . Let  $\hat{\Lambda}' = \Lambda' \cap \{\hat{\gamma}_2 = 0\} \cap \{F_{\text{sec}} = E_0\}$ , and denote by  $\hat{S}_\pm : \hat{\Lambda} \rightarrow \hat{\Lambda}'$  the scattering maps corresponding to the homoclinic channels  $\hat{\Gamma}_\pm$ . Denote by  $\hat{\phi}^t$  the time- $t$  map of the flow of the Hamiltonian function  $L_2^6 \hat{F}$  where the inner Hamiltonian  $\hat{F}$  is defined by (69). Proposition 3 of [33] implies that there are smooth functions  $\tau_\pm : \hat{\Lambda} \rightarrow \mathbf{R}$  such that

$$\hat{S}_\pm(\hat{z}) = \hat{\phi}^{\tau_\pm(\hat{z})} \circ S_\pm(\hat{z})$$

and  $\tau_\pm(\hat{z}) = O(1)$  for all  $\hat{z} \in \hat{\Lambda}$ . Indeed, the fact that  $\tau_\pm(\hat{z}) = O(1)$  is due to the functions  $\tau_\pm$  being constructed as return times with respect to the Hamiltonian  $L_2^6 \hat{F}$  to the section  $\{\hat{\gamma}_2 = 0\}$ ; since the time derivative of  $\hat{\gamma}_2$  with respect to  $L_2^6 \hat{F}$  is of order 1, so too is the return time.

Now, choose any point  $\hat{z} = (\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3) \in \hat{\Lambda}$ . This gives rise to a point  $z = (\hat{\gamma}_2, \hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3)$  in  $\Lambda$  where  $\hat{\gamma}_2 = 0$  and  $\hat{\Gamma}_2$  is a function of  $\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3$  and the energy  $E_0$ . Write  $\bar{z} = S_\pm(z)$ . Then the  $\bar{\Psi}_1, \bar{\Gamma}_3$  components of  $\bar{z}$  are

$$(115) \quad \bar{\Psi}_1 = \hat{\Psi}_1 + \frac{L_2^9}{L_3^6} \mathcal{S}_1^\pm(\hat{\psi}_1, \hat{\Gamma}_2) + \dots, \quad \bar{\Gamma}_3 = \hat{\Gamma}_3 + \frac{L_2^9}{L_3^6} \mathcal{S}_3^\pm(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) + \dots$$

Write  $\hat{z}^* = \hat{S}_\pm(\hat{z}) = \hat{\phi}^{\tau_\pm(\hat{z})}(\bar{z})$ . Since  $\tau_\pm(\hat{z})$  is of order 1, part 2 of Theorem 18 implies that the  $\hat{\Psi}_1^*, \hat{\Gamma}_3^*$  components of  $\hat{z}^*$  satisfy

$$(116) \quad (\hat{\Psi}_1^*, \hat{\Gamma}_3^*) = (\bar{\Psi}_1, \bar{\Gamma}_3) + O(\epsilon^{k_1-6} \mu^{k_2})$$

where  $k_1, k_2$  are assumed to be sufficiently large. Combining (115) and (116) completes the proof of the lemma.  $\square$

**7.2. Constructing transition chains of almost invariant tori.** — Observe that the normally hyperbolic invariant manifold  $\hat{\Lambda}$  constructed in Theorem 32 has a foliation by leaves

$$(117) \quad \mathcal{L}(\hat{\Psi}_1^*, \hat{\Gamma}_3^*) = \{(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3) \in \hat{\Lambda} : \hat{\Psi}_1 = \hat{\Psi}_1^*, \hat{\Gamma}_3 = \hat{\Gamma}_3^*\},$$

each of which is almost invariant by the inner map  $f$ , since it is of the form (111). The following result allows us to construct sequences of leaves  $\{\mathcal{L}_j\}$  of the foliation connected by the scattering maps  $\hat{S}_\pm$  (i.e. *transition chains*) such that the image of each leaf  $\mathcal{L}_j$  under one of the scattering maps  $\hat{S}_{\beta_j}$  (where  $\beta_j \in \{+, -\}$ ) is mapped transversely across  $\mathcal{L}_{j+1}$  in the following sense: there is a  $z \in \mathcal{L}_j$  such that  $\hat{S}_{\beta_j}(z) \in \mathcal{L}_{j+1}$ , and

$$(118) \quad T_{\hat{S}_{\beta_j}(z)} \Lambda = T_{\hat{S}_{\beta_j}(z)}(\hat{S}_{\beta_j}(\mathcal{L}_j)) \oplus T_{\hat{S}_{\beta_j}(z)} \mathcal{L}_{j+1}.$$

This, combined with the results of Section 7.1, will allow us to apply shadowing results from [26]. Recall that the first order terms  $\mathcal{S}_1^\pm, \mathcal{S}_3^\pm$  in the expansion of the images  $\bar{\Psi}_1, \bar{\Gamma}_3$  of a point  $(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3)$  under the scattering maps  $\hat{S}_\pm : \hat{\Lambda} \rightarrow \hat{\Lambda}'$  are given by (92), (93) respectively due to Theorem 32.

**Lemma 37.** — *Consider the set  $\hat{\Lambda} = \Lambda \cap \mathbf{M}$  defined in Theorem 32. Then there are constants  $\nu^\pm, \hat{\nu}^\pm, \xi^\pm, \hat{\xi}^\pm > 0$  and  $C^\pm > 0$  such that for  $L_2$  large enough (see (21)) and any leaf  $\mathcal{L}^* = \mathcal{L}(\hat{\Psi}_1^*, \hat{\Gamma}_3^*)$  of the foliation of  $\hat{\Lambda}$  there are open sets  $U_j^\pm \subset \mathcal{L}^* \simeq \mathbf{T}^2$  for  $j = 1, 2, 3, 4$  such that  $\mu(U_j^\pm) > C^\pm$  where  $\mu$  is the Lebesgue measure on  $\mathbf{T}^2$ , and:*

- (1) *For all  $z \in U_1^\pm$  we have  $\mathcal{S}_1^\pm(z) > \nu^\pm$ ,  $|\partial_{\hat{\psi}_1} \mathcal{S}_1^\pm(z)| > \hat{\nu}^\pm$ ,  $\mathcal{S}_3^\pm(z) > \xi^\pm$ ,  $|\partial_{\hat{\gamma}_3} \mathcal{S}_3^\pm(z)| > \hat{\xi}^\pm$ .*
- (2) *For all  $z \in U_2^\pm$  we have  $\mathcal{S}_1^\pm(z) < -\nu^\pm$ ,  $|\partial_{\hat{\psi}_1} \mathcal{S}_1^\pm(z)| > \hat{\nu}^\pm$ ,  $\mathcal{S}_3^\pm(z) < -\xi^\pm$ ,  $|\partial_{\hat{\gamma}_3} \mathcal{S}_3^\pm(z)| > \hat{\xi}^\pm$ .*
- (3) *For all  $z \in U_3^\pm$  we have  $\mathcal{S}_1^\pm(z) > \nu^\pm$ ,  $|\partial_{\hat{\psi}_1} \mathcal{S}_1^\pm(z)| > \hat{\nu}^\pm$ ,  $\mathcal{S}_3^\pm(z) < -\xi^\pm$ ,  $|\partial_{\hat{\gamma}_3} \mathcal{S}_3^\pm(z)| > \hat{\xi}^\pm$ .*
- (4) *For all  $z \in U_4^\pm$  we have  $\mathcal{S}_1^\pm(z) < -\nu^\pm$ ,  $|\partial_{\hat{\psi}_1} \mathcal{S}_1^\pm(z)| > \hat{\nu}^\pm$ ,  $\mathcal{S}_3^\pm(z) > \xi^\pm$ ,  $|\partial_{\hat{\gamma}_3} \mathcal{S}_3^\pm(z)| > \hat{\xi}^\pm$ .*

*Proof.* — We prove part 1 of the lemma for the scattering map  $\hat{S}_+$ , as the cases for parts 2-4 of the lemma and for  $\hat{S}_-$  are analogous. Since  $\hat{\Gamma}_2 \in (\zeta_1, \zeta_2)$  (see (109)), the coefficient of each term in (92) is bounded away from zero. It follows that  $\mathcal{S}_1^+$  is a nonconstant trigonometric polynomial with zero average, and with two different harmonics; therefore we have  $\partial_{\hat{\psi}_1} \mathcal{S}_1^+ = 0$  only on a set of measure 0 in each fixed  $\mathcal{L}^*$ . Moreover, this implies that there is an open set  $V_1^+ \subset \mathbf{T}$  of positive measure in  $\mathbf{T}$  and constants  $\nu^+, \hat{\nu}^+ > 0$  such that, for each  $\hat{\psi}_1 \in V_1^+$ , we have  $\mathcal{S}_1^+(\hat{\psi}_1, \hat{\Gamma}_2) > \nu^+$  and  $|\partial_{\hat{\psi}_1} \mathcal{S}_1^+(\hat{\psi}_1, \hat{\Gamma}_2)| > \hat{\nu}^+$ .

Since  $\hat{\Gamma}_2 \in (\zeta_1, \zeta_2)$  (see (109)), the coefficients in front of the two expressions in square brackets in (93) are bounded away from zero. Recall the quantities  $v_j$  for  $j = 0, 1, 2, 3$  appearing in (93) are defined by (51) and (52). Observe that  $v_j$  cannot all be zero simultaneously. Indeed,  $v_j$  depends only on  $\delta_1, \delta_3$ , where  $\delta_1 \in (0, 1)$  and  $\delta_3 > 0$ . Therefore  $v_0 = 0$  if and only if  $\delta_1 = \delta_3$ , in which case  $v_1 \neq 0$ . Since the harmonics appearing in the second set of square brackets in (93) are different from those appearing in the first set of square brackets, it follows that  $\mathcal{S}_3^+$  is a nonconstant trigonometric polynomial with zero average for any fixed value of  $\hat{\Gamma}_2$ .

Now, suppose  $\hat{\psi}_1 \in V_1^+$ , and choose  $\hat{\gamma}_3 \in \mathbf{T}$  such that  $\mathcal{S}_3^+(\hat{\psi}_1, \hat{\gamma}_3, \hat{\Gamma}_2) > 0$ . By slightly adjusting  $\hat{\psi}_1, \hat{\gamma}_3$  if necessary, we can obtain  $\hat{\psi}'_1, \hat{\gamma}'_3, \hat{\Gamma}'_2$  where  $\hat{\psi}'_1 \in V_1^+$ , such that  $\mathcal{S}_3^+(\hat{\psi}'_1, \hat{\gamma}'_3, \hat{\Gamma}'_2) > 0$  and  $|\partial_{\hat{\gamma}_3} \mathcal{S}_3^+(\hat{\psi}'_1, \hat{\gamma}'_3, \hat{\Gamma}'_2)| > 0$ . Therefore we can find constants  $\xi^+, \hat{\xi}^+ > 0$  and an open set  $U_1^+ \subset \mathbf{T}^2$  such that part 1 of the lemma holds.  $\square$

It is clear that Lemma 37 provides us with large open sets  $U_j^\pm$  in each leaf  $\mathcal{L}$  of the foliation of  $\hat{\Lambda}$  such that the scattering map can jump a fixed distance either up or down in the  $\hat{\Psi}_1, \hat{\Gamma}_3$  directions. In the next lemma we show that the conditions satisfied by the scattering maps in these sets  $U_j^\pm$  imply transversality of the foliations in the sense of (118).

**Lemma 38.** — *Suppose  $\hat{z} \in U_j^\pm \cap \mathcal{L}_0$  and  $\hat{z}^* = \hat{S}_\pm(\hat{z}) \in \mathcal{L}_1$ , where  $U_j^\pm$  are the open sets found in Lemma 37 and  $\mathcal{L}_j$  are leaves of the foliation of  $\hat{\Lambda}$ . Then  $\hat{S}_\pm$  maps  $\mathcal{L}_0$  transversely across  $\mathcal{L}_1$  at the point  $\hat{z}^* = \hat{S}_\pm(\hat{z})$  in the sense that*

$$T_{\hat{z}^*} \hat{\Lambda} = T_{\hat{z}^*}(\hat{S}_\pm(\mathcal{L}_0)) + T_{\hat{z}^*} \mathcal{L}_1.$$

*Proof.* — We prove the lemma for  $\hat{S}_+$ , assuming  $\hat{z} \in U_1^+$  and  $\hat{z}^* = \hat{S}_+(\hat{z})$ . The set  $T_{\hat{z}^*} \hat{\Lambda}$  is a 4-dimensional vector space, and its elements can be written in the form  $v = (Q, P)$  where  $Q \in \mathbf{R}^2$  is a tangent vector in the  $\hat{\psi}_1, \hat{\gamma}_3$  directions, and  $P \in \mathbf{R}^2$  is a tangent vector in the  $\hat{\Psi}_1, \hat{\Gamma}_3$  directions. Since  $\hat{S}_+$  is smooth we have  $T_{\hat{z}^*}(\hat{S}_+(\mathcal{L}_0)) = D_{\hat{z}} \hat{S}_+(T_{\hat{z}} \mathcal{L}_0)$ . With  $\hat{z} = (\hat{\psi}_1, \hat{\gamma}_3, \hat{\Psi}_1, \hat{\Gamma}_3)$  and  $\hat{z}^* = (\hat{\psi}_1^*, \hat{\gamma}_3^*, \hat{\Psi}_1^*, \hat{\Gamma}_3^*)$ , we can write

$$D_{\hat{z}} \hat{S}_+ = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$$

where

$$\begin{aligned} A_1 &= \frac{\partial(\hat{\psi}_1^*, \hat{\gamma}_3^*)}{\partial(\hat{\psi}_1, \hat{\gamma}_3)}, & A_2 &= \frac{\partial(\hat{\psi}_1^*, \hat{\gamma}_3^*)}{\partial(\hat{\Psi}_1, \hat{\Gamma}_3)}, \\ A_3 &= \frac{\partial(\hat{\Psi}_1^*, \hat{\Gamma}_3^*)}{\partial(\hat{\psi}_1, \hat{\gamma}_3)}, & A_4 &= \frac{\partial(\hat{\Psi}_1^*, \hat{\Gamma}_3^*)}{\partial(\hat{\Psi}_1, \hat{\Gamma}_3)}. \end{aligned}$$

Therefore

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \dots, \quad A_3 = \begin{pmatrix} \frac{L_2^9}{L_3^8} \partial_{\hat{\psi}_1} \mathcal{S}_1^+(\hat{z}) & 0 \\ \frac{L_2^{11}}{L_3^8} \partial_{\hat{\psi}_1} \mathcal{S}_3^+(\hat{z}) & \frac{L_2^{11}}{L_3^8} \partial_{\hat{\gamma}_3} \mathcal{S}_3^+(\hat{z}) \end{pmatrix} + \dots$$

since  $\mathcal{S}_1^+$  does not depend on  $\hat{\gamma}_3$ .

Now, let  $v^0 \in T_{\hat{z}} \mathcal{L}_0$  and  $v^1 \in T_{\hat{z}} \mathcal{L}_1$ . Due to the topology of the leaves  $\mathcal{L}_j$  we have  $v^j = (Q^j, 0)$  for  $j = 0, 1$ . Therefore we have

$$\begin{pmatrix} \bar{Q} \\ \bar{P} \end{pmatrix} = D_{\hat{z}} \hat{S}_+(v^0) + v^1 = \begin{pmatrix} A_1 Q^0 + Q^1 \\ A_3 Q^0 \end{pmatrix} + \dots$$

Since  $\hat{z} \in U_1^+$ , the entries on the diagonal of  $A_3$  are nontrivial by part 1 of Lemma 37, and therefore it is invertible. This completes the proof of the lemma.  $\square$

**7.3. Application of the shadowing theorems.** — In this section we show that the shadowing results of [26] (which are summarised in Appendix C for convenience) can be applied to (23) and, consequently, to (22). In fact, we have already proved that the secular Hamiltonian  $F_{\text{sec}}$  satisfies the assumptions [A1-3] of Theorem 47. Indeed, in Section 7.1, we constructed a Poincaré map  $F$  to the section  $\{F_{\text{sec}} = E_0\} \cap \{\hat{\gamma}_2 = 0\}$ ; we showed that  $F$  has a normally hyperbolic invariant cylinder  $\hat{\Lambda} \simeq \mathbf{T}^2 \times [0, 1]^2$ , that the restriction  $f = F|_{\hat{\Lambda}}$  is a near-integrable twist map in the sense of Definition 44 (satisfying a non-uniform twist condition of order  $\frac{L_2^8}{L_3^8}$ ), that the stable and unstable manifolds meet transversely (with an order of splitting equal to  $\frac{1}{L_2^2}$ ), and that there are two scattering maps  $\hat{S}_{\pm} : \hat{\Lambda} \rightarrow \hat{\Lambda}'$  (see Theorem 32). This implies conditions [A1] and [A2]. In Section 7.2 we constructed a foliation of  $\hat{\Lambda}$  by leaves of the form (117), each of which is almost invariant under the map  $f$ . In Lemma 37, we showed that we can make jumps of a fixed distance of order  $\frac{L_2^{11}}{L_3^8}$  (either up or down) in each of the  $\hat{\Psi}_1, \hat{\Gamma}_3$  directions using the scattering maps; choose some such sequence of leaves  $\{\mathcal{L}_j\}_{j \in \mathbf{N}}$  such that for each  $j \in \mathbf{N}$  there is  $\beta_j \in \{+, -\}$  such that  $\mathcal{L}_{j+1} \cap \hat{S}_{\beta_j}(\mathcal{L}_j) \neq \emptyset$ . By Lemma 38, the scattering map  $\hat{S}_{\beta_j}$  maps  $\mathcal{L}_j$  transversely across  $\mathcal{L}_{j+1}$ , and the angle of transversality is of order  $\frac{L_2^{11}}{L_3^8}$ . Thus assumption [A3] of Theorem 47 is also satisfied.

Recall in Section 3 we made the change of variables (29) to the ‘tilde’ variables, which were the basis for all further analysis in the paper. This change of variables, however, is local, whereas the drift in eccentricity and inclination described in Theorem 4 is global. In order to define these coordinates we introduced constants  $\delta_j$  for  $j = 1, 2, 3$ . Here  $\delta_2$  is the coefficient of total angular momentum, and is therefore fixed for the secular system. The constants  $\delta_1, \delta_3$  on the other hand are allowed to vary, and by varying them we simply obtain a different system of ‘tilde’ coordinates. It is not hard to see that the

subsequent analysis of this paper holds equally for any value of the constants<sup>12</sup>  $\delta_1 \in (0, 1)$  and  $\delta_3 \in (-1, 1)$ . Denote by  $\tilde{C}_{\delta_1, \delta_3}$  the system of ‘tilde’ coordinates corresponding to the values  $\delta_1, \delta_3$ . Then the results of Section 4 apply in  $\tilde{C}_{\delta_1, \delta_3}$  coordinates for each value of  $\delta_1, \delta_3$ , so we have a normally hyperbolic invariant manifold  $\Lambda_{\delta_1, \delta_3}$  in each such system of coordinates. Moreover, since the cylinder depends smoothly on the parameters  $\delta_1, \delta_3$ , this construction allows us to obtain one large normally hyperbolic invariant cylinder  $\Lambda^*$ .

Observe that the contents of Sections 7.1 and 7.2 apply equally in each system of coordinates  $\tilde{C}_{\delta_1, \delta_3}$ . Furthermore, since the  $\hat{\gamma}_2$  variable does not depend on  $\delta_1, \delta_3$ , the Poincaré section is global, and so we obtain a large 4-dimensional cylinder  $\hat{\Lambda}^*$  for the return map to the Poincaré section. We now fix a global transition chain on the cylinder  $\hat{\Lambda}^*$  such that the actions  $\hat{\Psi}_1, \hat{\Gamma}_3$  drift by an amount of order  $L_2$  along the chain, and choose some sequence  $\{\tilde{C}_{\delta_1^k, \delta_3^k}\}_{k=1, \dots, K}$  so that we have an appropriate system of coordinates to apply the analysis of the earlier sections near each torus in the chain. The analysis of Sections 7.1 and 7.2 applies in each coordinate system  $\tilde{C}_{\delta_1^k, \delta_3^k}$ . Note that the shadowing results of [26] apply equally well using the many different coordinate systems, as the coordinates used in that paper are purely local. Thus the assumptions of Theorems 47 and 48 apply to the global transition chain on the cylinder  $\hat{\Lambda}^*$ .

Denote by  $\{\mathcal{L}_i\}$  the global transition chain. Then, by Theorem 47, for any  $\eta > 0$ , all sufficiently large  $L_2$ , and all  $L_3$  satisfying (21), there exists a sequence  $\{z_j\}_{j \in \mathbf{N}}$  in the secular phase space and times  $t_j > 0$  such that

$$z_{i+1} = \phi_{\text{sec}}^{t_j}(z_i), \quad d(z_i, \mathcal{L}_i) < \eta$$

for each  $j \in \mathbf{N}$  where  $\phi_{\text{sec}}^t$  is the flow associated with the secular Hamiltonian (see (31)). Observe that this completes the proof of Proposition 5. Moreover, the time to move a distance of order  $L_2$  in the  $\hat{\Psi}_1, \hat{\Gamma}_3$  directions (which is the time to move a distance of order 1 in the eccentricity  $e_2$  and the inclination  $\theta_{23}$ ) is of order

$$(119) \quad L_2 L_2^6 \frac{L_3^{16}}{L_2^{22}} \frac{L_3^6}{L_2^8} \frac{L_3^8}{L_2^{11}} = \frac{L_3^{30}}{L_2^{34}}.$$

Indeed, this follows from the following facts: the distance we drift in the  $\hat{\Psi}_1, \hat{\Gamma}_3$  variables is of order  $L_2$ ; each  $t_j$  obtained via Theorem 47 is of order  $L_2^6$  as this is the order of the return time to the Poincaré section; the order of splitting of separatrices of the cylinder is  $\frac{1}{L_2^2}$  in the  $\hat{\Gamma}_3$  direction by Proposition 27; the twist property is of order  $\frac{L_3^8}{L_2^6}$  by Theorem 32; and the order of transversality of foliations is  $\frac{L_2^{11}}{L_3^8}$  by the proof of Lemma 38. Combining these values with the formula (124) gives the time estimate (119).

Consider now the Hamiltonian  $F_{4\text{bp}}$  (see (22)) of the four body problem after averaging the mean anomalies  $\ell = (\ell_1, \ell_2, \ell_3) \in \mathbf{T}^3$ . For  $j = 1, 2, 3$  fix some  $L_j^\pm \in \mathbf{R}$  such

<sup>12</sup> More properly, one has to consider  $\delta_1, \delta_3$  in closed intervals within the open intervals to be uniformly away from the singularities of the Deprit coordinates.

that  $0 < L_j^- < L_j^+$  and if  $L = (L_1, L_2, L_3) \in [L_1^-, L_1^+] \times [L_2^-, L_2^+] \times [L_3^-, L_3^+]$  then (21) is satisfied. Write  $\Sigma = \mathbf{T}^3 \times [L_1^-, L_1^+] \times [L_2^-, L_2^+] \times [L_3^-, L_3^+]$ . Recall from the beginning of Section 7.1 the definition of the domain  $\mathcal{D}$  in the secular phase space. Denote by  $\mathcal{E}_{4bp} = \{F_{4bp}(z, \ell, L) : z \in \mathcal{D}, (\ell, L) \in \Sigma\}$  the set of values of the energy of the full four-body problem that we consider. Fix  $E_1 \in \mathcal{E}_{4bp}$ , and denote by  $\Psi$  the Poincaré map of the flow of  $F_{4bp}$  to the section  $\mathcal{D} \times \Sigma \cap \{\hat{y}_2 = 0\} \cap \{F_{4bp} = E_1\}$ . With  $z$  denoting a point in  $\mathcal{D}$ , we write  $(\bar{z}, \bar{\ell}, \bar{L}) = \Psi(z, \ell, L)$  where  $\bar{z} = G(z, \ell, L)$  and  $(\bar{\ell}, \bar{L}) = \phi(z, \ell, L)$ . Since the Hamiltonian  $F_{4bp}$  is obtained by averaging the mean anomalies, there are  $\hat{k}_1, \hat{k}_2 \in \mathbf{N}$  such that the variables  $\ell_j$  do not appear in  $F_{4bp}$  until terms of order  $\epsilon^{\hat{k}_1} \mu^{\hat{k}_2}$  where  $\epsilon = \frac{1}{L_2}$  and  $\mu = \frac{L_2}{L_3}$ , and where we can choose  $\hat{k}_j$  to be as large as we like. It follows that the map  $G$  can be written in the form  $G(z, \ell, L) = \tilde{G}(z; L) + O(\epsilon^{\hat{k}_1-6} \mu^{\hat{k}_2})$  where the higher-order terms are uniformly bounded in the  $C^r$  topology (for any  $r \in \mathbf{N}$ ), and where for any fixed values of  $L$ , the map  $z \mapsto \tilde{G}(z; L)$  is a Poincaré map  $F$  of the type constructed above in Theorem 32, and therefore satisfies assumptions [A1-3] of Theorem 47. This in turn implies that the map  $\Psi$  satisfies assumption [B1] of Theorem 48. Moreover, writing  $\phi(z, \ell, L) = (\phi_1(z, \ell, L), \phi_2(z, \ell, L))$  such that  $\bar{\ell} = \phi_1(z, \ell, L)$  and  $\bar{L} = \phi_2(z, \ell, L)$ , we have  $\phi_2(z, \ell, L) = L + O(\epsilon^{\hat{k}_1-6} \mu^{\hat{k}_2})$  where the higher-order terms are uniformly bounded in the  $C^r$  topology. Therefore assumption [B2] is also satisfied. As explained in Appendix C, as a consequence of results of [32], the map  $\Psi$  has a normally hyperbolic (locally) invariant manifold  $\tilde{\Lambda}$  that is close to  $\hat{\Lambda} \times \Sigma$ . We can thus use the coordinates  $w, \ell, L$  on  $\tilde{\Lambda}$  where  $w$  are the coordinates on  $\hat{\Lambda}$ , and define a foliation of  $\tilde{\Lambda}$  by leaves

$$\tilde{\mathcal{L}}(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, L^*) = \{(w, \ell, L) \in \tilde{\Lambda} : w \in \mathcal{L}(\hat{\Psi}_1^*, \hat{\Gamma}_3^*), L = L^*\}$$

where  $\mathcal{L}(\hat{\Psi}_1^*, \hat{\Gamma}_3^*)$  is the leaf of the foliation of  $\hat{\Lambda}$  defined by (117). Fix  $\eta > 0$ ,  $K_1, K_2 \in \mathbf{N}$ , and choose some initial values  $L_*^1 = (L_1^1, L_2^1, L_3^1)$  of the  $L_j$  variables such that  $(\ell, L_*^1) \in \text{Int } \Sigma$  for any  $\ell \in \mathbf{T}^3$ . Choose  $N \leq \epsilon^{-K_1} \mu^{-K_2}$ , and values  $P_*^1, \dots, P_*^N$  of the actions  $\hat{\Psi}_1, \hat{\Gamma}_3$  such that the leaves  $\mathcal{L}_j = \mathcal{L}(P_*^j)$  of the foliation of  $\hat{\Lambda}$  are connected by the scattering maps (corresponding to the secular Hamiltonian with  $L = L_*^1$ ) in the sense of Lemma 37. Then by Theorem 48 there are  $L_*^2, \dots, L_*^N \in [L_1^-, L_1^+] \times [L_2^-, L_2^+] \times [L_3^-, L_3^+]$  such that, with  $\tilde{\mathcal{L}}_j = \tilde{\mathcal{L}}(P_*^j, L_*^j)$ , there are points  $(z^1, \ell^1, L^1), \dots, (z^N, \ell^N, L^N)$  in the phase space of the full four body problem and times  $t_j^* > 0$  such that

$$(z^{i+1}, \ell^{i+1}, L^{i+1}) = \phi_{4bp}^{t_j^*}(z^i, \ell^i, L^i), \quad d((z^i, \ell^i, L^i), \tilde{\mathcal{L}}_i) < \eta$$

where  $\phi_{4bp}^{t_j^*}$  is the flow of the Hamiltonian  $F_{4bp}$  of the full four body problem. Moreover, the time estimate (119) holds also in this case as a lower bound on the time to move a distance of order 1 in the inclination  $\theta_{23}$  and the eccentricity  $e_2$ .



## 8. Continuation in the planetary regime

We would like to discuss the robustness of our instability mechanism with respect to modifications of the four masses and the three semimajor axes. (There are other parameters, e.g. the eccentricities of the planets, but these do not play a similar role.) Due to the invariance of physical laws by change of mass and length units, without loss of generality we may set  $m_0 = a_1 = 1$  and are reduced to a 5 dimensional parameter space.

Up to now we have investigated the moderate hierarchical regime,

$$(120) \quad O(1) = a_1 \ll a_2 \rightarrow \infty \quad \text{and} \quad a_2^{\frac{11}{6}} \ll a_3 \ll a_2^2$$

(‘moderate’ meaning that while  $a_3$  is much smaller than  $a_2$ , it is not allowed to be arbitrarily small). Weakening the hypothesis (120) and considering the general hierarchical regime  $a_1 \ll a_2 \ll a_3$  would require another proof since the fast dynamics of planet 3 could be slower than the secular dynamics of planets 1 and 2, which would completely destroy the frequency hierarchy we have extensively used, for averaging and so on.

We have also assumed that  $m_0 \neq m_1$  and  $m_0 + m_1 \neq m_2$ , because  $m_1 - m_0$  and  $m_0 + m_1 - m_2$  are in factor of the two quadrupolar Hamiltonians respectively. Rather than fixing the masses, we may as well let  $(m_0, m_1, m_2, m_3)$  vary arbitrarily in some fixed compact subset of  $\mathbf{K} \subset \mathbf{M}_\eta$  (see (2)) and the conclusions our main theorems will hold, all our construction being uniform with respect to  $\mathbf{K}$ .

Another important regime in the parameter space is the so-called planetary regime, where

$$m_1, m_2, m_3 \rightarrow 0$$

while, in turn, semimajor axes are fixed (or vary in some compact subset) [40].

**Proposition 39.** — *The instability mechanism which we have shown to exist in the hierarchical regime in Sections 3-7 continues in the planetary regime and, as  $\rho$  tends to 0, the instability time is of the order of  $\rho^{-2}$ .*

*Proof.* — Write

$$m_j = \rho \bar{m}_j, \quad j = 1, 2, 3,$$

so that, when  $\mu \ll 1$ ,

$$M_j \sim 1, \quad \sigma_{0,j} \sim 1, \quad \sigma_{ij} \sim \rho \quad \text{and} \quad \mu_j \sim \rho \quad (i, j = 1, 2, 3)$$

(notations defined in Section 2). Our proof in the hierarchical regime assumed  $\rho$  constant, and we will now show why the construction holds when  $\rho$  is small.

The key remark is the following. Consider for instance the part of the perturbing function  $F_{\text{per}}^{12}$  regarding planets 1 and 2 (first introduced in equation (13)):

$$F_{\text{per}}^{12} = \frac{\mu_2 M_2}{\|q_2\|} - \frac{m_0 m_2}{\|q_2 + \sigma_{11} q_1\|} - \frac{m_1 m_2}{\|q_2 - \sigma_{01} q_1\|}.$$

The first two terms are  $O(\rho)$ , while the third one (describing the interaction of planets 1 and 2) is  $O(\rho^2)$ ; this could impair the construction of the hierarchical regime. But the third term is qualitatively so similar to the second one, that the relative smallness of  $m_1 m_2$  will not qualitatively change the dynamics. This can be seen explicitly in the expansion in Legendre polynomials:

$$F_{\text{per}}^{12} = -\frac{\mu_1 m_2}{\|q_2\|} \sum_{n=2}^{\infty} \tilde{\sigma}_{1,n} P_n(\cos \zeta_1) \left( \frac{\|q_1\|}{\|q_2\|} \right)^n$$

where

$$\tilde{\sigma}_{1,n} = \sigma_{01}^{n-1} + (-1)^n \sigma_{11}^{n-1} \sim 1,$$

so that the hierarchy of terms inside the infinite sum is unchanged when we move away from the hierarchical regime to the planetary regime. Similar is the case of  $F_{\text{per}}^{23}$ .

Because the secular frequencies are  $\rho^2$ -small compared to the Keplerian frequencies, the splitting is  $O(\rho^2)$  in the secular directions (while it is exponentially small with respect to  $\rho$  in the Keplerian directions).  $\square$

## 9. From Deprit coordinates to elliptic elements: proof of Theorem 4

Theorem 4 and the behaviour described in (3), (4), (5), (6) are a direct consequence of Theorems 7 and 8. We first explain how to obtain the evolution of the orbital elements and then deduce Theorem 4.

To obtain the evolution of  $\ell_2$  in (3), one just has to take into account its definition in (15). For  $\theta_{23}$ , note that (21) implies  $\theta_{23} = i_{23} + O(L_2^{-1})$  where  $i_{23}$  is defined by (17). Then, the evolution of  $\theta_{23}$  can be deduced from the evolution of the Deprit variables and (17) (and taking  $L_2$  large enough). The evolution of the node, even if not stated in Theorem 7 can be easily deduced from the shadowing argument explained in Section 7.3, by iterating the Poincaré map along leaves of the almost invariant foliation of the normally hyperbolic cylinder.

The behaviour in (4), (5) can be obtained as follows. First, note that the times  $\{t_k\}$  obtained in Theorem 7 are chosen so that the corresponding points are very close to the normally hyperbolic invariant cylinder. In particular,  $\Gamma_1$  is very close to  $L_1$ . This implies that  $e_1$  is very small.

Moreover, note that the energy must be preserved. The Keplerian terms in the Hamiltonian are constant up to an arbitrarily high order. This implies that the next term in the Hamiltonian expansion (see Proposition 11), that is the quadrupolar Hamiltonian must be preserved up to small errors. If one evaluates it (see (26) in Lemma 9) at the normally hyperbolic invariant manifold, one obtains the expression

$$\frac{(1 - e_1^2) \cos^2 \theta_{12}}{(1 - e_2^2)^{3/2}} + \dots = \text{constant}.$$

(note that for the first two planets  $\theta_{12}$  coincides with  $i_{12}$ , see (16)). Since  $e_1 \sim 0$  on the cylinder, the drift in  $e_2$  implies a drift in the inclination  $\theta_{12}$ . This is precisely what is described in (4), (5).

The evolution of  $e_3$  and  $a_j$  in (6) is a direct consequence of the evolution of  $\Gamma_3$  and  $L_j$  in Theorems 7 and 8.

Finally, note that (3), (4), (5), (6) imply Theorem 4 since the inclination, the node, and the eccentricity determine the normalised angular momentum vector. Note that in Theorem 4 we allow the orbit to shadow angular momenta corresponding to degeneracies of the Deprit variables. In order to do this, it is enough to choose  $\delta$  and the ratio of the semimajor axes adequately. Clearly we can choose these ratios polynomially in  $\delta$ . Then, the time estimates in Theorems 7 and 8 imply the time estimates (7) and (8).

## Appendix A: Deprit's coordinates

This appendix is a reminder on the Deprit coordinates. These coordinates are well suited for the reduction by 4 dimensions due to the symmetry of rotations. Here we show a direct way to check that they form a symplectic coordinate system, alternative to some other presentations [22, 36]. We restrict to 3 planets, as needed in this article, but the argument extends to any number of planets (it is not an induction on the number of bodies).

Denote by

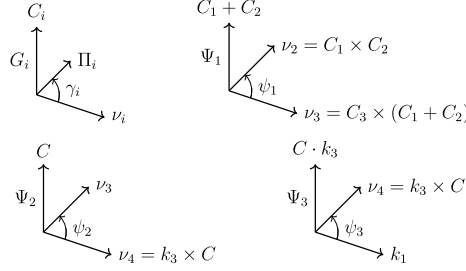
$$C_j = q_j \times p_j$$

the angular momentum of the  $j^{\text{th}}$  fictitious Keplerian body (Keplerian referring to  $F_{\text{Kep}}$ ), and let  $k_j$  be the  $j^{\text{th}}$  element of the standard orthonormal basis of  $\mathbf{R}^3$ . Define the nodes  $v_j$  by

$$v_1 = v_2 = C_1 \times C_2, \quad v_3 = (C_1 + C_2) \times C_3, \quad v_4 = k_3 \times C$$

where

$$C = C_1 + C_2 + C_3$$


 FIG. 5. — The flows of the  $G_i$ 's and the  $\Psi_i$ 's are rotations

is the total angular momentum vector. For a non-zero vector  $z \in \mathbf{R}^3$  and two non-zero vectors  $u, v$  lying in the plane orthogonal to  $z$ , denote by  $\alpha_z(u, v)$  the oriented angle between  $u, v$ , with orientation defined by the right hand rule with respect to  $z$ . Denote by  $\Pi_j$  the pericenter of  $q_j$  on its Keplerian ellipse. The Deprit variables  $(\ell_j, L_j, \gamma_j, \Gamma_j, \psi_j, \Psi_j)_{j=1,2,3}$  are defined as follows:

- $\ell_j$  is the mean anomaly of  $q_j$  on its Keplerian ellipse;
- $L_j = \mu_j \sqrt{M_j a_j}$ ;
- $\gamma_j = \alpha_{C_j}(v_j, \Pi_j)$ ;
- $\Gamma_j = \|C_j\|$ ;
- $\psi_1 = \alpha_{(C_1+C_2)}(v_3, v_2)$ ,  $\psi_2 = \alpha_C(v_4, v_3)$ ,  $\psi_3 = \alpha_{k_3}(k_1, v_4)$ ;
- $\Psi_1 = \|C_1 + C_2\|$ ,  $\Psi_2 = \|C_1 + C_2 + C_3\| = \|C\|$ ,  $\Psi_3 = C \cdot k_3$ .

The Deprit variables are analytic over the open subset  $\mathcal{D}$  over which the 3 terms of  $F_{\text{Kep}}$  are negative, the eccentricities of the Keplerian ellipses lie strictly between 0 and 1, and the nodes  $v_j$  are nonzero.

**Lemma 40.** — *The Deprit variables form a symplectic analytic coordinate system over  $\mathcal{D}$ .*

That Deprit's variables are independent, follows from their symplectic character. So, we only need to prove symplecticity.

The flow of  $L_i$  is a reparameterization of the Kepler flow of planet  $i$ , such that  $\{L_i, \ell_i\} = 1$  (as is known from the Delaunay coordinates) and the bracket of  $L_i$  with any other Deprit variable vanishes.

The following facts follow from the property that the angular momentum generates rotations, from rotational equivariance of the angular momentum and from the bare definition of the action variables (see Figure 5):

The time- $t$  map of the flow of  $G_i$  is a rotation of angle  $t$  of the ellipse  $i$  in its plane, around the center of attraction. So  $\{G_i, \gamma_i\} = 1$  and the bracket of  $G_i$  with any other variable vanishes.

$\Psi_1 = |C_1 + C_2|$  generates rotations of ellipses 1 and 2 around  $C_1 + C_2$ . So  $\{\Psi_1, \psi_1\} = 1$  and the bracket of  $\Psi_1$  with any other variable vanishes.

$\Psi_2 = |\mathbf{C}|$  generates rotations of all three ellipses around C. So  $\{\Psi_2, \psi_2\} = 1$  and the bracket of  $\Psi_2$  with any other variable vanishes.

$\Psi_3 = \mathbf{C} \cdot \mathbf{k}_3$  generates rotations of all three ellipses around the third axis. So  $\{\Psi_3, \psi_3\} = 1$  and the bracket of  $\Psi_3$  with any other variable vanishes.

It remains only to check that the brackets of pairs of angles vanish. Due to Jacobi's identity, it is sufficient to check these brackets when angles  $(\ell_j, \gamma_j, \psi_j)_{j=1,2,3}$  all vanish (cf. [39]), i.e. when planets are at their pericenters and all three pericenters  $\Pi_i$  lie on the half line generated by  $\mathbf{k}_1$ . From now on we restrict to this (allegedly Lagrangian) submanifold.

Recall that  $q_a = (q_{a,b})_{b=1,2,3}$  and  $p_a = (p_{a,b})_{b=1,2,3}$  are the position and impulsion of the  $a$ th fictitious Kepler body. Fix bodies  $j$  and  $k$ .

First,

$$\{\gamma_j, \gamma_k\} = \sum_a \left( \sum_{b=1} \left( \frac{\partial \gamma_j}{\partial q_{a,b}} \frac{\partial \gamma_k}{\partial p_{a,b}} - \frac{\partial \gamma_j}{\partial p_{a,b}} \frac{\partial \gamma_k}{\partial q_{a,b}} \right) + \sum_{b \neq 1} \left( \frac{\partial \gamma_j}{\partial q_{a,b}} \frac{\partial \gamma_k}{\partial p_{a,b}} - \frac{\partial \gamma_j}{\partial p_{a,b}} \frac{\partial \gamma_k}{\partial q_{a,b}} \right) \right) = 0;$$

indeed,

- when  $b = 1$ ,  $\frac{\partial \gamma_j}{\partial q_{a,b}} = 0$  because a variation of  $q_{a,b}$  leaves  $p_j$  orthogonal to  $q_j$  so the pericenter  $\Pi_j$  remains on the  $\mathbf{k}_1$ -axis,
- when  $b \neq 1$ ,  $\frac{\partial \gamma_j}{\partial p_{a,b}} = 0$  because after a variation of  $p_{a,b}$ ,  $p_j$  remains orthogonal to  $q_j$ ,
- $\gamma_j$  and  $\gamma_k$  play symmetric roles.

Second,

$$\{\gamma_j, \psi_k\} = \sum_a \left( \sum_{b=1} \left( \frac{\partial \gamma_j}{\partial q_{a,b}} \frac{\partial \psi_k}{\partial p_{a,b}} - \frac{\partial \gamma_j}{\partial p_{a,b}} \frac{\partial \psi_k}{\partial q_{a,b}} \right) + \sum_{b \neq 1} \left( \frac{\partial \gamma_j}{\partial q_{a,b}} \frac{\partial \psi_k}{\partial p_{a,b}} - \frac{\partial \gamma_j}{\partial p_{a,b}} \frac{\partial \psi_k}{\partial q_{a,b}} \right) \right) = 0$$

and

$$\{\psi_j, \psi_k\} = \sum_a \left( \sum_{b=1} \left( \frac{\partial \psi_j}{\partial q_{a,b}} \frac{\partial \psi_k}{\partial p_{a,b}} - \frac{\partial \psi_j}{\partial p_{a,b}} \frac{\partial \psi_k}{\partial q_{a,b}} \right) + \sum_{b \neq 1} \left( \frac{\partial \psi_j}{\partial q_{a,b}} \frac{\partial \psi_k}{\partial p_{a,b}} - \frac{\partial \psi_j}{\partial p_{a,b}} \frac{\partial \psi_k}{\partial q_{a,b}} \right) \right) = 0$$

for similar reasons.

## Appendix B: The scattering map of a normally hyperbolic invariant manifold

In this section we denote by  $M$  a  $C^r$  smooth manifold, and by  $\phi^t : M \rightarrow M$  a smooth flow with  $\frac{d}{dt}|_{t=0}\phi^t = X$  where  $X \in C^r(M, TM)$ . Let  $\Lambda \subset M$  be a compact  $\phi^t$ -invariant submanifold, possibly with boundary. By  $\phi^t$ -invariant we mean that  $X$  is tangent to  $\Lambda$ , but that orbits can escape through the boundary (a concept sometimes referred to as *local* invariance).

**Definition 41.** — *We call  $\Lambda$  a normally hyperbolic invariant manifold for  $\phi^t$  if there is  $0 < \lambda < \mu^{-1}$ , a positive constant  $C$  and an invariant splitting of the tangent bundle*

$$T_\Lambda M = T\Lambda \oplus E^s \oplus E^u$$

*such that:*

$$\begin{aligned} \|D\phi^t|_{E^s}\| &\leq C\lambda^t \quad \text{for all } t \geq 0, \\ \|D\phi^t|_{E^u}\| &\leq C\lambda^{-t} \quad \text{for all } t \leq 0, \\ \|D\phi^t|_{T\Lambda}\| &\leq C\mu^{|t|} \quad \text{for all } t \in \mathbf{R}. \end{aligned}$$

*Moreover,  $\Lambda$  is called an  $r$ -normally hyperbolic invariant manifold if it is  $C^r$  smooth, and*

$$(121) \quad 0 < \lambda < \mu^{-r} < 1$$

*for  $r \geq 1$ . This is called a large spectral gap condition.*

This definition guarantees the existence of stable and unstable invariant manifolds  $W^{s,u}(\Lambda) \subset M$  defined as follows. The local stable manifold  $W_{\text{loc}}^s(\Lambda)$  is the set of points in a small neighbourhood of  $\Lambda$  whose forward orbits never leave the neighbourhood, and tend exponentially to  $\Lambda$ . The local unstable manifold  $W_{\text{loc}}^u(\Lambda)$  is the set of points in the neighbourhood whose backward orbits stay in the neighbourhood and tend exponentially to  $\Lambda$ . We then define

$$W^s(\Lambda) = \bigcup_{t \geq 0} \phi^{-t}(W_{\text{loc}}^s(\Lambda)), \quad W^u(\Lambda) = \bigcup_{t \geq 0} \phi^t(W_{\text{loc}}^u(\Lambda)).$$

On the stable and unstable manifolds we have the strong stable and strong unstable foliations, the leaves of which we denote by  $W^{s,u}(x)$  for  $x \in \Lambda$ . For each  $x \in \Lambda$ , the leaf  $W^s(x)$  of the strong stable foliation is tangent at  $x$  to  $E_x^s$ , and the leaf  $W^u(x)$  of the strong unstable foliation is tangent at  $x$  to  $E_x^u$ . Moreover the foliations are invariant in the sense that  $\phi^t(W^s(x)) = W^s(\phi^t(x))$  and  $\phi^t(W^u(x)) = W^u(\phi^t(x))$  for each  $x \in \Lambda$  and  $t \in \mathbf{R}$ . We thus define the *holonomy maps*  $\pi^{s,u} : W^{s,u}(\Lambda) \rightarrow \Lambda$  to be projections along leaves of the strong stable and strong unstable foliations. That is to say, if  $x \in W^s(\Lambda)$  then there is a unique

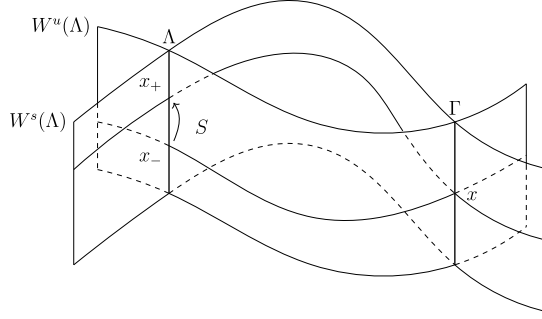


FIG. 6. — The scattering map  $S$  takes a point  $x_- \in \Lambda$ , follows the unique leaf of the strong unstable foliation passing through  $x_-$  to the point  $x$  in the homoclinic channel  $\Gamma$ , and from there follows the unique leaf of the strong stable foliation passing through  $x$  to the point  $x_+$  on  $\Lambda$

$x_+ \in \Lambda$  such that  $x \in W^s(x_+)$ , and so  $\pi^s(x) = x_+$ . Similarly, if  $x \in W^u(\Lambda)$  then there is a unique  $x_- \in \Lambda$  such that  $x \in W^u(x_-)$ , in which case  $\pi^u(x) = x_-$ .

Now, suppose that  $x \in (W^s(\Lambda) \cap W^u(\Lambda)) \setminus \Lambda$  is a transverse homoclinic point such that  $x \in W^s(x_+) \cap W^u(x_-)$ . We say that the homoclinic intersection at  $x$  is *strongly transverse* if

$$(122) \quad \begin{aligned} T_x W^s(x_+) \oplus T_x (W^s(\Lambda) \cap W^u(\Lambda)) &= T_x W^s(\Lambda), \\ T_x W^u(x_-) \oplus T_x (W^s(\Lambda) \cap W^u(\Lambda)) &= T_x W^u(\Lambda). \end{aligned}$$

In this case we can take a sufficiently small neighbourhood  $\Gamma$  of  $x$  in  $W^s(\Lambda) \cap W^u(\Lambda)$  so that (122) holds at each point of  $\Gamma$ , and the restrictions to  $\Gamma$  of the holonomy maps are bijections onto their images. We call  $\Gamma$  a *homoclinic channel* (see Figure 6). We can then define the scattering map as follows [32].

**Definition 42.** — Let  $y_- \in \pi^u(\Gamma)$ , let  $y = (\pi^u|_\Gamma)^{-1}(y_-)$ , and let  $y_+ = \pi^s(y)$ . The scattering map  $S : \pi^u(\Gamma) \rightarrow \pi^s(\Gamma)$  is defined by

$$S = \pi^s \circ (\pi^u)^{-1} : y_- \mapsto y_+.$$

Suppose now that the smoothness  $r$  of  $M$  and  $X$  is at least 2, suppose the normally hyperbolic invariant manifold  $\Lambda$  is a  $C^r$  submanifold of  $M$ , and suppose the large spectral gap condition (121) holds. This implies  $C^{r-1}$  regularity of the strong stable and strong unstable foliations [60], which in turn implies that the scattering map  $S$  is  $C^{r-1}$  [32].

In general, the scattering map may be defined only locally, as the transverse homoclinic intersection of stable and unstable manifolds can be very complicated; however in the setting of the present paper, the scattering maps we study turn out to be globally defined.

### Appendix C: A general shadowing argument

We follow the notation and exposition of [26]. Let  $M$  be a  $C^r$  manifold of dimension  $d = 2(m + n)$  where  $r \geq 4$ . Let  $F \in \text{Diff}^4(M)$ , and assume  $F$  depends smoothly on a small parameter  $\epsilon$ , with uniformly bounded derivatives. Suppose  $F$  has a normally hyperbolic invariant (or locally invariant) manifold  $\Lambda \subset M$  of dimension  $2n$  satisfying the large spectral gap condition (121); suppose moreover that  $\Lambda$  is diffeomorphic to  $\mathbf{T}^n \times [0, 1]^n$ . Furthermore, we assume that  $\dim W^s(\Lambda) = \dim W^u(\Lambda) = m + 2n$ . In order to state the remaining assumptions and the shadowing theorems, we must consider some definitions.

Suppose the scattering map  $S$  is defined relative to a homoclinic channel  $\Gamma$  for all sufficiently small  $\epsilon > 0$ . We allow for the possibility that the angle between  $W^{s,u}(\Lambda)$  along the homoclinic channel  $\Gamma$  goes to 0 as  $\epsilon \rightarrow 0$ . Denote by  $\alpha(v_1, v_2)$  the angle between two vectors  $v_1, v_2$  in the direction that yields the smallest result (i.e.  $\alpha(v_1, v_2) \in [0, \pi]$ ). For  $x \in \Gamma$ , let

$$\alpha_\Gamma(x) = \inf \alpha(v_+, v_-)$$

where the infimum is over all  $v_+ \in T_x W^s(\Lambda)^\perp$  and  $v_- \in T_x W^u(\Lambda)^\perp$  such that  $\|v_\pm\| = 1$ .

**Definition 43.** — For  $\sigma \geq 0$ , we say that the angle of the splitting along  $\Gamma$  is of order  $\epsilon^\sigma$  if there is a positive constant  $C$  (independent of  $\epsilon$ ) such that

$$\alpha_\Gamma(x) \geq C\epsilon^\sigma$$

for all  $x \in \Gamma$ .

Recall we have assumed that the normally hyperbolic invariant manifold  $\Lambda$  is diffeomorphic to  $\mathbf{T}^n \times [0, 1]^n$ , and denote by  $(q, p) \in \mathbf{T}^n \times [0, 1]^n$  smooth coordinates on  $\Lambda$ . Define  $f := F|_\Lambda$ , which also depends on the small parameter  $\epsilon$ .

**Definition 44.** — We say that  $f : \Lambda \rightarrow \Lambda$  is a near-integrable twist map if there is some  $k \in \mathbf{N}$  such that

$$f : \begin{cases} \bar{q} = q + g(p) + O(\epsilon^k) \\ \bar{p} = p + O(\epsilon^k) \end{cases}$$

where

$$\det Dg(p) \neq 0$$

for all  $p \in [0, 1]^n$ , and where the higher order terms are uniformly bounded in the  $C^1$  topology. If the higher order terms are 0 then  $f$  is an integrable twist map.



It follows from the definition that if  $f : \Lambda \rightarrow \Lambda$  is a near-integrable twist map, then there exist twist parameters  $T_+ > \tilde{T}_- > 0$  such that

$$\tilde{T}_- \|v\| \leq \|Dg(p)v\| \leq T_+ \|v\|$$

for all  $p \in [0, 1]^n$  and all  $v \in \mathbf{R}^n$ . We can always choose  $T_+$  to be independent of  $\epsilon$ . Our formulation of the problem allows the parameter  $\tilde{T}_-$  to depend on  $\epsilon$ : there is  $\tau \in \mathbf{N}_0$  and a strictly positive constant  $T_-$  (independent of  $\epsilon$ ) such that  $\tilde{T}_- = \epsilon^\tau T_-$ .

**Definition 45.** — Suppose  $f : \Lambda \rightarrow \Lambda$  is a near-integrable twist map. Denote by  $T_+ > \tilde{T}_- = \epsilon^\tau T_- > 0$  the twist parameters. We say that  $f$  satisfies:

- A uniform twist condition if  $\tau = 0$ ;
- A non-uniform twist condition (of order  $\epsilon^\tau$ ) if  $\tau > 0$ , and the order  $\epsilon^k$  of the error terms in the definition of the near-integrable twist map  $f$  is such that  $k > \tau$ .

In the coordinates  $(q, p)$ , we may define a foliation of  $\Lambda$ , the leaves of which are given by

$$(123) \quad \mathcal{L}(p^*) = \{(q, p) \in \Lambda : p = p^*\}.$$

If  $f : \Lambda \rightarrow \Lambda$  is a near-integrable twist map in the sense of Definition 44, then each leaf of the foliation is almost invariant under  $f$ , up to terms of order  $\epsilon^k$ . Denote by  $U \subset \Lambda$  the domain of definition of the scattering map  $S$ . We use the following notation: if  $M_1, M_2$  are submanifolds of  $M$  then  $M_1 \pitchfork M_2$  is the set of points  $x \in M_1 \cap M_2$  such that  $T_x M_1 \oplus T_x M_2 = T_x M$ .

**Definition 46.** — We say that the scattering map  $S$  is transverse to leaves along leaves, and that the angle of transversality is of order  $\epsilon^\nu$  (with respect to the leaves (123) of the foliation of  $\Lambda$ ) if there are  $c, C > 0$  such that for all  $p_0^* \in [0, 1]^n$  and all  $p^* \in [0, 1]^n$  satisfying  $\|p^* - p_0^*\| < c\epsilon^\nu$  we have

$$S(\mathcal{L}(p_0^*) \cap U) \pitchfork \mathcal{L}(p^*) \neq \emptyset$$

and there is  $x \in S(\mathcal{L}(p_0^*) \cap U) \pitchfork \mathcal{L}(p^*)$  such that

$$\inf \alpha(v_0, v) \geq C\epsilon^\nu$$

where the infimum is taken over all  $v_0 \in T_x S(\mathcal{L}(p_0^*) \cap U)$  and  $v \in T_x \mathcal{L}(p^*)$  such that  $\|v_0\| = \|v\| = 1$ .

Using these definitions, we may now state the main assumptions of the first shadowing theorem, which will be applied to the secular Hamiltonian (23) to prove the existence of drifting orbits for the secular Hamiltonian defined by (23).

- [A1] The stable and unstable manifolds  $W^{s,u}(\Lambda)$  have a strongly transverse homoclinic intersection along a homoclinic channel  $\Gamma$ , and so we have an open set  $U \subseteq \Lambda$  and a scattering map  $S : U \rightarrow \Lambda$ . The angle of the splitting along  $\Gamma$  is of order  $\epsilon^\sigma$ .
- [A2] The inner map  $f = F|_\Lambda$  is a near-integrable twist map with error terms of order  $\epsilon^k$  satisfying a non-uniform (or uniform) twist condition of order  $\epsilon^\tau$ .
- [A3] The scattering map  $S$  is transverse to leaves along leaves (with respect to the leaves (123) of the foliation of  $\Lambda$ ), and the angle of transversality is of order  $\epsilon^\nu$ .

**Theorem 47.** — Fix  $\eta > 0$ , let  $\epsilon > 0$  be sufficiently small, and suppose  $k \geq 2(\rho + \tau) + 1$  where  $\rho = \max\{2\sigma, 2\nu, \tau\}$ . Choose  $\{\rho_j^*\}_{j=1}^\infty \subset [0, 1]^n$  such that

$$S(\mathcal{L}_j \cap U) \cap \mathcal{L}_{j+1} \neq \emptyset,$$

and  $S(\mathcal{L}_j \cap U)$  is transverse to  $\mathcal{L}_{j+1}$ , where  $\mathcal{L}_j = \mathcal{L}(p_j^*)$ . Suppose the distance between  $\mathcal{L}_j$  and  $\mathcal{L}_{j+1}$  is of order  $\epsilon^\nu$  for each  $j$ . Then there are  $\{z_i\}_{i=1}^\infty \subset M$  and  $n_i \in \mathbf{N}$  such that  $z_{i+1} = F^{n_i}(z_i)$  and

$$d(z_i, \mathcal{L}_i) < \eta.$$

Moreover, the time to move a distance of order 1 in the  $p$ -direction is bounded from above by a term of order

$$(124) \quad \epsilon^{-\rho-\tau-\nu}.$$

Observe that Theorem 47 cannot be applied to (22). Indeed, a crucial assumption in Theorem 47 is that the scattering map  $S$  is transverse to leaves along leaves. For (22), we have no information about the behaviour of the scattering map in the  $L_i$  directions, and so we cannot check assumption [A3] for the Hamiltonian (22). Theorem 48 below generalises Theorem 47 to settings where transversality is only known in some directions, and thus allows us to complete the proof of Theorem 7.

To state Theorem 48 we consider, as before, a  $C^r$  manifold  $M$  of dimension  $2(m + n)$  where  $r \geq 4$  and  $m, n \in \mathbf{N}$ . Let  $\Sigma = \mathbf{T}^{\ell_1} \times [0, 1]^{\ell_2}$  for some  $\ell_1, \ell_2 \in \mathbf{N}_0$ , and denote by  $(\theta, \xi) \in \mathbf{T}^{\ell_1} \times [0, 1]^{\ell_2}$  coordinates on  $\Sigma$ . Write  $\tilde{M} = M \times \Sigma$ . Suppose  $\Psi \in \text{Diff}^4(\tilde{M})$  such that

$$\Psi(z, \theta, \xi) = (G(z, \theta, \xi), \phi(z, \theta, \xi))$$

where  $z \in M$ ,  $G \in C^4(\tilde{M}, M)$ , and  $\phi \in C^4(\tilde{M}, \Sigma)$ . Suppose  $\Psi$  depends on a small parameter  $\epsilon$ . We make the following assumptions on  $\Psi$ .

- [B1] There is some  $L \in \mathbf{N}$  such that

$$G(z, \theta, \xi) = \tilde{G}(z; \xi) + O(\epsilon^L)$$

where the higher order terms are uniformly bounded in the  $C^4$  topology, and for each  $\xi \in [0, 1]^{\ell_2}$  the map

$$\tilde{G}(\cdot; \xi) : z \in M \mapsto \tilde{G}(z; \xi) \in M$$

satisfies the assumptions [A1-3] of Theorem 47.

[B2] Moreover, the map  $\phi$  has the form

$$\phi : \begin{cases} \bar{\theta} = \phi_1(z, \theta, \xi) \\ \bar{\xi} = \phi_2(z, \theta, \xi) = \xi + O(\epsilon^L) \end{cases}$$

where the higher order terms are uniformly bounded in the  $C^4$  topology.

Results from [32] imply that  $\Psi$  has a normally hyperbolic invariant manifold  $\tilde{\Lambda}$  that is  $O(\epsilon^L)$  close in the  $C^4$  topology to  $\Lambda \times \Sigma$  where  $\Lambda \subset M$  is the normally hyperbolic invariant manifold of  $\tilde{G}(\cdot; \xi)$ . Moreover there is an open set  $\tilde{U} \subset \tilde{\Lambda}$  and a scattering map  $\tilde{S} : \tilde{U} \rightarrow \tilde{\Lambda}$  such that the  $z$ -component of  $\tilde{S}(z, \theta, \xi)$  is  $O(\epsilon^L)$  close in the  $C^3$  topology to  $S(z; \xi)$  where  $S(\cdot; \xi) : U \rightarrow \Lambda$  is the scattering map corresponding to  $\tilde{G}(\cdot; \xi)$ .

We use the coordinates  $(q, p, \theta, \xi)$  on  $\tilde{\Lambda}$  where  $(q, p)$  are the coordinates on  $\Lambda$  and  $(\theta, \xi)$  are the coordinates on  $\Sigma$ . Notice that the sets

$$\tilde{\mathcal{L}}(p^*, \xi^*) = \{(q, p, \theta, \xi) \in \tilde{\Lambda} : p = p^*, \xi = \xi^*\} = \mathcal{L}(p^*) \times \mathbf{T}^{\ell_1} \times \{\xi^*\}$$

for  $p^* \in [0, 1]^n$  and  $\xi^* \in [0, 1]^{\ell_2}$  define the leaves of a foliation of  $\tilde{\Lambda}$ , where  $\mathcal{L}(p^*)$  are the leaves of the foliation of  $\Lambda$  defined by (123).

**Theorem 48.** — Fix  $\eta > 0$  and  $K \in \mathbf{N}$  and let  $\epsilon > 0$  be sufficiently small. Choose  $N \in \mathbf{N}$  satisfying

$$N \leq \frac{1}{\epsilon^K}$$

$\xi_1^* \in \text{Int}([0, 1]^{\ell_2})$  so that  $\tilde{G}(\cdot; \xi_1^*)$  satisfies assumptions [A1-3], and  $p_1^*, \dots, p_N^* \in [0, 1]^n$  as in Theorem 47 such that

$$S(\mathcal{L}_j \cap U; \xi_1^*) \cap \mathcal{L}_{j+1} \neq \emptyset$$

and  $S(\mathcal{L}_j \cap U; \xi_1^*)$  is transverse to  $\mathcal{L}_{j+1}$ , where  $\mathcal{L}_j = \mathcal{L}(p_j^*)$ . Suppose the distance between  $\mathcal{L}_j$  and  $\mathcal{L}_{j+1}$  is of order  $\epsilon^\nu$  for each  $j$ , and  $L > 0$  is sufficiently large, depending on  $K$ . Then there are  $\xi_2^*, \dots, \xi_N^* \in [0, 1]^{\ell_2}$  such that, with  $\tilde{\mathcal{L}}_j = \tilde{\mathcal{L}}(p_j^*, \xi_j^*)$ , there are  $w_1, \dots, w_N \in \tilde{M}$  and  $n_i \in \mathbf{N}$  such that the  $\xi$  component of  $w_1$  is  $\xi_1^*$ ,

$$w_{i+1} = \Psi^{n_i}(w_i),$$

and

$$d(w_i, \tilde{\mathcal{L}}_i) < \eta$$

where  $\rho, \sigma, \tau$  are as in the statement of Theorem 47. Moreover, the time to move a distance of order 1 in the  $p$ -direction is of order  $\epsilon^{-\rho-\tau-\nu}$ .

Note that the transition chain obtained in Theorem 48 is only of finite length, while the one obtained in Theorem 47 may be infinite.

## Appendix D: Computation of the phase shift in $\tilde{\psi}_1$

Recall the secular Hamiltonian (see (31)) can be written

$$F_{\text{sec}} = C_{12} F_{\text{quad}}^{12} + O\left(\frac{1}{L_2^8}\right) \quad \text{where } C_{12} = -\frac{\mu_1 m_2}{(2\pi)^2}.$$

Since  $F_{\text{sec}} \sim L_2^{-6}$  (Proposition 11), we scale time to have an order one Hamiltonian,  $\tilde{F}_{\text{sec}} = L_2^6 F_{\text{sec}}$ . From now on in this section, we use this tilde notation to denote the scaled (secular, quadrupolar, octupolar etc.) Hamiltonians.

Denote by  $\Phi^t$  the flow of  $\tilde{F}_{\text{sec}}$  and let us consider points  $z^\pm$  in  $\Lambda$  such  $z^+ = S(z^-)$ . Then, there exists a point  $z^*$  in the homoclinic channel such that

$$|\Phi^t(z^*) - \Phi^t(z^\pm)| \lesssim e^{-\nu|t|} \quad \text{as } t \rightarrow \pm\infty$$

for some  $\nu > 0$  independent of  $L_2$ . The change in the  $\tilde{\psi}_1$  component from  $z_-$  to  $z_+$  is

$$(125) \quad \begin{aligned} \frac{1}{L_2^2} \Delta(\tilde{\psi}_1) &= \tilde{\psi}_1^+ - \tilde{\psi}_1^- = \int_0^{+\infty} (\partial_{\tilde{\psi}_1} \tilde{F}_{\text{sec}}(\Phi^t(z^*)) - \partial_{\tilde{\psi}_1} \tilde{F}_{\text{sec}}(\Phi^t(z^+))) dt \\ &\quad + \int_{-\infty}^0 (\partial_{\tilde{\psi}_1} \tilde{F}_{\text{sec}}(\Phi^t(z^*)) - \partial_{\tilde{\psi}_1} \tilde{F}_{\text{sec}}(\Phi^t(z^-))) dt \end{aligned}$$

The phase shift is the first order of this integral. Since  $\partial_{\tilde{\psi}_1} \tilde{F}_{\text{sec}}$  is of order  $L_2^{-1}$ , it may be expected that the phase shift will also be of order  $L_2^{-1}$ ; however, we will see below that the first order term does not contribute to the integral (125), and so we have to go the second order.

Now, notice that, since the first order of  $\tilde{F}_{\text{oct}}^{12}$  does not depend on  $\tilde{\Psi}_1$  (Lemma 14),

$$|\partial_{\tilde{\psi}_1} \tilde{F}_{\text{sec}} - C_{12} \partial_{\tilde{\psi}_1} \tilde{F}_{\text{quad}}^{12}| \lesssim \frac{1}{L_2^3}$$

and therefore, to compute  $\Delta(\tilde{\psi}_1)$ , one can replace  $F_{\text{sec}}^{12}$  by  $C_{12} \tilde{F}_{\text{quad}}^{12}$  in the integrals (125). Analogously, one can replace the flow  $\Phi^t$  by the flow  $\Phi_{\text{quad}}^t$  of  $C_{12} \tilde{F}_{\text{quad}}^{12}$ . That is, one can

conclude that

$$\begin{aligned} \frac{1}{L_2^2} \Delta(\tilde{\psi}_1) &= C_{12} \int_0^{+\infty} (\partial_{\tilde{\Psi}_1} \tilde{F}_{\text{quad}}^{12}(\Phi^t(z^*)) - \partial_{\tilde{\Psi}_1} \tilde{F}_{\text{quad}}^{12}(\Phi_{\text{quad}}^t(z^+))) dt \\ &\quad + C_{12} \int_{-\infty}^0 (\partial_{\tilde{\Psi}_1} \tilde{F}_{\text{quad}}^{12}(\Phi_{\text{quad}}^t(z^*)) - \partial_{\tilde{\Psi}_1} \tilde{F}_{\text{quad}}^{12}(\Phi_{\text{quad}}^t(z^-))) dt \\ &\quad + O(L_2^{-3}). \end{aligned}$$

Now, recall that  $\tilde{F}_{\text{quad}}^{12}$  is an integrable Hamiltonian which has a saddle with a homoclinic connection (in Poincaré variables). This homoclinic connection is  $L_2^{-1}$  close to that of  $H_0^{12}$  given in Lemma 17. We denote by  $z_{\text{quad}}^*$  and  $z_0^*$  the saddles of  $H_0^{12}$  and  $F_{\text{quad}}^{12}$  respectively and we denote by  $z_{\text{quad}}^h(t)$  and  $z_0^h(t)$  their homoclinic connections to the saddle. Therefore we can rewrite the phase shift as

$$(126) \quad \frac{1}{L_2^2} \Delta(\tilde{\psi}_1) = C_{12} \int_{-\infty}^{+\infty} (\partial_{\tilde{\Psi}_1} \tilde{F}_{\text{quad}}^{12}(z_{\text{quad}}^h(t)) - \partial_{\tilde{\Psi}_1} \tilde{F}_{\text{quad}}^{12}(z_{\text{quad}}^*)) dt + O\left(\frac{1}{L_2^3}\right).$$

Finally, it remains only to take advantage of the particular form of  $\tilde{F}_{\text{quad}}^{12}$  (see Lemma 13) to compute a first order for this integral. Write  $\hat{F}_{\text{quad}}^{12} = \frac{L_2^6}{\alpha_0^{12}} (-\tilde{c}_0^{12} + \tilde{F}_{\text{quad}}^{12})$ . Then we know from Lemma 13 that

$$\begin{aligned} \hat{F}_{\text{quad}}^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2, \tilde{\Psi}_1) &= H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) + L_2^{-1} \beta_1 H_1^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2, \tilde{\Psi}_1) \\ &\quad + L_2^{-2} \beta_2 \tilde{H}_2^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2, \tilde{\Psi}_1) + O(L_2^{-3}) \end{aligned}$$

where  $\beta_j = \alpha_j^{12} / \alpha_0^{12}$ .

Write

$$\begin{aligned} A(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) &= -4\tilde{\Gamma}_2 H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) + 3\tilde{\Gamma}_2 - \frac{\Gamma_1^2 \tilde{\Gamma}_2}{L_1^2} \\ B(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) &= \left(6 - 8H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) - 2\frac{\Gamma_1^2}{L_1^2}\right) \tilde{\Gamma}_2. \end{aligned}$$

It follows that

$$\begin{aligned} \partial_{\tilde{\Psi}_1} \hat{F}_{\text{quad}}^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2, \tilde{\Psi}_1) &= L_2^{-1} \beta_1 (3H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) - 1) \\ &\quad + L_2^{-2} \beta_2 [2\tilde{\Psi}_1 (3H_0^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) - 1) \\ &\quad + B(\gamma_1, \Gamma_1, \tilde{\Gamma}_2)] + O(L_2^{-3}). \end{aligned}$$

Notice that  $\frac{\beta_1^2}{\beta_2} = \frac{1}{2}$ . Using these formulas one can write  $\partial_{\tilde{\psi}_1} \hat{\mathbf{F}}_{\text{quad}}^{12}$  as

$$\begin{aligned} \partial_{\tilde{\psi}_1} \hat{\mathbf{F}}_{\text{quad}}^{12}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2, \tilde{\psi}_1) &= \mathbf{L}_2^{-1} \beta_1 (3\hat{\mathbf{F}}_{\text{quad}}^{12}(\gamma_1, \Gamma_1, \Gamma_2) - 1) \\ &\quad + \mathbf{L}_2^{-2} \beta_2 \left[ 2(3\hat{\mathbf{F}}_{\text{quad}}^{12} - 1)\tilde{\psi}_1 + \mathbf{B}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) - 3\frac{\beta_1^2}{\beta_2} \mathbf{H}_1^{12} \right] + \mathbf{O}(\mathbf{L}_2^{-3}) \\ &= \mathbf{L}_2^{-1} \beta_1 (3\hat{\mathbf{F}}_{\text{quad}}^{12}(\gamma_1, \Gamma_1, \Gamma_2) - 1) \\ &\quad + \mathbf{L}_2^{-2} \beta_2 \left[ \frac{1}{2}(3\hat{\mathbf{F}}_{\text{quad}}^{12} - 1)\tilde{\psi}_1 + \mathbf{B}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) - \frac{3}{2}\mathbf{A}(\gamma_1, \Gamma_1, \tilde{\Gamma}_2) \right] \\ &\quad + \mathbf{O}(\mathbf{L}_2^{-3}). \end{aligned}$$

Now, since  $\hat{\mathbf{F}}_{\text{quad}}^{12}$  is constant along its solutions,

$$\hat{\mathbf{F}}_{\text{quad}}^{12}(z_{\text{quad}}^h(t)) = \hat{\mathbf{F}}_{\text{quad}}^{12}(z_{\text{quad}}^*) \quad \forall t \in \mathbf{R}$$

and therefore the integral (126) can be rewritten as

$$\begin{aligned} \Delta(\tilde{\psi}_1) &= \mathbf{C}_{12} \beta_2 \left[ \int_{-\infty}^{+\infty} [\mathbf{B}(z_{\text{quad}}^h(t)) - \mathbf{B}(z_{\text{quad}}^*)] dt \right. \\ &\quad \left. - \frac{3}{2} \int_{-\infty}^{+\infty} [\mathbf{A}(z_{\text{quad}}^h(t)) - \mathbf{A}(z_{\text{quad}}^*)] dt \right] + \mathbf{O}(\mathbf{L}_2^{-1}). \end{aligned}$$

Finally, using that the separatrix and saddle of  $\hat{\mathbf{F}}_{\text{quad}}^{12}$  and  $\mathbf{H}_0^{12}$  are  $\mathbf{L}_2^{-1}$ -close one can conclude that

$$\begin{aligned} \Delta(\tilde{\psi}_1) &= \mathbf{C}_{12} \beta_2 \left[ \int_{-\infty}^{+\infty} [\mathbf{B}(z_0^h(t)) - \mathbf{B}(z_0^*)] dt \right. \\ &\quad \left. - \frac{3}{2} \int_{-\infty}^{+\infty} [\mathbf{A}(z_0^h(t)) - \mathbf{A}(z_0^*)] dt \right] + \mathbf{O}(\mathbf{L}_2^{-1}). \end{aligned}$$

where  $z_0^h$  is the separatrix analysed in Lemma 17 and  $z_0^*$  the saddle to which it is asymptotic.<sup>13</sup> Using the fact that  $\mathbf{H}_0^{12}(z_0^h(t)) = \mathbf{H}_0^{12}(z_0^*(t))$ , and that  $\tilde{\Gamma}_2$  is constant with respect to  $\mathbf{H}_0^{12}$ , we see that

$$\Delta(\tilde{\psi}_1) = -\frac{1}{2} \frac{\tilde{\Gamma}_2}{\mathbf{L}_1^2} \mathbf{C}_{12} \beta_2 \int_{-\infty}^{+\infty} (\Gamma_1(t)^2 - \mathbf{L}_1^2) dt$$

<sup>13</sup> In Deprit coordinates it is asymptotic to two different saddles in the past and future due to the blow up of circular motions. However the values of A and B on the two saddles are the same and therefore we abuse notation and we identify them.

where the value  $\Gamma_1(t)$  of  $\Gamma_1$  on the separatrix is given by (59). It follows that

$$\begin{aligned}\Delta(\tilde{\psi}_1) &= -\frac{1}{2A_2} \frac{\tilde{\Gamma}_2}{L_1^2} C_{12} \beta_2 \int_{-\infty}^{+\infty} \frac{\frac{5}{3}\tilde{\Gamma}_2 - L_1^2}{\cosh^2 \tau} d\tau \\ &= \frac{L_1}{6} \sqrt{\frac{3}{2}} C_{12} \beta_2 \tilde{\Gamma}_2 \sqrt{1 - \frac{5}{3} \frac{\tilde{\Gamma}_2^2}{L_1^2}}\end{aligned}$$

where we have used the formula (57) for  $A_2$ .

## Appendix E: Expansion of the inner Hamiltonian

In the proof of Lemma 25, we performed an expansion of the restriction  $K_2$  of  $F_{\text{quad}}^{23}$  to  $\Lambda$ , after averaging over all of the angles. The terms defined in (90) are as follows.

$$\begin{aligned}\bar{K}_{0,0} &= -\frac{(18\delta_1^2 - 30)\delta_3^2 - 6\delta_1^4 + 10\delta_1^2}{3\delta_1^2\delta_2^3} \\ \bar{K}_{0,1} &= \frac{(12\tilde{\Gamma}_2\delta_1^2 - 20\tilde{\Psi}_1)\delta_3^2 + (-12\tilde{\Gamma}_3\delta_1^3 + 20\tilde{\Gamma}_3\delta_1)\delta_3 + (4\tilde{\Psi}_1 - 4\tilde{\Gamma}_2)\delta_1^4}{\delta_1^3\delta_2^3} \\ \bar{K}_{0,2} &= -\frac{1}{\delta_1^4\delta_2^3} \left[ ((12\tilde{\Gamma}_2\tilde{\Psi}_1 - 9L_1^2 + 15\tilde{\Gamma}_2^2)\delta_1^2 \right. \\ &\quad \left. - 30\tilde{\Psi}_1^2 + 15L_1^2 - 15\tilde{\Gamma}_2^2)\delta_3^2 \right. \\ &\quad \left. + (-24\tilde{\Gamma}_2\tilde{\Gamma}_3\delta_1^3 + 40\tilde{\Gamma}_3\tilde{\Psi}_1\delta_1)\delta_3 \right. \\ &\quad \left. + (-2\tilde{\Psi}_1^2 + 4\tilde{\Gamma}_2\tilde{\Psi}_1 + 3L_1^2 + 6\tilde{\Gamma}_3^2 - 5\tilde{\Gamma}_2^2)\delta_1^4 \right. \\ &\quad \left. + (-5L_1^2 - 10\tilde{\Gamma}_3^2 + 5\tilde{\Gamma}_2^2)\delta_1^2 \right] \\ \bar{K}_{1,0} &= -\frac{(24\delta_1^2 - 40)\delta_3^3 + (-12\delta_1^4 + 20\delta_1^2)\delta_3}{\delta_1^2\delta_2^4} \\ \bar{K}_{1,1} &= \frac{1}{\delta_1^3\delta_2^4} \left( (48\tilde{\Gamma}_2\delta_1^2 - 80\tilde{\Psi}_1)\delta_3^3 + (-72\tilde{\Gamma}_3\delta_1^3 + 120\tilde{\Gamma}_3\delta_1)\delta_3^2 \right. \\ &\quad \left. + (24\tilde{\Psi}_1 - 24\tilde{\Gamma}_2)\delta_1^4\delta_3 + 12\tilde{\Gamma}_3\delta_1^5 - 20\tilde{\Gamma}_3\delta_1^3 \right) \\ \bar{K}_{1,2} &= -\frac{1}{\delta_1^4\delta_2^4}\end{aligned}$$

$$\begin{aligned}
 & \times \left[ \begin{aligned} & ((48 \tilde{\Gamma}_2 \tilde{\Psi}_1 - 36 L_1^2 + 60 \tilde{\Gamma}_2^2) \delta_1^2 - 120 \tilde{\Psi}_1^2 + 60 L_1^2 \\ & - 60 \tilde{\Gamma}_2^2) \delta_3^3 \\ & + (-144 \tilde{\Gamma}_2 \tilde{\Gamma}_3 \delta_1^3 + 240 \tilde{\Gamma}_3 \tilde{\Psi}_1 \delta_1) \delta_3^2 \\ & + ((-12 \tilde{\Psi}_1^2 + 24 \tilde{\Gamma}_2 \tilde{\Psi}_1 + 18 L_1^2 + 72 \tilde{\Gamma}_3^2 - 30 \tilde{\Gamma}_2^2) \delta_1^4 \\ & + (-30 L_1^2 - 120 \tilde{\Gamma}_3^2 + 30 \tilde{\Gamma}_2^2) \delta_1^2) \delta_3 \\ & + (-24 \tilde{\Gamma}_3 \tilde{\Psi}_1 + 24 \tilde{\Gamma}_2 \tilde{\Gamma}_3) \delta_1^5 \end{aligned} \right] \\
 \bar{K}_{2,0} = & - \frac{(123 \delta_1^2 - 205) \delta_3^4 + (-78 \delta_1^4 + 130 \delta_1^2) \delta_3^2 + 3 \delta_1^6 - 5 \delta_1^4}{2 \delta_1^2 \delta_2^5} \\
 \bar{K}_{2,1} = & \frac{1}{\delta_1^3 \delta_2^5} \left[ (123 \tilde{\Gamma}_2 \delta_1^2 - 205 \tilde{\Psi}_1) \delta_3^4 + (-246 \tilde{\Gamma}_3 \delta_1^3 + 410 \tilde{\Gamma}_3 \delta_1) \delta_3^3 \right. \\
 & + (78 \tilde{\Psi}_1 - 78 \tilde{\Gamma}_2) \delta_1^4 \delta_3^2 \\
 & \left. + (78 \tilde{\Gamma}_3 \delta_1^5 - 130 \tilde{\Gamma}_3 \delta_1^3) \delta_3 + (-6 \tilde{\Psi}_1 + 3 \tilde{\Gamma}_2) \delta_1^6 + 5 \tilde{\Psi}_1 \delta_1^4 \right] \\
 \bar{K}_{2,2} = & - \frac{1}{4 \delta_1^4 \delta_2^5} \\
 & \times \left[ \begin{aligned} & ((492 \tilde{\Gamma}_2 \tilde{\Psi}_1 - 369 L_1^2 + 615 \tilde{\Gamma}_2^2) \delta_1^2 \\ & - 1230 \tilde{\Psi}_1^2 + 615 L_1^2 - 615 \tilde{\Gamma}_2^2) \delta_3^4 \\ & + (-1968 \tilde{\Gamma}_2 \tilde{\Gamma}_3 \delta_1^3 + 3280 \tilde{\Gamma}_3 \tilde{\Psi}_1 \delta_1) \delta_3^3 \\ & + ((-156 \tilde{\Psi}_1^2 + 312 \tilde{\Gamma}_2 \tilde{\Psi}_1 + 234 L_1^2 \\ & + 1476 \tilde{\Gamma}_3^2 - 390 \tilde{\Gamma}_2^2) \delta_1^4 \\ & + (-390 L_1^2 - 2460 \tilde{\Gamma}_3^2 + 390 \tilde{\Gamma}_2^2) \delta_1^2) \delta_3^2 \\ & + (-624 \tilde{\Gamma}_3 \tilde{\Psi}_1 + 624 \tilde{\Gamma}_2 \tilde{\Gamma}_3) \delta_1^5 \delta_3 \\ & + (36 \tilde{\Psi}_1^2 - 36 \tilde{\Gamma}_2 \tilde{\Psi}_1 - 9 L_1^2 - 156 \tilde{\Gamma}_3^2 + 15 \tilde{\Gamma}_2^2) \delta_1^6 \\ & + (-10 \tilde{\Psi}_1^2 + 15 L_1^2 + 260 \tilde{\Gamma}_3^2 - 15 \tilde{\Gamma}_2^2) \delta_1^4 \end{aligned} \right].
 \end{aligned}$$

## Appendix F: Corrigendum of [41]

As this paper uses several ideas and formulas from [41], we include here corrections to some errata in that paper.

- (1) In Lemma 2.1, equation (12), the sign of the last term  $\frac{\Gamma^2}{L_1^2}$  should be +, and so the Hamiltonian  $H_0$  should look the same as the Hamiltonian  $H_0^{12}$  defined in equation (33) of the present paper.
- (2) In Lemma 3.1, equation (26), the sign should be +, and so that lemma is equivalent to Lemma 17 of the present paper.
- (3) In Lemma 5.1, the sign of  $\sin \gamma^2$  should be +.
- (4) Lemma 5.2 should be as follows:



*Lemma.* — The function  $\mathcal{F}^+$  can be written, on the separatrix, as a function of  $g_1$  as  $\mathcal{F}^+ = \frac{1}{2}(\mathcal{F}_1 + i\mathcal{F}_2)$  with

$$\begin{cases} \mathcal{F}_1 = C_1 \frac{\sqrt{1-\frac{5}{3}(1+\chi^2)\cos^2 g_1}}{1-\frac{5}{3}\cos^2 g_1} \cos g_1 \\ \mathcal{F}_2 = C_2 \frac{\sqrt{1-\frac{5}{3}(1+\chi^2)\cos^2 g_1}}{1-\frac{5}{3}\cos^2 g_1} \left(-\frac{21}{5} + \frac{1}{3} \frac{\Gamma^2}{L_1^2} \frac{15-13\cos^2 g_1}{1-\frac{5}{3}\cos^2 g_1}\right) \end{cases}$$

where

$$C_1 = \frac{20}{3} \sqrt{\frac{2}{3}} \frac{\Gamma^3}{L_1^3 \chi} A_{\text{oct}}, \quad C_2 = \frac{\sqrt{10}}{3} \frac{\Gamma}{L_1 \chi} A_{\text{oct}},$$

$$A_{\text{oct}} = -\frac{15}{64} \frac{a_1^3}{a_2^4} \frac{e_2}{(1-e_2^2)^{\frac{5}{2}}}.$$

- (5) As a consequence, there are many cancellations in the Poincaré-Melnikov computation in Section 5 of [41], the complex integrand has only one singularity, and the integral  $\mathcal{L}(\gamma^0)$  defined in equation (34) takes the simple form

$$\mathcal{L}(\gamma^0) = \sqrt{\frac{3}{2}} \frac{\pi A_{\text{oct}} e^{\frac{\pi \Gamma}{\Lambda_2 L_1^2}}}{6 \sqrt{15} L_1 (1 + e^{\frac{2\pi \Gamma}{\Lambda_2 L_1^2}})} (24 L_1^2 - 37 \Gamma^2) \sin \gamma^0.$$

Note that the notation from [41] is  $\Gamma = \tilde{\Gamma}_2$ , and moreover the function in front of  $\sin \gamma^0$ , called  $\tilde{\mathcal{L}}_2^{12}$  in Proposition 27 of the present paper, does not vanish under condition (55).

## Acknowledgements

The authors warmly thank the referees for their reading and suggestions.

## Funding

A. Clarke and M. Guardia are supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 757802). M. Guardia is also supported by the Catalan Institution for Research and Advanced Studies via an ICREA Academia Prize 2018 & 2023. This work is part of the grant PID-2021-122954NB-I00 funded by MCIN/AEI/10.13039/501100011033 and “ERDF A way of making Europe”. This work is also supported by the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M).

This work is also partially supported by the project of the French Agence Nationale pour la Recherche CoSyDy (ANR-CE40-0014).

## Declarations:

## Competing Interests

The authors declare no competing interests.

## Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## REFERENCES

1. A. ALBOUY, Histoire des équations de la mécanique analytique: repères chronologiques et difficultés, in *Siméon-Denis Poisson*, Hist. Math. Sci. Phys., pp. 229–280, Ed. Éc. Polytech, Palaiseau, 2013.
2. V. M. ALEKSEEV, Quasirandom dynamical systems. I, II, III. Math. USSR, **5**, **6**, **7**, 1968–1969.
3. V. I. ARNOLD, Small denominators and problems of stability of motion in classical and celestial mechanics, *Usp. Mat. Nauk*, **18** (1963), 91–192.
4. V. I. ARNOLD, Instability of dynamical systems with many degrees of freedom, *Dokl. Akad. Nauk SSSR*, **156** (1964), 9–12.
5. V. I. ARNOLD, *Mathematical Methods of Classical Mechanics*, vol. 60, Springer, Berlin, 1978.
6. K. BATYGIN, A. MORBIDELLI and M. J. HOLMAN, Chaotic disintegration of the inner solar system, *Astrophys. J.*, **799** (2015), 120.
7. P. BERNARD, The dynamics of pseudographs in convex Hamiltonian systems, *J. Am. Math. Soc.*, **21** (2008), 615–669.
8. P. BERNARD, V. KALOSHIN and K. ZHANG, Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders, *Acta Math.*, **217** (2016), 1–79.
9. S. BOLOTIN, Symbolic dynamics of almost collision orbits and skew products of symplectic maps, *Nonlinearity*, **19** (2006), 2041–2063.
10. S. BOLOTIN and D. TRESCHÉV, Unbounded growth of energy in nonautonomous Hamiltonian systems, *Nonlinearity*, **12** (1999), 365–388.
11. G. BOUÉ and J. LASKAR, A collisionless scenario for Uranus tilting, *Astrophys. J. Lett.*, **712** (2010), L44.
12. G. BOUÉ, J. LASKAR and F. FARAGO, A simple model of the chaotic eccentricity of Mercury, *Astron. Astrophys.* (2012).
13. M. J. CAPIŃSKI, Computer assisted existence proofs of Lyapunov orbits at  $L_2$  and transversal intersections of invariant manifolds in the Jupiter-Sun PCR3BP, *SIAM J. Appl. Dyn. Syst.*, **11** (2012), 1723–1753.

14. M. J. CAPIŃSKI and M. GIDEA, Arnold diffusion, quantitative estimates, and stochastic behavior in the three-body problem, *Commun. Pure Appl. Math.*, **76** (2023), 616–681.
15. M. J. CAPIŃSKI and P. ZGLICZYŃSKI, Cone conditions and covering relations for topologically normally hyperbolic invariant manifolds, *Discrete Contin. Dyn. Syst.*, **30** (2011), 641–670.
16. M. J. CAPIŃSKI, M. GIDEA and R. de la LLAVE, Arnold diffusion in the planar elliptic restricted three-body problem: mechanism and numerical verification, *Nonlinearity*, **30** (2017), 329–360.
17. J. CHAZY, Sur l’allure du mouvement dans le problème des trois corps quand le temps croît indéfiniment, *Ann. Sci. Éc. Norm. Supér.*, **3** (1922), 29–130.
18. C. CHENG, Dynamics around the double resonance, *Camb. J. Math.*, **5** (2017), 153–228.
19. C. CHENG and J. YAN, Existence of diffusion orbits in a priori unstable Hamiltonian systems, *J. Differ. Geom.*, **67** (2004), 457–517.
20. T. CHENY, A. CLARKE and M. GUARDIA, *Arnold diffusion in the planetary 5-body problem*, preprint.
21. L. CHIERCHIA and G. GALLAVOTTI, Drift and diffusion in phase space, *Ann. Inst. Henri Poincaré A, Phys. Théor.*, **60** (1994), 144.
22. L. CHIERCHIA and G. PINZARI, Deprit’s reduction of the nodes revisited, *Celest. Mech. Dyn. Astron.*, **109** (2011), 285–301.
23. L. CHIERCHIA and G. PINZARI, The planetary N-body problem: symplectic foliation, reductions and invariant tori, *Invent. Math.*, **186** (2011), 1–77.
24. B. V. CHIRIKOV, The passage of a nonlinear oscillating system through resonance, *Sov. Phys. Dokl.*, **4** (1959), 390–394.
25. A. CLARKE and D. TURAEV, Arnold diffusion in multidimensional convex billiards, *Duke Math. J.*, **172** (2023), 1813–1878.
26. A. CLARKE, J. FEJOZ and M. GUARDIA, Topological shadowing methods in Arnold diffusion: weak torsion and multiple time scales, *Nonlinearity*, **36** (2023), 426–457.
27. A. CLARKE, J. FEJOZ and M. GUARDIA, A counterexample to the theorem of Laplace–Lagrange on the stability of semimajor axes, *Arch. Ration. Mech. Anal.*, **248** (2024), 19.
28. A. DELSHAMS and G. HUGUET, Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems, *Nonlinearity*, **22** (2009), 1997–2077.
29. A. DELSHAMS, R. de la LLAVE and T. SEARA, A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by a potential of generic geodesic flows of  $\mathbf{T}^2$ , *Commun. Math. Phys.*, **209** (2000), 353–392.
30. A. DELSHAMS, R. de la LLAVE and T. SEARA, Orbits of unbounded energy in quasi-periodic perturbations of geodesic flows, *Adv. Math.*, **202** (2006), 64–188.
31. A. DELSHAMS, R. de la LLAVE and T. M. SEARA, 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** (2006), viii+141.
32. A. DELSHAMS, R. De La LLAVE and T. M. SEARA, Geometric properties of the scattering map of a normally hyperbolic invariant manifold, *Adv. Math.*, **217** (2008), 1096–1153.
33. A. DELSHAMS, M. GIDEA and P. ROLDAN, Transition map and shadowing lemma for normally hyperbolic invariant manifolds, *Discrete Contin. Dyn. Syst.*, **33** (2013), 1089–1112.
34. A. DELSHAMS, R. de la LLAVE and T. M. SEARA, Instability of high dimensional Hamiltonian systems: multiple resonances do not impede diffusion, *Adv. Math.*, **294** (2016), 689–755.
35. A. DELSHAMS, V. KALOSHIN, A. de la ROSA and T. M. SEARA, Global instability in the restricted planar elliptic three body problem, *Commun. Math. Phys.*, **366** (2019), 1173–1228.
36. A. DEPRIT, Elimination of the nodes in problems of  $n$  bodies, *Celest. Mech.*, **30** (1983), 181–195.
37. J. FEJOZ, Quasiperiodic motions in the planar three-body problem, *J. Differ. Equ.*, **183** (2002), 303–341.
38. J. FÉJOZ, Démonstration du ‘théorème d’Arnold’ sur la stabilité du système planétaire (d’après Herman), *Ergod. Theory Dyn. Syst.*, **24** (2004), 1521–1582.
39. J. FEJOZ, On action-angle coordinates and the Poincaré coordinates, *Regul. Chaotic Dyn.*, **18** (2013), 703–718.
40. J. FÉJOZ, Celestial Mechanics, chapter the N-body problem. Encyclopedia of life support systems, *Unesco-EOLSS* (2015).
41. J. FEJOZ and M. GUARDIA, Secular instability in the three-body problem, *Arch. Ration. Mech. Anal.*, **221** (2016), 335–362.
42. J. FEJOZ and M. GUARDIA, A remark on the onset of resonance overlap, *Regul. Chaotic Dyn.*, **28** (2023), 578–584.
43. J. FÉJOZ, M. GUARDIA, V. KALOSHIN and P. ROLDAN, Kirkwood gaps and diffusion along mean motion resonances in the restricted planar three-body problem, *J. Eur. Math. Soc.*, **18** (2016), 2315–2403.
44. N. FENICHEL, Persistence and smoothness of invariant manifolds for flows, *Indiana Univ. Math. J.*, **21** (1971), 193–226.

45. N. FENICHEL, Asymptotic stability with rate conditions, *Indiana Univ. Math. J.*, **23** (1974), 1109–1137.
46. N. FENICHEL, Asymptotic stability with rate conditions, ii, *Indiana Univ. Math. J.*, **26** (1977), 81–93.
47. E. FONTICH and P. MARTÍN, Arnold diffusion in perturbations of analytic exact symplectic maps, *Nonlinear Anal.*, **42** (2000), 1397–1412.
48. E. FONTICH and P. MARTÍN, Arnold diffusion in perturbations of analytic integrable Hamiltonian systems, *Discrete Contin. Dyn. Syst.*, **7** (2001), 61–84.
49. C. FROESCHLE and H. SCHOLL, The three principal secular resonances nu(5), nu(6), and nu(16) in the asteroidal belt, *Celest. Mech. Dyn. Astron.*, **46** (1989), 231–251.
50. V. GELFREICH and D. TURAEV, Unbounded energy growth in Hamiltonian systems with a slowly varying parameter, *Commun. Math. Phys.*, **283** (2008), 769–794.
51. V. GELFREICH and D. TURAEV, Arnold diffusion in a priori chaotic symplectic maps, *Commun. Math. Phys.*, **353** (2017), 507–547.
52. M. GIDEA and R. de la LLAVE, Topological methods in the instability problem of Hamiltonian systems, *Discrete Contin. Dyn. Syst.*, **14** (2006), 295.
53. M. GIDEA, R. de la LLAVE and T. M. SEARA, A general mechanism of diffusion in Hamiltonian systems: qualitative results, *Commun. Pure Appl. Math.*, **73** (2020), 150–209.
54. M. GUARDIA, P. MARTÍN and T. M. SEARA, Homoclinic solutions to infinity and oscillatory motions in the restricted planar circular three body problem, in *Progress and Challenges in Dynamical Systems*, Springer Proc. Math. Stat., vol. 54, pp. 265–280, Springer, Heidelberg, 2013.
55. M. GUARDIA, P. MARTÍN and T. M. SEARA, Oscillatory motions for the restricted planar circular three body problem, *Invent. Math.*, **203** (2016), 417–492.
56. M. GUARDIA, J. PARADELA, T. M. SEARA and C. VIDAL, Symbolic dynamics in the restricted elliptic isosceles three body problem, *J. Differ. Equ.*, **294** (2021), 143–177.
57. M. GUARDIA, P. MARTÍN, J. PARADELA and T. M. SEARA, Hyperbolic dynamics and oscillatory motions in the 3 body problem, preprint (2022), <https://arxiv.org/abs/2207.14351>.
58. R. S. HARRINGTON, Dynamical evolution of triple stars, *Astron. J.* (1968), 190–194.
59. M. HERMAN, Some open problems in dynamical systems, in *Proceedings of the International Congress of Mathematicians (Berlin, 1998)*, volume Extra Vol. II, pp. 797–808 (electronic), 1998.
60. M. W. HIRSCH, C. C. PUGH and M. SHUB, Invariant manifolds, *Bull. Am. Math. Soc.*, **76** (1970).
61. V. I. ARNOLD, *Hygens and Barrow, Newton and Hooke*, Birkhäuser, Basel, 1990, Pioneers in mathematical analysis and catastrophe theory from evolvents to quasicrystals, Translated from the Russian by Eric J. F. Primrose..
62. T. ITO and K. OHTSUKA, The Lidov-Kozai oscillation and Hugo von Zeipel, *Monogr. Environ. Earth Planets*, **7** (2019), 1–113.
63. W. H. JEFFERYS and J. MOSER, Quasi-periodic solutions for the three-body problem, *Astron. J.*, **71** (1966), 568–578.
64. V. KALOSHIN and K. ZHANG, *Arnold Diffusion for Smooth Systems of Two and a Half Degrees of Freedom*, Annals of Mathematics Studies., vol. 208, Princeton University Press, Princeton, 2020.
65. Y. KOZAI, Secular perturbations of asteroids with high inclination and eccentricity, *Astron. J.* (1962).
66. P.-S. LAPLACE, Théorie de Jupiter et de Saturne. *Mém. Acad. Royale des sciences de Paris*, Œuvres complètes, Tome XI, 1785–1788.
67. J. LASKAR, A numerical experiment on the chaotic behavior of the Solar System, *Nature*, **338** (1989), 237–238.
68. J. LASKAR, *Sfogliando la Mécanique analitique*, *Giornata di studio su Louis Lagrange*, chapter Lagrange et la stabilité du système solaire, Edizioni Universitarie di Lettere Economia Diritto, Milano, 2006.
69. J. LASKAR, Chaotic diffusion in the Solar System, *Icarus*, **196** (2008), 1–15.
70. J. LASKAR, Le système solaire est-il stable? in *Le Chaos, Number XIV in Séminaire Poincaré*, pp. 221–246, Birkhäuser, Basel, 2010.
71. J. LASKAR and P. ROBUTEL, The chaotic obliquity of the planets, *Nature*, **361** (1993), 608–612.
72. L. LAZZARINI, J.-P. MARCO and D. SAUZIN, Measure and capacity of wandering domains in Gevrey near-integrable exact symplectic systems, *Mem. Am. Math. Soc.*, **257** (2019), vi+110.
73. M. LECAR, F. A. FRANKLIN, M. J. HOLMAN and N. J. MURRAY, Chaos in the solar system, *Annu. Rev. Astron. Astrophys.*, **39** (2001), 581–631.
74. M. L. LIDOV, The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies, *Planet. Space Sci.*, **9** (1962), 719–759.
75. J. LLIBRE and C. SIMÓ, Oscillatory solutions in the planar restricted three-body problem, *Math. Ann.*, **248** (1980), 153–184.
76. J. LLIBRE and C. SIMÓ, Some homoclinic phenomena in the three-body problem, *J. Differ. Equ.*, **37** (1980), 444–465.

77. R. MOECKEL, Chaotic dynamics near triple collision, *Arch. Ration. Mech. Anal.*, **107** (1989), 37–69.
78. R. MOECKEL, Generic drift on Cantor sets of annuli, in *Celestial Mechanics (Evanston, IL, 1999)*, Contemp. Math., vol. 292, pp. 163–171, Am. Math. Soc., Providence, 2002.
79. R. MOECKEL, Symbolic dynamics in the planar three-body problem, *Regul. Chaotic Dyn.*, **12** (2007), 449–475.
80. F. MOGAVERO and J. LASKAR, The origin of chaos in the solar system through computer algebra, *Astron. Astrophys.*, **662** (2022), L3.
81. R. MONTGOMERY, The zero angular momentum, three-body problem: all but one solution has syzygies, *Ergod. Theory Dyn. Syst.*, **27** (2007), 1933–1946.
82. R. MONTGOMERY, Oscillating about coplanarity in the 4 body problem, *Invent. Math.*, **218** (2019), 113–144.
83. A. MORBIDELLI, *Modern Celestial Mechanics. Aspects of Solar System Dynamics*, CRC Press, Boca Raton, 2002.
84. J. MOSER, *Stable and Random Motions in Dynamical Systems. Princeton Landmarks in Mathematics*, Princeton University Press, Princeton, 2001, With special emphasis on celestial mechanics, Reprint of the 1973 original, with a foreword by Philip J. Holmes.
85. S. NAOZ, The eccentric Kozai-Lidov effect and its applications, *Annu. Rev. Astron. Astrophys.*, **54** (2016), 441–489.
86. I. NEWTON, *Philosophiae naturalis principia mathematica*, Maclehose, 1871.
87. L. NIEDERMAN, Stability over exponentially long times in the planetary problem, *Nonlinearity*, **9** (1996), 1703–1751.
88. G. PINZARI, On the Kolmogorov set for many-body problems, PhD thesis, Università degli Studi di Roma Tre, 2009.
89. H. POINCARÉ, *Les Méthodes Nouvelles de la Mécanique Céleste*, Gauthier-Villars, Paris, 1892.
90. P. ROBUTEL, Stability of the planetary three-body problem. II. KAM theory and existence of quasiperiodic motions, *Celest. Mech. Dyn. Astron.*, **62** (1995), 219–261.
91. K. SITNIKOV, The existence of oscillatory motions in the three-body problems, *Sov. Phys. Dokl.*, **5** (1960), 647–650.
92. D. TRESCHÉV, Evolution of slow variables in a priori unstable Hamiltonian systems, *Nonlinearity*, **17** (2004), 1803–1841.
93. D. TRESCHÉV, Arnold diffusion far from strong resonances in multidimensional a priori unstable Hamiltonian systems, *Nonlinearity*, **25** (2012), 2717–2757.
94. H. von ZEIPPEL, Sur l’application des séries de M. Lindstedt à l’étude du mouvement des comètes périodiques, *Astron. Nachr.*, **183** (1910).
95. J. XUE, Arnold diffusion in a restricted planar four-body problem, *Nonlinearity*, **27** (2014), 2887–2908.
96. L. ZHAO, Quasi-periodic solutions of the spatial lunar three-body problem, *Celest. Mech. Dyn. Astron.*, **119** (2014), 91–118.
97. S. L. ZIGLIN, Secular evolution of the orbit of a planet in a binary-star system, 1975.

## Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A. C.

Escola d’Enginyeria de Barcelona Est, Campus Diagonal Besòs, Edifici A (EEBE),  
 Universitat Politècnica de Catalunya,  
 Av. Eduard Maristany 16,  
 08019, Barcelona, Spain  
[andrew.michael.clarke@upc.edu](mailto:andrew.michael.clarke@upc.edu)

J. F.

Université Paris Dauphine PSL,  
 CEREMADE,  
 Place du Maréchal de Lattre de Tassigny,  
 75016, Paris, France

ANDREW CLARKE, JACQUES FEJOZ, MARCEL GUARDIA

and  
Observatoire de Paris PSL,  
IMCCE,  
77 avenue Denfert Rochereau,  
75014, Paris, France  
[jacques.fejoz@dauphine.psl.eu](mailto:jacques.fejoz@dauphine.psl.eu)

M. G.  
Departament de Matemàtiques i Informàtica,  
Universitat de Barcelona, Gran Via, 585,  
08007, Barcelona, Spain  
and  
Centre de Recerca Matemàtica,  
Edifici C, Campus Bellaterra,  
08193, Bellaterra, Spain  
[guardia@ub.edu](mailto:guardia@ub.edu)

*Manuscrit reçu le 19 octobre 2022*

*Version révisée le 4 juin 2024*

*Manuscrit accepté le 6 septembre 2024*