A counterexample to the theorem of Laplace-Lagrange on the stability of semimajor axes

Andrew Clarke   Jacques Fejoz   Marcel Guardia

March 7, 2024

Abstract

A longstanding belief has been that the semimajor axes, in the Newtonian planetary problem, are stable. In the course of the XIX century, Laplace, Lagrange and others gave stronger and stronger arguments in this direction, thus culminating in what has commonly been referred to as the first Laplace-Lagrange stability theorem. In the problem with 3 planets, we prove the existence of orbits along which the semimajor axis of the outer planet undergoes large random variations thus disproving the conclusion of the Laplace-Lagrange theorem. The time of instability varies as a negative power of the masses of the planets. The orbits we have found fall outside the scope of the theory of Nekhoroshev-Niederman because they are not confined by the conservation of angular momentum and because the Hamiltonian is not (uniformly) convex with respect to the Keplerian actions.

Contents

1 Introduction

1.1 On the stability of semimajor axes

Consider the 4-body problem, namely the motion of 4 bodies, numbered from 0 to 3, moving in the 3-dimensional space and subject to the Newtonian universal attraction:

$$\ddot{x}_j = \sum_{\substack{0 \le i < 3 \\ i \ne j}} m_i \frac{x_i - x_j}{\|x_i - x_j\|^3}, \tag{1}$$

where $x_j \in \mathbb{R}^3$ is the position and $m_j > 0$ the mass of body j. Of particular interest is the planetary problem, where the masses of bodies $1, 2, 3$ (planets) are small with respect to body 0 (Sun), and where each planet revolves around the Sun along an approximate, slowly deforming Keplerian ellipse. In the first approximation, the problem consists of three uncoupled Kepler problems whose ellipses are fixed in space, together with their geometric elements determining the shape of the ellipses and their position in space. The question is to determine the long term influence of the mutual attraction of planets on the elliptical positions and elements. In this article, we will also consider the hierarchical problem, where masses are fixed (or within some compact set of $(0, \infty)$) and successive semimajor axes' ratios a_j/a_{j+1} are small.

Euler and Lagrange had failed to prove the stability of semimajor axes of planets in the Solar System. In 1776, in a commendable *tour de force* Laplace was able to overcome the difficulties his predecessors had met. He wrote [46]:

> J'ai trouvé [que l'inégalité séculaire des demi grands axes est] absolument nulle; d'où je conclus que l'altération du mouvement moyen de Jupiter, si elle existe, n'est point due à l'action de Saturne.[1]

Here Laplace is neglecting second order terms in the masses of the planets, as well as third order terms in the eccentricities and inclinations of planets.

Lagrange later proved that this result holds for arbitrary eccentricities and inclinations. This is the "first stability theorem of Laplace and Lagrange" [3, Example 6.16]. About the 1808 Mémoire of Lagrange [45], Arago commented: "Le 17 août 1808, [Lagrange] lit au Bureau des longitudes, et le lundi suivant 22, à l'Académie des sciences, un des plus admirables Mémoires qu'ait jamais tracés la plume d'un mathématicien" (F. Arago, *Œuvres complètes*, 1854, p. 654).[2]

Poisson later proved that the conclusion of the theorem holds at the second order in the masses of the planets [54]. His proof is a cornerstone of Hamiltonian perturbation theory, but is lengthy and complicated. Lagrange simplified it substantially (see his Œuvres, t. VI, p. 735), but to the point where Lagrange's argument is flawed, as his editor M. Serret mentions. The later correction made in [44] is not satisfactory either, as Mathieu noticed [48]... (see [47, 29] and references therein).

Nowadays the first stability theorem of Laplace-Lagrange is a simple consequence of the existence of the Delaunay coordinates for the two-body problem. In these symplectic coordinates, the variable which

---

[1]In modern English: I have found that the variations of the semimajor axis of Jupiter, under the influence of Saturn, have zero average.

[2]On 17 August 1808, and on the following Monday 22, at the Académie des sciences, [Lagrange] reads one of the most magnificient memoirs ever written by a mathematician. This work was entitled: Memoir on the theory of the variations of planets' elements, and in particular of the variations of semimajor axes of their orbits.

is conjugate to the fast Keplerian angle (mean anomaly) is a function of the semimajor axis. So, outside Keplerian resonances, for the (first order) secular system obtained by averaging out the mean anomalies, semimajor axes are first integrals.

In order to explain the irregularities of Jupiter and Saturn, Laplace called on comets. Comets had unknown masses, so it was a convenient argument (which actually was a fortunate motivation for Laplace to get interested in probabilities). Yet, there is an intricate interplay between small parameters in the parameter space (masses of the planets, distance to mean motion resonances, distance to circular motions, etc.). It is a mistake to infer the stability of the semimajor axes from the low order analysis that had been carried out. First, there is an issue with the confusion between normal forms and the full vector field. Second, averaging out the outer mean anomalies becomes irrelevant when the mean motion of outer planets is slower than the secular dynamics of inner planets.

Poincaré is perfectly aware of the first issue, for example when he writes "Ce résultat aurait été envisagé par Laplace ou Lagrange comme établissant complètement la stabilité du système solaire. Nous sommes plus difficiles aujourd'hui parce que la convergence des développements n'est pas démontrée"[3] [53, Vol. 2, Chap. 10, Paragraph 132]. But he seems less aware of the second issue, when he writes "Il est clair que tout ce qui précède s'applique, sans qu'on ait rien à y changer, au cas où l'on aurait plus de trois corps"[4] (ibid.).

More recently, after the proof of Arnold's theorem on the existence of a set of positive Lebesgue measure of invariant tori in the planetary problem [1, 26], Herman has speculated that "in some respect Lagrange and Laplace, against Newton, are correct in the sense of measure theory and that in the sense of topology, the above question [on the stability], in some respect, could show Newton is correct" [41].

It is the purpose of the present article to disprove the belief in the conclusion of the Laplace-Lagrange stability of the semimajor axis, as well as Herman's conjectural dichotomy: random instability does occur on a set of positive Lebesgue measure of the 4-body problem in the planetary regime, in a time which is an inverse power of the masses of the planets. More precisely, we exhibit some dynamical behavior which is inconsistent with the averaged dynamics, in that one semimajor axis, as well as other quantities, display large, random variations. A further step would be to estimate the local probability of instability in some given time (in the light of the theory of adiabatic invariants [3, Section 6.4], or else).

1.2 Main results

Consider 4 bodies whose motion is governed by Newton's equation (1.1). We will assume that $m_0 \neq m_1$.[5] For the sake of simplicity, let us first focus on the "hierarchical regime"; it is the asymptotic regime where masses are fixed, while body 2 revolves around and far away from bodies 0 and 1, and body 3 revolves around and even farther away from bodies 0, 1 and 2. (We will make some more precise hypotheses below.) 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, 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.

Some notation: let a_1, a_2 and a_3 be the semimajor axes of the planets, e_1, e_2 and e_3 be their eccentricities, and C_1, C_2 and C_3 their angular momenta. In the hierarchical regime, for eccentricities bounded away from 1, $a_1 \ll a_2 \ll a_3$. Even further (and unlike in [13]), we will consider a *strongly hierarchical regime*, where not only the semimajor axes ratios $\alpha_i = a_i/a_{i+1}$ are small, but even the ratios of the ratios α_i/α_{i+1} are small, in the following quantitative manner:

$$a_1 = O(1) \ll a_2 \ll a_3^{1/3}. \tag{2}$$

Here is the rough description of the scales of times:

---

[3]This result would have been considered by Laplace or Lagrange as proving the stability of the solar system. We are more careful today because the convergence of expansions has not been proved.

[4]It is clear that the above proof applies, with nothing to be changed, to the case of more than three bodies

[5]If the four masses are not equal to each other, this condition is always satisfied up to renumbering the masses, i.e. up to switching the roles of bodies.

- The fastest frequencies are the mean motions (Keplerian frequencies) of the two inner planets. Since $a_1 \ll a_2$, these inner mean motions do not interfere, which allows us to average out the mean anomalies, without resonances.

- The next frequencies are the secular frequencies of the two inner planets. They govern the rotation of the plane of the ellipses around their angular momentum vector $C_1 + C_2$, and the rotation of the ellipses in their plane, as well as the quasiperiodic oscillations of the corresponding inclinations and eccentricities. The dynamics of the truncated relevant normal form ("quadrupolar dynamics" of planets 1 and 2) is still integrable, as noticed by Harrington [40], due to the fact that the quadrupolar Hamiltonian does not depend on the argument of the outer pericenter g_2.

- In the strongly hierarchical regime, the outer semimajor axis is so large that the mean motion of planet 3 is slower than secular frequencies of the two inner planets.

- Then come the secular frequencies of the (outer) planet 3, approximately determined by the quadrupolar Hamiltonian of planets 2 and 3. The conservation of the total angular momentum vector $C = C_1 + C_2 + C_3 \simeq C_3$ prevents significant changes in the plane of the outer ellipse, or of the product $a_3\sqrt{1 - e_3^2}$. On the other hand, it does not prevent major (joint) changes in a_3 or e_3.

Similarly to the regime studied in [13], along the orbits we will prove the existence of the two inner planets will be close to the hyperbolic secular singularity of the quadrupolar Hamiltonian or to the associated stable and unstable manifolds. In particular, their mutual inclination will be large.

We will pay special attention to two quantities:

- the semimajor axis a_3 of the outer planet

- the normalized angular momentum $\tilde{C}_2 \in \mathbb{B}^3$ of planet 2, defined as the vector orthogonal to the plane of its ellipse and whose norm is $\|\tilde{C}_2\| = \sqrt{1 - e_2^2}$.

Theorem 1 (Main theorem). *Consider the 4-body problem with masses $m_j > 0$, $j = 0, 1, 2, 3$ with $m_0 \neq m_1$. For every finite itinerary $\tilde{C}_2^1, ..., \tilde{C}_2^k \in \mathbb{B}^3$, $a_3^1, ..., a_3^k \in [1, +\infty)$ and every $\delta > 0$, there exists an open set of initial conditions whose trajectories realise the prescribed itinerary up to precision δ.*

This theorem is a consequence of Theorems 2 and 3 below, which contain a more detailed description of the diffusing orbits.

Let us make some comments on Theorem 1.

- The drifting time needed to follow the prescribed itinerary in the theorem satisfies

$$0 < T < C(m_0, m_1, m_2, m_3) \frac{N}{\delta^\kappa}, \tag{3}$$

where C is a constant depending only on the masses and the exponent $\kappa > 0$ does not depend on N nor on the itinerary. To be more precise, call $\alpha_i = a_i/a_{i+1}$, $i = 1, 2$, the semimajor axis ratios. As δ tends to zero, the α_i's will be chosen polynomially smaller, and the drifting time itself depends polynomially on the α_i's.

- As stated, Theorem 1 assumes small semimajor axis ratios, for fixed masses. In Section 8.1, we provide asymptotic estimates when we let the masses of the planets tend to 0, i.e. in the planetary regime where $m_j = \rho \tilde{m}_j$ for $j = 1, 2, 3$ with $\rho > 0$ small. Then, one possibility is to let the semimajor axes of planets 1 and 2 tend to 0 as $\rho \to 0$. In that case, the drifting time satisfies

$$0 < T < C(m_0, \tilde{m}_1, \tilde{m}_2, \tilde{m}_3) \frac{N}{\delta^\kappa \rho^\nu}. \tag{4}$$

Another possibility is to place planets 1 and 2 at a uniform distance (with respect to ρ) from the Sun and place planet 3 very far away, so that $a_3 \sim \rho^{-2/3}$. That is the setting considered in Theorem 3 below, where we provide the concrete exponent $\nu = 35/3$.

Note that the instability time is polynomial with respect to the masses of the planets. See Section 1.3 below for the comparison of the regime of Theorem 1 with those regimes where Nekhoroshev Theory can be applied to prove exponential stability of the semimajor axes.

- The novelty of the unstable behavior presented in this paper compared to that of [13] is the evolution of the semimajor axis a_3 of the third planet. Indeed, in the moderately hierarchical regime considered in [13], a_3 is stable, whereas in the strongly hierarchical regime it can follow any prescribed itinerary (see Section 1.4 below for a comparison between the two regimes).

  On the contrary, the changes in the normalized angular momentum of the second planet are the same in both regimes. Let us briefly describe what the changes in $\tilde{C}_2$ imply in terms of the orbital elements of the second planet. Indeed, fixing a prescribed itinerary $\tilde{C}_2^0, ..., \tilde{C}_2^N \in B^3$ is equivalent to prescribing any itinerary in: the eccentricity e_2^k, the mutual inclination θ_{23}^k between planets 2 and 3, and the longitude h_2^k of the node of planet 2, for $k = 0 \ldots N$. Then, we can construct an orbit and times $t_0 < t_1 < \cdots < t_N$ such that the osculating orbital elements satisfy

  $$|e_2(t_k) - e_2^k| \le \delta, \qquad |\theta_{23}(t_k) - \theta_{23}^k| \le \delta, \qquad |h_2(t_k) - h_2^k| \le \delta \qquad \text{for} \quad k = 0, 1, ..., N.$$

  As already mentioned, the angular momentum of the third body is almost constant and therefore, the evolution of e_3 is determined by the evolution of a_3.

  Finally, the evolution of the eccentricity e_1 of the first planet, and the mutual inclination θ_{12} between planets 1 and 2, cannot be controled since they are prescribed by the diffusion mechanism. Let us briefly mention that:

    - The eccentricity e_1 does change but it can start arbitrarily close to 0. That is, the initial configuration can have all planets performing close to circular motion.

    - The mutual inclination i_{12} always stays above 55 degrees.

  One can see [13] for a more detailed description of the evolution of e_1 and i_{12}.

- In our Solar System, semimajor axes seem very stable. There are some exceptions. Notably, the semimajor axis of the Moon is drifting. But this is due to non-Hamiltonian, tidal effects [24]. Also, at the early stages of our Solar System, planets migrated towards the exterior of the Solar System. But this migration too is a non-conservative phenomenon, explained by the interaction with the planetesimal disk [37].

  Orbits described in theorem 1 show wild variations of elliptical elements, and, plausibly, subsequent collisions of neighboring planets and their accretion. We may conjecture that only the observation of many extra-solar systems might exhibit one day such transient behavior.

1.3   Remark on Nekhoroshev theory and weak convexity

Due to the proper degeneracy of the Keplerian approximation, standard Nekhoroshev theory does not apply in a straightforward way to the planetary problem. Yet it has been successfully extended to the planetary problem [50, 51] (see also [3][6.3.4]). In particular, Niederman proved the following conditional result regarding a Hamiltonian perturbation of a properly degenerate integrable system: provided that the actions in the degenerate (i.e. secular) directions remain in some bounded region, the actions conjugate to the fast angles are stable over an exponentially long time. He then showed that this model can be applied to the planetary problem. In the neighborhood of coplanar and circular ellipses (the maximum of the angular momentum), the conservation of the angular momentum prevents the degenerate actions (encoding eccentricities and inclinations) to undergo any substantial instability, so the actions conjugate to the fast angles (encoding the semimajor axes) are indeed stable over an exponentially long time.

The regimes of the 4-body problem studied in the present paper differ from Niederman's work in two respects:

- In the planetary regime, the conservation of the angular momentum does not prevent secular variables from drifting because of the high inclination of the two inner planets.

- In the hierarchical problem, the convexity of the fast, Keplerian part is weak, because of the large outer semimajor axes.

Hence neither Nekhoroshev theory nor Niederman's adaptation applies. (Incidentally, a proof *ad absurdum* is that the conclusion of Theorem 1 would contradict Nekhoroshev theory.)

Regarding the weak convexity (for a numerical investigation, see [39]), let us mention the following open question. Consider the toy Hamiltonian

$$H(\theta, r) = r_1^2 + r_2^2 + \epsilon^\alpha r_3^2 + \epsilon f(\theta, r),$$

where $\epsilon \ll 1$ and $0 \le \alpha \le 1$. H is Nekhoroshev-stable for $\alpha = 0$ and, trivially, unstable for $\alpha = 1$. But, more precisely, how does the radius of confinement of r deteriorate as α grows from 0 to 1? The classical proof as well as more recent examples should provide a precise answer to this question.

1.4 Main ideas of the proof of Theorem 1 and moderately versus strongly hierarchical regimes

The orbits constructed in Theorem 1 rely on an Arnold diffusion mechanism [2]. Progress in the understanding of Arnold diffusion in nearly-integrable Hamiltonian systems in these last decades has been remarkable, especially for two and a half degrees of freedom (see [4, 6, 10, 16, 17, 18, 21, 35, 43, 49, 55], or [5, 9, 15, 20, 33, 34, 36, 56] for results in higher dimension). However, most of these results deal with generic nearly-integrable Hamiltonian systems in C^r or C^∞ regularity whereas results in the analytic category, including results in Celestial Mechanics, are rather scarce. See the discussion in Section 1.1 of [13] for more details.

Indeed, even if Arnold in his seminal paper conjectured that his diffusion mechanism should be present in the 3-body problem, as far as the authors know, the *only complete analytical* proofs of Arnold diffusion in celestial mechanics are [13, 22, 38]. Other works in the field rely on computer-assisted computations [8, 28], on computer-assisted proofs [7], or on the assumption of a plausible transversality hypothesis [57].

In order to prove Theorem 1 we adapt what are usually referred to as the geometric and topological methods of Arnold diffusion. Although some of the geometric ideas could now be considered classical, others are contemporary (in particular a topological shadowing result proven recently by the same authors in [14]). While explaining the overview of the proof, we will compare the moderately hierarchical regime considered in [13] with the strongly hierarchical regime considered in the present paper.

The classical geometric approach used to prove Arnold diffusion both in the present paper and in [13] can be broken down into the following steps.

- Prove the existence of a normally hyperbolic invariant cylinder (See Appendix C for the definition). The "vertical" components of the cylinder are the actions in which we want to drift.

- Prove that the invariant manifolds of the cylinder intersect transversely along homoclinic channels. Orbits in the channel are heteroclinic orbits between different orbits in the cylinder. This is encoded in a *scattering map* [19]. One can obtain an asymptotic formula for it through Poincaré-Melnikov Theory.

- Construct an iterated function system consisting of the *inner dynamics* and the scattering map, and show that its orbits (called *pseudo-orbits*) display a drift in the action variables.

- Use a shadowing argument to obtain orbits which follow closely these pseudo-orbits.

To carry out these steps in the 4 body problem, both in [13] and in the present paper we consider the hierarchical regime which makes the 4-body problem nearly integrable. In [13] we consider what we call the *moderately hierarchical* regime, where we assume

$$a_1 = O(1) \ll a_2 \ll a_3^{6/11} \ll a_2^{12/11} \tag{5}$$

whereas in the present paper we consider the strongly hierarchical regime (1.2). Both regimes lead to a nearly integrable setting. However, they lead to different hierarchies of time scales, and to different first-order effective models. Let us describe them. To this end, we express the 4-body problem in a good set of coordinates, which reduces the dimension of the model by eliminating its first integrals. First, we consider Jacobi coordinates to eliminate the translation invariance and then we use the Deprit coordinates to perform the symplectic reduction by rotational symmetry (see [12]). After this reduction, the 4-body problem becomes

a Hamiltonian system with seven degrees of freedom. In Section 2 we state the main results of this paper in Deprit coordinates.

The two faster frequencies of the 4-body problem in Deprit coordinates are, in both the strongly and moderately hierarchical regimes, the mean anomalies of the first two planets. Moreover, they evolve at different time scales to one another. This implies that they can be averaged out up to arbitrarily high order in a_2^{-1}. If one ignores the higher order terms, one can reduce the dimension by 2 and end up with a five degree of freedom Hamiltonian depending on a_1 and a_2, which can be treated as parameters.

In the moderately hierachical regime (1.4), the third fastest frequency is the mean anomaly of the third planet. Proceeding analogously, in that regime it too can be averaged out up to high order which leads to a 4 degree of freedom Hamiltonian. This Hamiltonian is usually called the secular Hamiltonian since it models the slow evolution of the osculating ellipses.

On the contrary, in the strongly hierachical regime, since the third planet is placed much further away, the third mean anomaly becomes slower and it cannot be averaged out. For this reason, in the present paper we analyse the 5 degree of freedom Hamiltonian which, by an abuse of language, we also call the secular Hamiltonian. It models the slow evolution of the osculating ellipses of the three planets plus the motion of the third planet on its osculating ellipse.

In both regimes, the next step is to expand the secular Hamiltonian in powers of $1/a_2$ and a_2/a_3 using the Legendre polynomials. This is done in Section 3. The first term in the expansion is the so-called quadrupolar Hamiltonian of the first two planets, which is integrable, and the second term is the so-called octupolar Hamiltonian which captures the next order of interaction between planets 1 and 2. The subsequent orders in the expansion involve the interaction between planets 2 and 3. It is at these orders where the analysis of the moderately and strongly hierarchical regimes differs considerably. In [13] we need both the quadrupolar and octupolar Hamiltonians associated to planets 2 and 3 whereas in the present paper the approximate dynamics does not depend on the octupolar term. The reason is that, since we do not average the mean anomaly ℓ_3, the quadrupolar term adds "more non-integrability" to the model. Indeed, the quadrupolar Hamiltonian depends on all the secular variables which was not the case in [13].

Next, we analyse the normally hyperbolic cylinder and its invariant manifolds for the secular Hamiltonian. The first appropriate approximation is that of the quadrupolar Hamiltonian of planets 1 and 2 (see Section 4). It is well known that it posesses a hyperbolic singularity, which appears when the mutual inclination between planets 1 and 2 is larger than around 40 degrees. This hyperbolic singularity corresponds to a normally hyperbolic invariant cylinder in the full phase space. Moreover, the integrability implies that its stable and unstable manifolds coincide along a homoclinic manifold.

Fenichel Theory [30, 31, 32] implies that the normally hyperbolic invariant cylinder is persistent. In Section 5, we analyse the induced dynamics of the secular Hamiltonian on the cylinder, usually called the *inner dynamics*. We prove that it is integrable up to an arbitrarily high order and that it has torsion, provided that the mutual inclination of planets 1 and 2 is larger than 55 degrees. Note that the cylinder in the present paper has two dimensions more than the cylinder considered in [13], as the mean anomaly of planet 3 and the semimajor axis a_3 provide additional inner variables.

The results in [27] combined with classical perturbation techniques imply that the stable and unstable invariant manifolds of the cylinder of the secular Hamiltonian intersect transversely along two homoclinic channels. Orbits in these channels are heteroclinic orbits which "connect" different points in the cylinder. Such connections are encoded in the scattering maps (see [19] and Appendix C for the definition). Section 6 is devoted to the computation of the first order of these maps by means of Poincaré-Melnikov Theory.

Once the inner dynamics and the outer dynamics (i.e. the scattering maps) have been analysed, the last step is to combine them to achieve drift in the actions. This is done in Section 7. First, we construct pseudo-orbits (i.e. orbits of the iterated function system consisting of a Poincaré map induced by the inner dynamics and the two scattering maps) that follow any prescribed itinerary in the actions such that the scattering maps map "approximately invariant tori" of the inner dynamics transversely across other such tori. Then, referring to an argument contained in a previous paper by the authors [14] which provides rather flexible shadowing results, we show that there are orbits of both the secular Hamiltonian and the four-body problem Hamiltonian which follow closely the pseudo-orbits. Moreover, the shadowing methods in [14] allow us also to obtain time estimates.

Finally, in Section 8.1 we explain how to deal with the planetary regime where the masses of the three planets are arbitrarily small.

2 Main results in Deprit coordinates

The first step towards a proof of Theorem 1 is to find a suitable set of coordinates in which we can analyse the problem. In particular, as is well-known, the 4-body problem has many symmetries which can be exploited to reduce the dimension of the phase space. To this end, in Section 2.1 we explain how to express the 4-body problem in Jacobi coordinates, thus reducing by translational symmetry, and then pass to Deprit coordinates to reduce by rotational symmetry. In Section 2.2 we state a more detailed version of Theorem 1 in Deprit coordinates.

2.1 The Jacobi and Deprit coordinates

The 4-body problem is a Hamiltonian system with respect to the Hamiltonian

$$H = \sum_{0 \le j \le 3} \frac{y_j^2}{2m_j} - \sum_{0 \le i < j \le 3} \frac{m_i m_j}{\|x_j - x_i\|}, \tag{6}$$

and the symplectic form $\Omega = dq \wedge dp$ where $x_j \in \mathbb{R}^3$ is the position of body j and $y_j \in \mathbb{R}^3$ its conjugate linear momentum.

The Jacobi coordinates $(q_j, p_j) \in \mathbb{R}^3 \times \mathbb{R}^3$, $j = 0, 1, 2, 3$, are defined as

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

where

$$\sigma_{ij} = \frac{m_i}{M_j} \qquad \text{and} \qquad M_j = \sum_{i=0}^{j} m_i. \tag{7}$$

A direct computation implies that this transformation is symplectic, in the sense that $dq \wedge dp = dx \wedge dy$. The Hamiltonian (2.1) expressed in these coordinates does not depend on q_0, and therefore p_0 is a first integral. Without loss of generality, we may restrict to $p_0 = 0$ and consider the reduced phase space with coordinates $(q_j, p_j)_{j=1,2,3}$. Then, the Hamiltonian (2.1) becomes

$$H = F_{\text{Kep}} + F_{\text{per}} \tag{8}$$

where

$$F_{\text{Kep}} = \sum_{j=1}^{3} \left(\frac{p_j^2}{2\mu_j} - \frac{\mu_j M_j}{\|q_j\|} \right) \tag{9}$$

$$F_{\text{per}} = \sum_{j=2}^{3} \frac{\mu_j M_j}{\|q_j\|} - \frac{m_0 m_2}{\|q_2 + \sigma_{11} q_1\|} - \frac{m_0 m_3}{\|q_3 + \sigma_{22} q_2 + \sigma_{11} q_1\|} - \frac{m_1 m_2}{\|q_2 - \sigma_{01} q_1\|}$$
$$- \frac{m_1 m_3}{\|q_3 + \sigma_{22} q_2 + (\sigma_{11} - 1) q_1\|} - \frac{m_2 m_3}{\|q_3 + (\sigma_{22} - 1) q_2\|} \tag{10}$$

with the reduced masses μ_j defined, for each $j = 1, 2, 3$, by

$$\mu_j^{-1} = M_{j-1}^{-1} + m_j^{-1}.$$

The next step is to pass to Deprit coordinates, which are well suited to the symmetry of rotations. These coordinates were discovered originally by Deprit [23], but their efficacy in the N-body problem was noticed only recently by Chierchia and Pinzari [11]. Let us denote by

$$C_j = q_j \times p_j$$

8

the angular momentum of the jth fictitious body and let

$$C = C_1 + C_2 + C_3$$

be the total angular momentum vector. Let k_j be the jth element of the standard orthonormal basis of $\mathbb{R}^3$ and define the nodes ν_j by

$$\nu_1 = \nu_2 = C_1 \times C_2, \quad \nu_3 = (C_1 + C_2) \times C_3, \quad \nu_4 = k_3 \times C.$$

For a non-zero vector $z \in \mathbb{R}^3$ 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 Π_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:

- ℓ_j is the mean anomaly of q_j on its Keplerian ellipse;

- $L_j = \mu_j \sqrt{M_j a_j}$;

- $\gamma_j = \alpha_{C_j}(\nu_j, \Pi_j)$;

- $\Gamma_j = \|C_j\|$;

- $\psi_1 = \alpha_{(C_1+C_2)}(\nu_3, \nu_2)$, $\psi_2 = \alpha_C(\nu_4, \nu_3)$, $\psi_3 = \alpha_{k_3}(k_1, \nu_4)$;

- $\Psi_1 = \|C_1 + C_2\|$, $\Psi_2 = \|C_1 + C_2 + C_3\| = \|C\|$, $\Psi_3 = C \cdot k_3$.

The Deprit variables are analytic and symplectic over the open subset $\mathcal{D}$ in which the 3 terms of F_{Kep} are negative, the eccentricities of the Keplerian ellipses belong to $(0, 1)$ and the nodes ν_j are nonzero (see [11, 23] or [13, Appendix A]). Actions Ψ_2 and Ψ_3 are commuting first integrals.

The orbital elements can be expressed in terms of Deprit cordinates:

- The osculating eccentricities are defined by

$$e_j = \sqrt{1 - \frac{\Gamma_j^2}{L_j^2}}, \qquad j = 1, 2, 3. \tag{11}$$

- The mutual inclination i_{12} between planets 1 and 2, measured as the oriented angle between C_1 and C_2, is defined via its cosine by

$$\cos i_{12} = \frac{\Psi_1^2 - \Gamma_1^2 - \Gamma_2^2}{2\,\Gamma_1\,\Gamma_2}. \tag{12}$$

- The mutual inclination i_{23} between planet 3 and the inner two planets, measured as the oriented angle between $S_1 = C_1 + C_2$ and C_3, is defined via its cosine by

$$\cos i_{23} = \frac{\Psi_2^2 - \Gamma_3^2 - \Psi_1^2}{2\Gamma_3\Psi_1}. \tag{13}$$

2.2 Arnold diffusion in Deprit coordinates

In this section we give a precise reformulation of Theorem 1 in terms of the Deprit coordinates. To this end, let us recall that we assume the masses $m_0, m_1, m_2, m_3 > 0$ are fixed and satisfy $m_0 \neq m_1$. We consider a regime of increasingly separated bodies. In terms of the semimajor axes, we assume that

$$a_1 \ll a_2 \ll a_3^{1/3} \tag{14}$$

which, in terms of Deprit, reads

$$L_1 \ll L_2 \ll L_3^{1/3}. \tag{15}$$

We assume that the eccentricities of the bodies are uniformly bounded away from 0 and 1, and therefore the other Deprit actions Γ_j, $j = 1, 2, 3$, and Ψ_j, $j = 1, 2$, have significantly different sizes. Indeed they satisfy

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

Let us explain which actions may drift under these assumptions. The semimajor axes L_1 and L_2 are almost constant due to the fact that the conjugate angles ℓ_1, ℓ_2 evolve faster than all other variables, and on different time scales to one another. As a result, ℓ_1, ℓ_2 can be averaged out of the Hamiltonian up to arbitrarily high order and any splitting of separatrices in the L_1, L_2 directions is exponentially small. Moreover, recall that Ψ_2 is the total angular momentum, which is a first integral. Since $\Gamma_1, \Gamma_2 \ll \Gamma_3$ this implies that $\Gamma_3 \sim \Psi_2$ is almost constant. However in this case, as in [13], relatively small proportions of angular momentum can transfer from Γ_3 to Γ_2 to create a significant change in the orbital elements of the second planet. Indeed, Γ_3 can drift from

$$\Gamma_3 \sim \Psi_2 - \Psi_1 \qquad \text{to} \qquad \Gamma_3 \sim \Psi_2 + \Psi_1.$$

This corresponds to having C_3 and $C_1 + C_2$ close to parallel and either with the same sign or with opposite sign. That is, this evolution in Γ_3 implies a large drift in the inclination i_{23} (see (2.1)), or equivalently in θ_{23} (as defined in Section 1.2).

Since $\Gamma_2 \ll \Gamma_3$, this transfer of angular momentum between bodies can cause dramatic changes in $\Gamma_2 \in (0, L_2)$. Indeed, we are able to show that it drifts from

$$\Gamma_2 \sim L_2 \qquad \text{to} \qquad \Gamma_2 \sim 0.$$

Equivalently, the orbital ellipse of the second planet can swing from being near-circular ($e_2 \sim 0$) to being highly eccentric ($e_2 \sim 1$).

Finally, note that L_3 must satisfy $L_3 > \Gamma_3 \sim \Psi_2$. Moreover, by taking $L_3 \to +\infty$ while Ψ_2 is fixed (recall that it is a first integral) one has that $e_3 \to 1$. Since we are considering a hierarchical regime, where the orbits of the bodies are increasingly (and uniformly) separated, one has to constrain the possible growth of L_3. If one fixes $0 < \kappa \ll 1$, then one can consider L_3 transitioning from

$$L_3 \sim (1 + \kappa)\Psi_2 \qquad \text{to} \qquad L_3 \sim \frac{\Psi_2}{\kappa}. \tag{16}$$

Since Γ_3 is almost constant, this is equivalent to changing at the same time the semimajor axis as above (i.e. transitioning from $a_3 \sim (1 + \kappa)^2 \Psi_2^2$ to $a_3 \sim \frac{\Psi_2^2}{\kappa^2}$) and the eccentricity e_3 as

$$e_3 \sim \sqrt{\kappa} \qquad \text{to} \qquad e_3 \sim \sqrt{1 - \kappa^2}.$$

The next theorem shows that such transitions are possible and that one can freely vary Γ_2, Γ_3 and L_3 within their "allowed" ranges.

Theorem 2. *Fix masses $m_0, m_1, m_2, m_3 > 0$ such that*

$$m_0 \neq m_1. \tag{17}$$

There exists ξ with $0 < \xi \ll 1$ and $\alpha_1, \alpha_2, \beta > 0$ such that the following is satisfied.

Fix $N \geq 1$ any $\{\nu^k\}_{k=0}^N \subset (0, 1)$, $\{\eta^k\}_{k=0}^N \subset (-1, 1)$, $\{\zeta^k\}_{k=0}^N \subset (1, +\infty)$ and constants L_1^0, L_2^0 and Ψ_2^0 satisfying

$$L_1^0 \in [1 - \xi, 1 + \xi], \qquad L_1^0 \ll L_2^0 \qquad (L_2^0)^3 \ll \Psi_2^0 \qquad \text{and} \qquad |\Psi_1^0 - \nu_0 L_2^0| \leq \frac{L_1^0}{\sqrt{3}} + \xi.$$

Then, there exists an orbit of the Hamiltonian H in (2.1) expressed in Deprit coordinates and times $\{t_k\}_{k=0}^N$ satisfying

$$t_0 = 0 \qquad \text{and} \qquad |t_k| \leq \left(L_2^0\right)^{\alpha_1} \left(\Psi_2^0\right)^{\alpha_2}, \, k \geq 1$$

such that
$$|\Gamma_2(t_k) - \nu_k L_2^0| \le (L_2^0)^{-\beta}$$
$$|\Psi_2^0 - \Gamma_3(t_k) - \eta_k \Psi_1(t_k)| \le (L_2^0)^{-\beta}$$
$$|L_3(t_k) - \zeta_k \Psi_2^0| \le (L_2^0)^{-\beta}.$$

Moreover,
$$|\Gamma_1(t_k) - L_1^0| \le (L_2^0)^{-\beta}, \qquad |\Psi_1(t_k) - \Gamma_2(t_k) - M_k| \le (L_2^0)^{-\beta}$$
where $M_k \in \left(0, \frac{L_1^0}{\sqrt{3}} + \xi\right)$ *is determined by*

$$\frac{M_k^2}{(L_2^0 - \Gamma_2(t_k))^{3/2}} = \frac{M_0^2}{(L_2^0 - \Gamma_2^0)^{3/2}} \qquad and \qquad M_0 = \Psi_1^0 - \Gamma_2^0$$

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

$$|\Gamma_3(t) - \Gamma_3^0| \le 2L_2^0, \qquad and \qquad |L_j(t) - L_j^0| \le (L_2^0)^{-\beta} \quad for \quad j = 1, 2.$$

Theorem 2 is proved in Sections 3-7. It pertains to fixed values of the masses, in the sense that the increasing separation of the semimajor axes depend on the mass choices. We now want to obtain an analogous statement in the planetary regime, namely when bodies 1, 2, and 3 are assumed to have small mass. More concretely, $m_0 \sim 1$ and $m_i = \rho \tilde{m}_i$ for $i = 1, 2, 3$ with $\tilde{m}_i \sim 1$ and $0 < \rho \ll 1$.

To deal with the planetary regime we consider scaled Deprit coordinates. Indeed, for fixed semimajor axes, the Deprit actions all have size ρ. Then, to be able to capture their drift along the diffusing orbits, it is convenient to perform the conformally symplectic scaling

$$L = \rho \check{L}, \quad \Gamma = \rho \check{\Gamma}, \quad \Psi = \rho \check{\Psi}, \tag{18}$$

Theorem 3. *Fix* $m_0, \widetilde{m}_1, \widetilde{m}_2, \widetilde{m}_3 > 0$ *and consider the Hamiltonian* H *in* (2.1) *expressed in Deprit coordinates with masses* $m_0, m_j = \rho \widetilde{m}_j$ *with* $j = 1, 2, 3$. *Then, there exists* $0 < \kappa \ll 1$, $\alpha_1, \alpha_2, \beta > 0$, *such that the following is satisfied.*

Fix $N \ge 1$ *any* $\{\nu^k\}_{k=0}^N \subset (0, 1)$, $\{\eta^k\}_{k=0}^N \subset (-1, 1)$, $\{\zeta^k\}_{k=0}^N \subset (1, +\infty)$ *and constants* $\check{L}_1^0$, $\check{L}_2^0$ *and* $\check{\Psi}_2^0$ *satisfying*

$$\check{L}_1^0 \in \left[\frac{1}{2}, 2\right], \qquad \check{L}_1^0 \ll \check{L}_2^0, \qquad \check{\Psi}_2^0 \gg \rho^{-1/3} \qquad and \qquad |\check{\Psi}_1^0 - \nu_0 \check{L}_2^0| \le \kappa.$$

Then, there exists an orbit of the Hamiltonian H *in* (2.1) *expressed in scaled Deprit coordinates and times* $\{t_k\}_{k=0}^N$ *satisfying*
$$t_0 = 0 \qquad and \qquad |t_k| \le C(\check{L}_2^0)\rho^{-35/3}, \ k \ge 1,$$
where $C(\check{L}_2^0)$ *is a constant depending on* $\check{L}_2^0$ *but independent of* ρ, *such that*

$$|\check{\Gamma}_2(t_k) - \nu_k \check{L}_2^0| \le (\check{L}_2^0)^{-\beta}$$
$$|\check{\Psi}_2^0 - \check{\Gamma}_3(t_k) - \eta_k \check{\Psi}_1(t_k)| \le (\check{L}_2^0)^{-\beta}$$
$$|\check{L}_3(t_k) - \zeta_k \check{\Psi}_2^0| \le (\check{L}_2^0)^{-\beta}.$$

Moreover,
$$|\check{\Gamma}_1(t_k) - \check{L}_1^0| \le (\check{L}_2^0)^{-\beta}, \qquad |\check{\Psi}_1(t_k) - \check{\Gamma}_2(t_k) - M_k| \le (\check{L}_2^0)^{-\beta}$$
where $M_k \in (0, \kappa)$ *is determined by*

$$\frac{M_k^2}{(\check{L}_2^0 - \check{\Gamma}_2(t_k))^{3/2}} = \frac{M_0^2}{(\check{L}_2^0 - \check{\Gamma}_2^0)^{3/2}} \qquad and \qquad M_0 = \check{\Psi}_1^0 - \check{\Gamma}_2^0$$

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

$$|\check{\Gamma}_3(t) - \check{\Gamma}_3^0| \le 2\check{L}_2^0, \qquad and \qquad |\check{L}_j(t) - \check{L}_j^0| \le (\check{L}_2^0)^{-\beta} \quad for \quad j = 1, 2.$$

3 Computation of the secular Hamiltonian

In this section we compute the secular Hamiltonian in three steps: first, we expand the perturbing function (2.1) using the Legendre polynomials; next, we observe that the mean anomalies ℓ_1, ℓ_2 are faster than the other variables, and we use this fact to perform a near-identity symplectic coordinate transformation that averages the angles ℓ_1, ℓ_2 out of the perturbing function up to arbitrarily high order; finally, we make a further symplectic coordinate transformation so that the new action variables are all of order 1, and we use these variables to expand the Taylor series of the secular Hamiltonian.

This section has strong similarities with the computation of the secular Hamiltonian in [13]. We do include it for the convenience of the reader and because terms describing the motion of planet 3 differ significantly.

3.1 Expansion of the perturbing function in Legendre polynomials

Since $\|q_j\| = O(a_j) = O(L_j^2)$, the assumption (2.2) implies that $\|q_1\| \ll \|q_2\| \ll \|q_3\|$. Denote by ζ_j the angles between q_j and q_{j+1} for $j = 1, 2$, and denote by P_n the Legendre polynomial of degree n. Observe that we can write the perturbing function F_{per}, defined by (2.1), as

$$F_{\text{per}} = F_{\text{per}}^{12} + F_{\text{per}}^{23} + O\left(\frac{1}{a_3^3}\right) \tag{19}$$

where

$$F_{\text{per}}^{12} = \frac{\mu_2 M_2}{\|q_2\|} - \frac{m_0\, m_2}{\|q_2 + \sigma_{11}\, q_1\|} - \frac{m_1\, m_2}{\|q_2 - \sigma_{01}\, q_1\|} = -\frac{\mu_1 m_2}{\|q_2\|} \sum_{n=2}^{\infty} \tilde{\sigma}_{1,n} P_n(\cos\zeta_1) \left(\frac{\|q_1\|}{\|q_2\|}\right)^n \tag{20}$$

is the perturbing function of the inner 3-body problem, and where

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

describes the interactions between bodies 2 and 3, with

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

and σ_{ij} are defined in (2.1).

3.2 Averaging of the mean anomalies ℓ_1 and ℓ_2

In Deprit coordinates the Kepler Hamiltonian F_{Kep}, defined by (2.1), is given by

$$F_{\text{Kep}} = -\sum_{j=1}^{3} \frac{\mu_j^3\, M_j^2}{2\, L_j^2}. \tag{21}$$

From Hamilton's equations of motion we see that the first order of $\dot{\ell}_j$ is $\frac{\partial F_{\text{Kep}}}{\partial L_j} = \frac{\mu_j^3\, M_j^2}{L_j^3}$. Since the first order term in F_{per} is of order $\frac{\|q_1\|^2}{\|q_2\|^3} = O\left(L_2^{-6}\right)$, it follows that the angles ℓ_1, ℓ_2 are faster than all other variables. Therefore standard averaging arguments imply that we can perform a near-identity coordinate transformation so that, in the new coordinates, the Hamiltonian does not depend on the angles ℓ_1, ℓ_2 up to arbitrarily (but finitely) high order terms. Effecting the coordinate transformation, the Hamiltonian $H = F_{\text{Kep}} + F_{\text{per}}$ becomes

$$F = F_{\text{Kep}} + \tilde{F}_{\text{sec}} + \frac{1}{L_2^{10}} R_1 + \frac{1}{L_3^6} R_2, \tag{22}$$

where F_{Kep} is given by (3.2), the Hamiltonian $\tilde{F}_{\text{sec}}$ is defined by

$$\tilde{F}_{\text{sec}} = F_{\text{sec}}^{12} + F_{\text{sec}}^{23} + O\left(\frac{1}{L_3^6}\right), \tag{23}$$

with

$$F_{\text{sec}}^{12} = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} F_{\text{per}}^{12} \, d\ell_1 \, d\ell_2 = -\frac{1}{(2\pi)^2} \frac{\mu_1 m_2}{\|q_2\|} \sum_{n=2}^{\infty} \tilde{\sigma}_{1,n} \int_{\mathbb{T}^2} P_n(\cos \zeta_1) \left(\frac{\|q_1\|}{\|q_2\|} \right)^n d\ell_1 \, d\ell_2 \tag{24}$$

$$F_{\text{sec}}^{23} = \frac{1}{2\pi} \int_{\mathbb{T}} F_{\text{per}}^{12} d\ell_2 = -\frac{1}{2\pi} \frac{\mu_2 m_3}{\|q_3\|} \sum_{n=2}^{\infty} \tilde{\sigma}_{2,n} \int_{\mathbb{T}} P_n(\cos \zeta_2) \left(\frac{\|q_2\|}{\|q_3\|} \right)^n d\ell_2. \tag{25}$$

Moreover the remainder term R_1 depends only on the variables on which F_{per}^{12} depends.

Remark 4. *We use the following terminology and notation.*

1. *Using the usual terminology from the literature, we refer to the terms in the expansion of F_{sec}^{12} obtained by setting $n = 2, 3$ in (3.2) as the quadrupolar, octupolar (respectively) Hamiltonians of the interaction between bodies 1 and 2, and we write F_{quad}^{12}, F_{oct}^{12} (respectively) to denote these Hamiltonians.*

2. *In addition, we refer to the $n = 2$ term in the expansion (3.2) of F_{sec}^{23} as the quadrupolar Hamiltonian of the interaction between bodies 2 and 3, and we write F_{quad}^{23} to denote this Hamiltonian. Note that this terminology is generally reserved in the literature for the term obtained by averaging both ℓ_2 and ℓ_3 (see for example [13]); however, in this paper, the angle ℓ_3 is slower than, for example γ_1 (see Proposition 7 below), and so it cannot be averaged from the perturbing function. We therefore consider this terminology and notation to be appropriate in this instance.*

In this paper we require only the Hamiltonians F_{quad}^{12}, F_{oct}^{12}, F_{quad}^{23}. Expanding the first two terms of (3.2) and (3.2) and using the notation of Remark 4 we obtain

$$F_{\text{sec}}^{12} = -\frac{\mu_1 m_2}{(2\pi)^2} \left(F_{\text{quad}}^{12} + \tilde{\sigma}_{1,3} \, F_{\text{oct}}^{12} + O\left(\frac{a_1^4}{a_2^5} \right) \right), \quad F_{\text{sec}}^{23} = -\frac{\mu_2 m_3}{2\pi} \left(F_{\text{quad}}^{23} + O\left(\frac{a_2^3}{a_3^4} \right) \right),$$

where

$$\begin{cases} F_{\text{quad}}^{12} = \int_{\mathbb{T}^2} P_2\left(\cos \zeta_1 \right) \frac{\|q_1\|^2}{\|q_2\|^3} \, d\ell_1 \, d\ell_2, \quad F_{\text{oct}}^{12} = \int_{\mathbb{T}^2} P_3\left(\cos \zeta_1 \right) \frac{\|q_1\|^3}{\|q_2\|^4} \, d\ell_1 \, d\ell_2 \\ F_{\text{quad}}^{23} = \int_{\mathbb{T}} P_2\left(\cos \zeta_2 \right) \frac{\|q_2\|^2}{\|q_3\|^3} \, d\ell_2 \end{cases}$$

since $\tilde{\sigma}_{j,2} = 1$ for $j = 1, 2$. In the following Lemma (proved in [13]), we compute F_{quad}^{12} and F_{oct}^{12} explicitly in terms of Deprit coordinates; we could perform similar computations to compute F_{quad}^{23}, but the resulting expression would be very long. Instead, we compute F_{quad}^{23} in Section 3.3 via a Taylor expansion, after making a suitable linear symplectic coordinate transformation.

Lemma 5. *The quadrupolar and octupolar Hamiltonians of bodies 1 and 2 are given by*

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

and

$$F_{\text{oct}}^{12} = -\frac{15}{64} \frac{a_1^3}{a_2^4} \frac{e_1 \, e_2}{\left(1 - e_2^2 \right)^{\frac{5}{2}}}$$

$$\times \left\{ \begin{array}{l} \cos \gamma_1 \, \cos \gamma_2 \left[\begin{array}{l} \dfrac{\Gamma_1^2}{L_1^2} \left(5 \, \sin^2 i_{12} \left(6 - 7 \cos^2 \gamma_1 \right) - 3 \right) \\[2mm] -35 \, \sin^2 \gamma_1 \, \sin^2 i_{12} + 7 \end{array} \right] \\[6mm] + \sin \gamma_1 \, \sin \gamma_2 \, \cos i_{12} \left[\begin{array}{l} \dfrac{\Gamma_1^2}{L_1^2} \left(5 \, \sin^2 i_{12} \left(4 - 7 \cos^2 \gamma_1 \right) - 3 \right) \\[2mm] -35 \, \sin^2 \gamma_1 \, \sin^2 i_{12} + 7 \end{array} \right] \end{array} \right\} \tag{27}$$

respectively, where the eccentricity e_j of the jth Keplerian ellipse is defined by (2.1), and where i_{12} is the mutual inclination of Keplerian bodies 1 and 2, defined by

$$\cos i_{12} = \frac{\Psi_1^2 - \Gamma_1^2 - \Gamma_2^2}{2\,\Gamma_1\,\Gamma_2}.$$

3.3   Taylor expansion of the secular Hamiltonian

In order to perform the subsequent analysis, we divide the phase space into strips where the actions live in some bounded region. In each strip we perform an affine coordinate transformation so that the new actions have order 1, thus allowing us to perform a further Taylor expansion of the secular Hamiltonian.

Recall we assume that the semimajor axes satisfy (2.2), which in Deprit coordinates corresponds to the assumption (2.2) (if the masses are assumed to be fixed). Now, the variables Γ_2, Ψ_1 are of order L_2, while Γ_3, Ψ_2 are of order L_3. Fix some large positive value L_3^* of L_3. The total angular momentum Ψ_2 is conserved, so we write

$$\Psi_2 = \delta_2\,L_3^* \tag{28}$$

for some fixed $\delta_2 > 0$. We make the symplectic change of variables:

$$\begin{cases} L_3 = L_3^* + \tilde{L}_3, & \tilde{\ell}_3 = \ell_3 \\ \tilde{\Psi}_1 = \Psi_1 - \delta_1\,L_2, & \tilde{\psi}_1 = \psi_1 + \gamma_2 \\ \tilde{\Gamma}_2 = \Psi_1 - \Gamma_2, & \tilde{\gamma}_2 = -\gamma_2 \\ \tilde{\Gamma}_3 = \Psi_2 - \Gamma_3 - \delta_3\,L_2, & \tilde{\gamma}_3 = -\gamma_3 \end{cases} \tag{29}$$

where $\delta_1, \delta_3 > 0$ are constant with respect to the Hamiltonian $\tilde{F}_{\text{sec}}$. Note that this symplectic transformation does not modify the variables γ_1, Γ_1 (or indeed ℓ_3). We assume that L_3^* is chosen so that $\tilde{L}_3 = O(1)$.

Moreover, we assume that

$$\tilde{\Gamma}_2 > 0 \tag{30}$$

as the case where $\tilde{\Gamma}_2$ is negative can be treated analogously. Furthermore, 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^*.

Remark 6. *We make the following remarks regarding this coordinate transformation and notation.*

1. *By choosing different values of the constants $\delta_1, \delta_2, \delta_3$ (and adjusting L_3^* so that (3.3) still holds) we can focus on any relevant region of the phase space, in order to make the coordinate transformation (3.3). This fact will be of importance in Section 7.3, where we construct trajectories that drift through many of these different regions.*

2. *Observe that we could equally have used the total angular momentum Ψ_2 as a parameter instead of δ_2, as Ψ_2, L_3^* are constant, and (3.3) implies that $\delta_2 = \frac{\Psi_2}{L_3^*}$. The reason that we have used the notation δ_2 instead is to maintain consistency with [13]; indeed, we use several results from [13] and it is easier to compare the formulas if the notation is the same.*

Using the coordinates (3.3), the Keplerian Hamiltonian F_{Kep} takes the form

$$F_{\text{Kep}} = \tilde{F}_{\text{Kep}} + \frac{1}{(L_3^*)^3}\,\alpha_{\text{Kep}}\,\tilde{L}_3 - \frac{1}{(L_3^*)^4}\frac{3}{2}\alpha_{\text{Kep}}\,\tilde{L}_3^2 + O\left(\frac{1}{(L_3^*)^5}\right) \tag{31}$$

where

$$\tilde{F}_{\text{Kep}} = -\sum_{j=1}^{2}\frac{\mu_j^3\,M_j^2}{2\,L_j^2} - \frac{\mu_3^3\,M_3^2}{2\,(L_3^*)^2}, \quad \alpha_{\text{Kep}} = \mu_3^3\,M_3^2. \tag{32}$$

Since the angle $\tilde{\ell}_3$ evolves slower than some of the secular angles (i.e. $\gamma_1, \tilde{\gamma}_2, \tilde{\psi}_1$; see Proposition 7 below) we must consider it as a secular variable. Therefore, we include $F_{\text{Kep}} - \tilde{F}_{\text{Kep}}$ in the secular Hamiltonian F_{sec}, which we define by

$$F_{\text{sec}} = \tilde{F}_{\text{sec}} + \left(F_{\text{Kep}} - \tilde{F}_{\text{Kep}}\right) \tag{33}$$

14

where $\tilde{F}_{\text{sec}}$ is defined by (3.2).

In the following proposition, we identify the first appearance of each of the secular variables in the secular Hamiltonian F_{sec}. In addition, we identify the first appearance of products of trigonometric functions of $\tilde{\psi}_1$, $\tilde{\gamma}_3$, $\tilde{\ell}_3$ with functions of γ_1, Γ_1, $\tilde{\gamma}_2$; this is of significance as these are the terms that will contribute to the Poincaré-Melnikov computation in Section 6.

Proposition 7. *The secular Hamiltonian F_{sec}, defined by (3.3), can be expanded in the form*

$$F_{\text{sec}} = c + \sum_{i,j=0}^{\infty} \varepsilon^i \, \mu^j \, F_{ij}$$

where $\varepsilon = \frac{1}{L_2}$, $\mu = \frac{L_2}{L_3}$, and where the terms in the 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 α_i^{12} are nontrivial constants, and where the Hamiltonians H_0^{12}, H_1^{12} are defined by (10) and (10) 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_{quad}^{12} (see (5)). 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 (11), and is the first order term in the expansion of F_{oct}^{12} (see (5)), up to a multiplicative constant.*

3. *The action $\tilde{L}_3$ first appears in $F_{3,3}$, specifically in the term of order $(L_3^*)^{-3}$ in the expansion of $F_{\text{Kep}} - \tilde{F}_{\text{Kep}}$ (see (3.3), (3.3), and (3.3)).*

4. *Each of the angles $\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3$ first appears in the Hamiltonian H_0^{23} (see (12)) which is contained in $F_{2,6}$.*

5. *The action $\tilde{\Gamma}_3$ first appears in the Hamiltonian $\tilde{\Gamma}_3 \tilde{H}_3$ (see (12) and (12)) which is contained in $F_{3,6}$.*

6. *The Hamiltonian H_1^{23}, contained in $F_{3,6}$, presents the first products of functions of each of the angles $\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3$ with functions of $\gamma_1, \Gamma_1, \tilde{\gamma}_2$.*

Proof. The proposition follows from Lemmas 10, 11, and 12 below, upon comparing the orders of the coefficients of each Hamiltonian using the assumptions (2.2). $\qquad\square$

Notation 8. *Throughout this paper, in order to simplify notation, we use ellipsis to mean the following. Fix some sufficiently large integer $r \in \mathbb{N}$. The notation $F = \epsilon^i \, \mu^j \, G + \cdots$ means that there are $\eta_1, \eta_2 \in \mathbb{N}_0$, not both 0, and a positive constant C such that*

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

Moreover, we use the expression nontrivial constant *to mean a constant depending only on the masses and the parameters δ_j that is nonzero for all $m_0, m_1, m_2, m_3 > 0$ satisfying (2), all $\delta_1, \delta_2 \in (0,1)$, and all $\delta_3 \in (-1,1)$.*

Remark 9. *Although the secular Hamiltonian depends on the mean anomaly $\tilde{\ell}_3$, the most natural way to compute such Hamiltonians is by using the true anomaly, denoted by v_3. Throughout this paper, we will write the dependence of the Hamiltonian on the variable $\tilde{\ell}_3$ (and indeed later on similar variables ℓ_3', $\hat{\ell}_3$ obtained via near-identity coordinate transformations); however the dependence of the Hamiltonian on this variable will be seen only implicitly through the Hamiltonian's dependence on the true anomaly v_3. When we need to differentiate such a Hamiltonian, say K, with respect to the mean anomaly $\tilde{\ell}_3$, we obtain $\frac{\partial K}{\partial \tilde{\ell}_3} = \frac{\partial K}{\partial v_3} \frac{\partial v_3}{\partial \tilde{\ell}_3}$, and we notice that, due to Kepler's second law, we have*

$$\frac{\partial v_3}{\partial \tilde{\ell}_3} = \frac{\partial v_3}{\partial \ell_3} = \left(1 - e_3^2\right)^{-\frac{3}{2}} \left(1 + e_3 \cos v_3\right)^2 = \delta_2^{-3} \left(1 + \sqrt{1 - \delta_2^2} \cos v_3\right)^2 + O\left(\frac{L_2}{L_3^*}\right),$$

where we have expanded the eccentricity e_3 using (2.1) and (3.3).

15

The quadrupolar and octopolar Hamiltonians F^{12}_{quad} and F^{12}_{oct} of bodies 1 and 2 take the same form as in [13], as described in the following two lemmas (see [13] for the proofs).

Lemma 10. *The Hamiltonian F^{12}_{quad} can be written in the variables (3.3) as*

$$F^{12}_{\text{quad}} = \tilde{c}^{12}_0 + \frac{1}{L^6_2}\,\alpha^{12}_0\,H^{12}_0 + \frac{1}{L^7_2}\,\alpha^{12}_1\,H^{12}_1 + \frac{1}{L^8_2}\,\tilde{\alpha}_2\,\tilde{H}_2 + \cdots$$

where

$$H^{12}_0 = \left(1 - \frac{\Gamma^2_1}{L^2_1}\right)\left[2 - 5\left(1 - \frac{\tilde{\Gamma}^2_2}{\Gamma^2_1}\right)\sin^2\gamma_1\right] + \frac{\tilde{\Gamma}^2_2}{L^2_1} \tag{34}$$

$$H^{12}_1 = \left(3H^{12}_0\left(\gamma_1, \Gamma_1, \tilde{\Gamma}_2\right) - 1\right)\tilde{\Psi}_1 - 4\tilde{\Gamma}_2 H^{12}_0\left(\gamma_1, \Gamma_1, \tilde{\Gamma}_2\right) + 3\tilde{\Gamma}_2 - \frac{\Gamma^2_1\tilde{\Gamma}_2}{L^2_1} \tag{35}$$

$$\tilde{H}_2 = \left(3H^{12}_0\left(\gamma_1, \Gamma_1, \tilde{\Gamma}_2\right) - 1\right)\tilde{\Psi}^2_1 + \left(6 - 8H^{12}_0\left(\gamma_1, \Gamma_1, \tilde{\Gamma}_2\right) - 2\frac{\Gamma^2_1}{L^2_1}\right)\tilde{\Gamma}_2\tilde{\Psi}_1$$

$$+ \frac{1}{8}\left[\sin^2\gamma_1\left(5\,\Gamma^2_1 - \frac{5\,\Gamma^4_1}{L^2_1} + \frac{210\,\Gamma^2_1\tilde{\Gamma}^2_2}{L^2_1} - \frac{205\,\tilde{\Gamma}^4_2}{L^2_1} + \frac{205\,\tilde{\Gamma}^4_2}{\Gamma^2_1}\right)\right.$$

$$\left. + \frac{\Gamma^4_1}{L^2_1} - \frac{66\,\Gamma^2_1\tilde{\Gamma}^2_2}{L^2_1} + \frac{41\,\tilde{\Gamma}^4_2}{L^2_1} + 40\,\tilde{\Gamma}^2_2\right]$$

and

$$\alpha^{12}_0 = \frac{3\,L^4_1\,M^3_2\,\mu^6_2}{8\,M^2_1\,\delta^3_1\,\mu^4_1}, \quad \alpha^{12}_1 = -\frac{3\,L^4_1\,M^3_2\,\mu^6_2}{8\,M^2_1\,\delta^4_1\,\mu^4_1}, \quad \tilde{\alpha}_2 = \frac{3\,L^4_1\,M^3_2\,\mu^6_2}{4\,M^2_1\,\delta^5_1\,\mu^4_1}.$$

Moreover F^{12}_{quad} is integrable.

Lemma 11. *The Hamiltonian F^{12}_{oct} can be written in the rescaled variables (3.3) as*

$$F^{12}_{\text{oct}} = \frac{1}{L^8_2}\,\alpha^{12}_2\,H^{12}_2 + \cdots$$

where

$$H^{12}_2 = \sqrt{1 - \frac{\Gamma^2_1}{L^2_1}}\left\{\begin{array}{l}\cos\gamma_1\,\cos\tilde{\gamma}_2\left[\begin{array}{l}\frac{\Gamma^2_1}{L^2_1}\left(5\left(1 - \frac{\tilde{\Gamma}^2_2}{\Gamma^2_1}\right)\left(6 - 7\cos^2\gamma_1\right) - 3\right) \\ -35\sin^2\gamma_1\left(1 - \frac{\tilde{\Gamma}^2_2}{\Gamma^2_1}\right) + 7\end{array}\right] \\ +\frac{\tilde{\Gamma}_2}{\Gamma_1}\,\sin\gamma_1\,\sin\tilde{\gamma}_2\left[\begin{array}{l}\frac{\Gamma^2_1}{L^2_1}\left(5\left(1 - \frac{\tilde{\Gamma}^2_2}{\Gamma^2_1}\right)\left(4 - 7\cos^2\gamma_1\right) - 3\right) \\ -35\sin^2\gamma_1\left(1 - \frac{\tilde{\Gamma}^2_2}{\Gamma^2_1}\right) + 7\end{array}\right]\end{array}\right\} \tag{36}$$

and

$$\alpha^{12}_2 = -\frac{15}{64}\frac{L^6_1\mu^8_2 M^4_2}{\mu^6_1 M^3_1}\frac{\sqrt{1 - \delta^2_1}}{\delta^5_1}.$$

The quadrupolar Hamiltonian F^{23}_{quad} of bodies 2 and 3 is expanded in the following lemma.

Lemma 12. *We can write the Hamiltonian F^{23}_{quad} as*

$$F^{23}_{\text{quad}} = \frac{L^4_2}{(L^*_3)^6}\,\alpha^{23}_0\,K_0 + \frac{L^3_2}{(L^*_3)^6}\,\alpha^{23}_1\,K_1 + \cdots \tag{37}$$

16

where the Hamiltonians K_0, K_1 are themselves defined via the expansions

$$\begin{cases} K_0 = H_0^{23}\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right) + \dfrac{1}{L_2}\left[\tilde{\Gamma}_3\,\tilde{H}_3\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right) + \tilde{H}_4\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3, \tilde{\Psi}_1\right)\right] + \cdots, \\[2mm] K_1 = H_1^{23}\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3, \Gamma_1, \tilde{\Gamma}_2\right) + \cdots, \end{cases} \tag{38}$$

and we have

$$\begin{cases} H_0^{23} = \left(1 + \sqrt{1 - \delta_2^2}\,\cos v_3\right)^3 \left[A_0\left(\tilde{\gamma}_3, v_3\right)\cos^2\tilde{\psi}_1 + B_0\left(\tilde{\gamma}_3, v_3\right)\cos\tilde{\psi}_1\,\sin\tilde{\psi}_1 + C_0\left(\tilde{\gamma}_3, v_3\right)\right], \\[2mm] H_1^{23} = \left(1 + \sqrt{1 - \delta_2^2}\,\cos v_3\right)^3 \sqrt{\Gamma_1^2 - \tilde{\Gamma}_2^2}\left[A_1\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right)\cos\tilde{\gamma}_2 + B_1\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right)\sin\tilde{\gamma}_2\right] \end{cases} \tag{39}$$

where the trigonometric polynomials A_0, B_0, C_0 are given by

$$A_0\left(\tilde{\gamma}_3, v_3\right) = 15\left[\left(\delta_3^2 - \frac{\delta_3^2}{\delta_1^2} - \left(1 - \delta_1^2\right)\right)\sin^2\left(v_3 - \tilde{\gamma}_3\right) + \left(1 - \delta_1^2\right)\right] \tag{40}$$

$$B_0\left(\tilde{\gamma}_3, v_3\right) = 30\,\frac{\delta_3}{\delta_1}\left(1 - \delta_1^2\right)\cos\left(v_3 - \tilde{\gamma}_3\right)\sin\left(v_3 - \tilde{\gamma}_3\right) \tag{41}$$

$$C_0\left(\tilde{\gamma}_3, v_3\right) = \left(15\,\frac{\delta_3^2}{\delta_1^2} - 12\,\delta_3^2 - 3\,\delta_1^2\right)\sin^2\left(v_3 - \tilde{\gamma}_3\right) + 6\,\delta_1^2 - 5$$

and the trigonometric polynomials A_1, B_1, $\tilde{H}_3$ are given by

$$A_1\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right) = \delta_1\,\delta_3\,\cos\tilde{\psi}_1\,\sin^2\left(v_3 - \tilde{\gamma}_3\right) - \delta_1^2\,\sin\tilde{\psi}_1\,\cos\left(v_3 - \tilde{\gamma}_3\right)\sin\left(v_3 - \tilde{\gamma}_3\right)$$

$$B_1\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right) = \left(4\,\delta_1\,\delta_3 - 5\,\frac{\delta_3}{\delta_1}\right)\sin\tilde{\psi}_1\,\sin^2\left(v_3 - \tilde{\gamma}_3\right) + \left(4\,\delta_1^2 - 5\right)\cos\tilde{\psi}_1\,\cos\left(v_3 - \tilde{\gamma}_3\right)\sin\left(v_3 - \tilde{\gamma}_3\right)$$

$$\tilde{H}_3\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right) = \left(1 + \sqrt{1 - \delta_2^2}\,\cos v_3\right)^3 \left[\left(30\,\delta_3\left(1 - \frac{1}{\delta_1^2}\right)\cos^2\tilde{\psi}_1 + 30\,\frac{\delta_3}{\delta_1^2} - 24\,\delta_3\right)\sin^2\left(v_3 - \tilde{\gamma}_3\right)\right.$$
$$\left. + 30\,\frac{1}{\delta_1}\left(1 - \delta_1^2\right)\cos\tilde{\psi}_1\,\sin\tilde{\psi}_1\,\cos\left(v_3 - \tilde{\gamma}_3\right)\sin\left(v_3 - \tilde{\gamma}_3\right)\right].$$

Proof. By definition, we have

$$F_{\text{quad}}^{23} = \int_{\mathbb{T}} P_2\left(\cos\zeta_2\right)\frac{\|q_2\|^2}{\|q_3\|^3}\,d\ell_2. \tag{42}$$

Denote by $\mathcal{R}_1(\theta)$, $\mathcal{R}_3(\theta)$ the rotation matrix by an angle θ 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 ℓ_j. By Proposition 4.1 of [52], we have

$$\bar{q}_2 = \mathcal{R}_3(\psi_3)\,\mathcal{R}_1(i)\,\mathcal{R}_3(\psi_2)\,\mathcal{R}_1(\tilde{i}_2)\,\mathcal{R}_3(\psi_1)\,I_3\,\mathcal{R}_1(i_2)\,\bar{Q}_2$$
$$\bar{q}_3 = \mathcal{R}_3(\psi_3)\,\mathcal{R}_1(i)\,\mathcal{R}_3(\psi_2)\,I_3\,\mathcal{R}_1(i_3)\,\bar{Q}_3$$

where

$$\cos i = \frac{\Psi_3}{\Psi_2}, \qquad \cos\tilde{i}_2 = \frac{\Psi_2^2 + \Psi_1^2 - \Gamma_3^2}{2\,\Psi_1\,\Psi_2}, \qquad \cos i_2 = \frac{\Gamma_2^2 + \Psi_1^2 - \Gamma_1^2}{2\,\Psi_1\,\Gamma_2}, \qquad \cos i_3 = \frac{\Gamma_3^2 + \Psi_2^2 - \Psi_1^2}{2\,\Psi_2\,\Gamma_3}.$$

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,

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

$$\cos i_3 = 1 + O\left(\frac{L_2}{L_3^*}\right), \quad \sin i_3 = O\left(\frac{L_2}{L_3^*}\right).$$

Therefore we can write

$$\mathcal{R}_1(i_2) = \mathrm{Id} + L_2^{-1} M_1 + O\left(L_2^{-2}\right), \quad \mathcal{R}_1(i_3) = \mathrm{Id} + O\left(\frac{L_2}{L_3^*}\right)$$

where

$$M_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -b_1 \\ 0 & b_1 & 0 \end{pmatrix} \quad \text{with} \quad b_1 = \frac{\sqrt{\Gamma_1^2 - \tilde{\Gamma}_2^2}}{\delta_1}.$$

We thus obtain the expression

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

where

$$W_0 = \mathcal{R}_1(\tilde{i}_2)\, \mathcal{R}_3(\psi_1)\, I_3\, \bar{Q}_2 \cdot I_3\, \bar{Q}_3,$$
$$W_1 = \mathcal{R}_1(\tilde{i}_2)\, \mathcal{R}_3(\psi_1)\, I_3\, M_1\, \bar{Q}_2 \cdot I_3\, \bar{Q}_3.$$

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

$$P_2(\cos \zeta_2) = P_2(W_0) + 3\,\frac{1}{L_2}\, W_0\, W_1 + O\left(\frac{L_2}{L_3^*}\right). \tag{43}$$

In what follows, we integrate the two terms on the right-hand side of (3.3) separately using the technique introduced in Appendix C of [25]. We have

$$\begin{cases} \|q_2\| = a_2\, \rho_2, \quad \rho_2 = 1 - e_2 \cos u_2 \\ \|q_3\| = a_3\left(1 - e_3^2\right)\varrho_3, \quad \dfrac{1}{\varrho_3} = 1 + e_3 \cos v_3 \end{cases}$$

where u_2 is the eccentric anomaly of body 2 and v_3 is the true anomaly of body 3. By differentiating the Kepler equation $\ell_2 = u_2 - e_2 \sin u_2$ we obtain $d\ell_2 = \rho_2\, du_2$. Combining these formulas with (3.3) we see that

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

where

$$K_0 = \varrho_3^{-3} \int_{\mathbb{T}} P_2\left(W_0\right) \rho_2^3\, du_2, \quad K_1 = \varrho_3^{-3} \int_{\mathbb{T}} W_0\, W_1\, \rho_2^3\, du_2.$$

Using the expressions

$$\rho_2 \cos v_2 = \cos u_2 - e_2, \quad \rho_2 \sin v_2 = \sqrt{1 - e_2^2}\, \sin u_2,$$

we can eliminate the angle v_2 from the integrands of K_0 and K_1. The result is a trigonometric polynomial which can be computed by quadrature and combined with the expansions

$$e_2^2 = (1 - \delta_1^2) + \frac{1}{L_2}\, 2\delta_1\left(\tilde{\Gamma}_2 - \tilde{\Psi}_1\right) + O\left(L_2^{-2}\right), \quad e_3^2 = 1 - \delta_2^2 + O\left(\frac{L_2}{L_3^*}\right),$$

and

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

to complete the proof of the lemma. □

4 First-order integrable dynamics

The purpose of this section is to establish the existence of a normally hyperbolic invariant manifold for the first-order term H_0^{12}, defined by (10), in the expansion of the secular Hamiltonian.

This section is similar to the corresponding section in [13]. We include it for the sake of completeness.

The Deprit coordinates conveniently produce in the four-body problem the same Hamiltonian F_{quad}^{12} (see (5)) as the quadrupolar Hamiltonian of the three-body problem, expressed in Delaunay coordinates (see for example [27]; see also Section 4 of [13]). Therefore we can use the analysis and results of [27] (up to some errata; see Appendix F of [13]). In summary, we will show that H_0^{12} has a hyperbolic periodic orbit with a homoclinic connection (in some system of coordinates); since H_0^{12} is integrable (notice that it does not depend on $\tilde{\gamma}_2$), this homoclinic connection is a separatrix, lying in the non-transverse homoclinic intersection of the stable and unstable manifolds.

The Hamiltonian vector field of H_0^{12} in the γ_1, Γ_1 directions, obtained by differentiating (10), is

$$\begin{cases} \dot{\gamma}_1 = \dfrac{\partial H_0^{12}}{\partial \Gamma_1} = \dfrac{2\,\Gamma_1}{L_1^2} \left[5 \left(1 - \dfrac{\tilde{\Gamma}_2^2}{\Gamma_1^2} \right) \sin^2 \gamma_1 - 2 \right] - 10 \left(1 - \dfrac{\Gamma_1^2}{L_1^2} \right) \dfrac{\tilde{\Gamma}_2^2}{\Gamma_1^3} \sin^2 \gamma_1 \\[2ex] \dot{\Gamma}_1 = -\dfrac{\partial H_0^{12}}{\partial \gamma_1} = 5 \left(1 - \dfrac{\Gamma_1^2}{L_1^2} \right) \left(1 - \dfrac{\tilde{\Gamma}_2^2}{\Gamma_1^2} \right) \sin 2\gamma_1 . \end{cases} \tag{44}$$

This vector field has (among others) two equilibria whenever $\Gamma_1 = L_1$ and

$$\tilde{\Gamma}_2 < L_1 \sqrt{\frac{3}{5}} . \tag{45}$$

Indeed, whenever (4) holds there are two solutions $\gamma_1^{\min} \in (0, \pi)$ and $\gamma_1^{\max} = \pi - \gamma_1^{\min}$ to the equation

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

Therefore $(\gamma_1, \Gamma_1) = (\gamma_1^{\min,\max}, L_1)$ are equilibria of the Hamiltonian vector field (4). Moreover, these equilibria are hyperbolic; this can be seen clearly below when we pass to Poincaré variables (4). The Hamiltonian H_0^{12} also depends on $\tilde{\Gamma}_2$; indeed, we have $\frac{d\tilde{\gamma}_2}{dt} = \frac{\partial H_0^{12}}{\partial \tilde{\Gamma}_2} = \frac{2\tilde{\Gamma}_2}{L_1^2}$. Therefore, lifting the equilibria to the full phase space of H_0^{12} by including the variables $\tilde{\gamma}_2$, $\tilde{\Gamma}_2$, we obtain the periodic orbits

$$Z_{\min,\max}^0 \left(t, \tilde{\gamma}_2^0 \right) = \left(\gamma_1^{\min,\max}, L_1, \tilde{\gamma}_2^0 + \tilde{\gamma}_2^1(t), \tilde{\Gamma}_2 \right) \tag{46}$$

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

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

Remark 13. *The assumptions we have discussed in the introduction regarding inclination are seen mathematically in (4) (see also Remark 25 below for a refinement of this remark). Indeed, from (2.1) and the change of coordinates (3.3), we see that $\cos i_{12} = \frac{\tilde{\Gamma}_2}{\Gamma_1} + O\left(L_2^{-1} \right)$. Therefore on the circular ellipse $\{\Gamma_1 = L_1\}$, the assumptions (3.3) and (4) imply that $|\cos i_{12}| < \sqrt{\frac{3}{5}} + O\left(L_2^{-1} \right)$. This implies that i_{12} is more than roughly $40°$.*

Suppose (4) holds, and recall moreover we have assumed in (3.3) that $\tilde{\Gamma}_2 > 0$. Define the positive constants

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

The proof of the following result is identical to the proof of Lemma 3.1 in [27] (see also Appendix F of [13]).

Lemma 14. *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

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

where

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

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

and

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

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 (what is the perihelion of a circle?). We therefore introduce the Poincaré variables

$$\xi = \sqrt{2 \, (L_1 - \Gamma_1)} \cos \gamma_1, \quad \eta = -\sqrt{2 \, (L_1 - \Gamma_1)} \sin \gamma_1. \tag{47}$$

This is a symplectic change of variables, in the sense that $d\xi \wedge d\eta = d\Gamma_1 \wedge d\gamma_1$. In these variables, the Hamiltonian H_0^{12} becomes

$$\tilde{H}_0^{12} = \frac{1}{L_1} \left[2 \, \xi^2 - \left(3 - 5 \frac{\tilde{\Gamma}_2^2}{L_1^2}\right) \eta^2 \right] + \frac{\tilde{\Gamma}_2^2}{L_1^2} + O_2 \left(\xi^2 + \eta^2\right) \tag{48}$$

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

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

Moreover, the heteroclinic connection established in Lemma 14 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 (4). In other words, the Hamiltonian $\tilde{H}_0^{12}$ has a normally hyperbolic invariant manifold given by

$$\Lambda_0 = \left\{ \left(\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2\right) : (\xi, \eta) = (0, 0), \tilde{\gamma}_2 \in \mathbb{T}, \tilde{\Gamma}_2 \in [\zeta_1, \zeta_2] \right\}$$

where ζ_1, ζ_2 satisfy

$$0 < \zeta_1 < \zeta_2 < L_1 \sqrt{\frac{3}{5}}. \tag{49}$$

Moreover the stable and unstable manifolds of Λ_0 coincide.

5   Analysis of the inner dynamics

We can lift the normally hyperbolic invariant manifold Λ_0 to the full secular phase space by increasing its dimension to include the remaining secular variables $\tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3$, in order to obtain

$$\tilde{\Lambda}_0 = \left\{ \left(\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3 \right) : \xi = \eta = 0, \ \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3 \in \mathbb{T}, \ \tilde{\Gamma}_2 \in [\zeta_1, \zeta_2], \ \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3 \in [-1, 1] \right\}.$$

It is clear that this set remains a normally hyperbolic invariant manifold for H_0^{12}: the variables $\tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3$ are constant with respect to H_0^{12} so invariance follows; and the hyperbolicity in the normal directions ξ, η is preserved. In addition, the variables $\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3 \right)$ define coordinates on $\tilde{\Lambda}_0$.

In a neighbourhood of the normally hyperbolic cylinder $\tilde{\Lambda}_0$ the symplectic form is

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

and so the restriction to $\tilde{\Lambda}_0$ of Ω is

$$\Omega_0 = \Omega|_{\tilde{\Lambda}_0} = d\tilde{\Gamma}_2 \wedge d\tilde{\gamma}_2 + d\tilde{\Gamma}_3 \wedge d\tilde{\gamma}_3 + d\tilde{\Psi}_1 \wedge d\tilde{\psi}_1 + d\tilde{L}_3 \wedge d\tilde{\ell}_3. \tag{50}$$

The following theorem is the main result of this section.

Theorem 15. *For any $r \geq 2$ there is $L_2^* > 0$ such that for any $L_2 \geq L_2^*$ and $L_3 \gg L_2^3$ we have the following.*

1. *The flow of the secular Hamiltonian F_{sec} has a normally hyperbolic invariant manifold Λ that is $O\left(\frac{1}{L_2} \right)$-close to $\tilde{\Lambda}_0$ in the C^r topology. The restriction to Λ of the symplectic form Ω is closed and nondegenerate. The variables $\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3 \right)$ define coordinates on Λ, with respect to which $\Omega|_\Lambda$ is not necessarily in Darboux form.*

2. *Choose $k_1, k_2 \in \mathbb{N}$. There is a coordinate transformation*

$$\Phi : \left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3 \right) \longmapsto \left(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3, \hat{\ell}_3, \hat{L}_3 \right)$$

   on Λ that is $O\left(\varepsilon^3 \right)$-close to the identity in the C^r topology such that

$$\Omega|_\Lambda = d\hat{\Gamma}_2 \wedge d\hat{\gamma}_2 + d\hat{\Psi}_1 \wedge d\hat{\psi}_1 + d\hat{\Gamma}_3 \wedge d\hat{\gamma}_3 + d\hat{L}_3 \wedge d\hat{\ell}_3,$$

   and the restriction to Λ of the secular Hamiltonian F_{sec} is

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

   where $\hat{F}_0 = \varepsilon^6 \, \hat{c}_0 \, \hat{\Gamma}_2^2 + \varepsilon^7 \, \hat{h}_0 \left(\hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3; \varepsilon, \mu \right)$, where $\varepsilon = \frac{1}{L_2}$, $\mu = \frac{L_2}{L_3^*}$, and where $\hat{F}_j$ are uniformly bounded in the C^r topology as $\varepsilon, \mu \to 0$ for $j = 0, 1$.

The rest of the section is dedicated to the proof of Theorem 15. Fenichel theory guarantees the persistence of the normally hyperbolic manifold $\tilde{\Lambda}_0$ for F_{sec} [30, 31, 32]. In particular, the existence is guaranteed of a function $\rho : \tilde{\Lambda}_0 \to \mathbb{R}^2$ that is $O\left(L_2^{-1} \right)$ small in the C^r topology such that the set

$$\Lambda = \mathrm{graph}(\rho) = \left\{ (\rho(x), x) : x \in \tilde{\Lambda}_0 \right\} \tag{51}$$

is a normally hyperbolic invariant manifold for F_{sec}. The following lemma provides us with information regarding the order at which each secular variable appears in the Taylor expansion of ρ.

Lemma 16. *The function ρ admits a Taylor expansion of the form*

$$\rho = \frac{1}{L_2}\,\rho_0 + \frac{1}{L_2^2}\,\rho_1 + \frac{L_2^{10}}{(L_3^*)^6}\,\rho_2 + \frac{L_2^9}{(L_3^*)^6}\,\rho_3 + \frac{L_2^{10}}{(L_3^*)^7}\,\rho_4$$

with

$$
\begin{cases}
\rho_0 = & \rho_0\left(\tilde{\Gamma}_2\right) \\
\rho_1 = & \rho_1\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\Psi}_1\right) \\
\rho_2 = & \rho_2\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3\right) \\
\rho_3 = & \rho_2\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3\right) \\
\rho_4 = & \rho_2\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3\right)
\end{cases}
$$

where each ρ_j is uniformly bounded in the C^r topology as $L_2, L_3 \to \infty$ (with $L_3 \gg L_2^3$).

Proof. The proof is analogous to the proof of Lemma 21 in [13], and we do not repeat it here. $\square$

Lemma 17. *There is a coordinate transformation*

$$\Phi^{(1)} : \left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right) \longmapsto (\gamma_2', \Gamma_2', \gamma_3', \Gamma_3', \psi_1', \Psi_1', \ell_3', L_3') \tag{52}$$

on Λ that is $O\left(\frac{1}{L_2^3}\right)$ close to the identity satisfying

$$
\begin{cases}
\Gamma_2' = & \tilde{\Gamma}_2 + \frac{1}{L_2^3}\, P_0'\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right) \\[4pt]
\Psi_1' = & \tilde{\Psi}_1 + \frac{L_2^9}{(L_3^*)^6}\, P_1'\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right) \\[4pt]
\Gamma_3' = & \tilde{\Gamma}_3 + \frac{L_2^9}{(L_3^*)^6}\, P_2'\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right) \\[4pt]
L_3' = & \tilde{L}_3 + \frac{L_2^9}{(L_3^*)^6}\, P_3'\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right)
\end{cases}
$$

such that

$$\Omega|_\Lambda = d\Gamma_2' \wedge d\gamma_2' + d\Gamma_3' \wedge d\gamma_3' + d\Psi_1' \wedge d\psi_1' + dL_3' \wedge d\ell_3'.$$

Proof. Denote by $U \subset \mathbb{R}^2$ the domain of the Poincaré variables (ξ, η), and by $V \subset \mathbb{R}^4$ the domain of the actions $\left(\tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right)$. Define the inclusion map $P : \mathbb{T}^4 \times V \to U \times \left(\mathbb{T}^4 \times V\right)$ by $P(x) = (\rho(x), x)$ where ρ is the function that parametrises the normally hyperbolic cylinder Λ via (5). Then $\Omega_1 = P^*\Omega$ is the pullback to Λ of Ω in the coordinates (3.3).

The Liouville 1-form λ is given by $\lambda = \xi\,d\eta + \tilde{\Gamma}_2\,d\tilde{\gamma}_2 + \tilde{\Gamma}_3\,d\tilde{\gamma}_3 + \tilde{\Psi}_1\,d\tilde{\psi}_1 + \tilde{L}_3\,d\tilde{\ell}_3$, and we have $\Omega = d\lambda$. Define $\lambda_1 = P^*\lambda$. Then Ω_1 is exact, because $d\lambda_1 = d\left(P^*\lambda\right) = P^*d\lambda = P^*\Omega = \Omega_1$.

Denote by ρ_ξ, ρ_η the ξ, η components of ρ respectively, and by $\rho_{j,\xi}, \rho_{j,\eta}$ the ξ, η components of ρ_j respectively for each $j = 0, 1, 2, 3, 4$. Using (5) we see that

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

with

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

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

where we have used the fact that ρ_0 depends only on $\tilde{\Gamma}_2$. Then R_1 is of order $\frac{1}{L_2^3}$, and depends only on $\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\Psi}_1$, whereas R_2 is of order $\frac{L_2^9}{\left(L_3^*\right)^6}$ (i.e. the order of ρ_2 times the order of ρ_1) and depends on all of the secular variables. In what follows, we construct the coordinate transformation (17) in two steps: the first one eliminates R_1 from Ω_1, and the second eliminates R_2.

Suppose there is a coordinate transformation

$$h_0 : \left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3\right) \longmapsto \left(\gamma_2^*, \Gamma_2^*, \psi_1^*, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3\right) \tag{53}$$

that changes only the $\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1$ variables, such that

$$h_0^* \, \Omega' = \Omega_0 \quad \text{where} \quad \Omega' = \Omega_0 + R_1. \tag{54}$$

We use Moser's trick from his proof of Darboux's theorem: suppose $h_0 = \phi_{\hat{\varepsilon}}$ where ϕ_t is the time-t map of some nonautonomous vector field X_t and where $\hat{\varepsilon} = L_2^{-3}$. Upon differentiating (5) with respect to $\hat{\varepsilon}$ and using Cartan's magic formula we get

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

where $\mathcal{L}_{X_{\hat{\varepsilon}}}$ is the Lie derivative with respect to $X_{\hat{\varepsilon}}$, and $i_{X_{\hat{\varepsilon}}}$ is the contraction operator of $X_{\hat{\varepsilon}}$. By the same argument as for Ω_1 above, $\Omega' = d\lambda'$ is exact, and so, since $d\Omega' = 0$, we obtain

$$di_{X_{\hat{\varepsilon}}}\Omega' = -\frac{d}{d\hat{\varepsilon}}\Omega' = -\frac{d}{d\hat{\varepsilon}}d\lambda' = -d\left(\frac{d}{d\hat{\varepsilon}}\lambda'\right).$$

Observe that this equation is satisfied by vector fields X_t for which

$$i_{X_{\hat{\varepsilon}}}\Omega' = -\frac{d}{d\hat{\varepsilon}}\lambda'.$$

By inverting Ω' this can be solved explicitly for X_t. Its flow exists at least for a time $\hat{\varepsilon}$, and its time-$\hat{\varepsilon}$ map gives the required map h_0 as in (5). Indeed, the argument in Appendix A implies that this coordinate transformation does not affect $\tilde{\Psi}_1$, because ρ_0, ρ_1 do not depend on $\tilde{\psi}_1$.

In the new coordinates, the symplectic form Ω_1 becomes $\hat{\Omega}_1 = \Omega_0 + \hat{R}_2$ where $\hat{R}_2 = h_0^* R_2$. Since R_2 is of order $\frac{L_2^9}{\left(L_3^*\right)^6}$, we may repeat the above procedure with $\hat{\varepsilon} = \frac{L_2^9}{\left(L_3^*\right)^6}$ to complete the proof of the lemma. $\quad\square$

Lemma 18. *Choose $k_1, k_2 \in \mathbb{N}$. There is a symplectic coordinate transformation*

$$\Phi^{(2)} : (\gamma_2', \Gamma_2', \gamma_3', \Gamma_3', \psi_1', \Psi_1', \ell_3', L_3') \longmapsto \left(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\gamma}_3, \hat{\Gamma}_3, \hat{\psi}_1, \hat{\Psi}_1, \hat{\ell}_3, \hat{L}_3\right)$$

on Λ that is $O\left(\varepsilon^3\right)$ close to the identity, such that the restriction $\hat{F} = F_{\text{sec}}|_\Lambda$ of the secular Hamiltonian to Λ becomes

$$\hat{F} = \hat{F}_0 \left(\hat{\Gamma}_2, \hat{\Gamma}_3, \hat{\Psi}_1, \hat{L}_3; \varepsilon, \mu\right) + \varepsilon^{k_1} \mu^{k_2} F_1 \left(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\gamma}_3, \hat{\Gamma}_3, \hat{\psi}_1, \hat{\Psi}_1, \hat{\ell}_3, \hat{L}_3; \varepsilon, \mu\right)$$

where $\hat{F}_0 = \varepsilon^6 c_0 \hat{\Gamma}_2^2 + \varepsilon^7 \hat{h}_0 \left(\hat{\Gamma}_2, \hat{\Gamma}_3, \hat{\Psi}_1, \hat{L}_3; \varepsilon, \mu\right)$, where $\varepsilon = \frac{1}{L_2}$, $\mu = \frac{L_2}{L_3^*}$, and where the C^2 norm of $\hat{F}$ is uniformly bounded in ε, μ for $j = 0, 1$. Moreover we can approximate the transformations of the actions at

first order by

$$\hat{\Gamma}_2 = \Gamma_2' + O\left(\frac{1}{L_2^3}\right) \tag{55}$$

$$\hat{\Gamma}_3 = \Gamma_3' + O\left(\frac{L_2^3}{(L_3^*)^2}\right)$$

$$\hat{\Psi}_1 = \Psi_1' - \frac{L_2^{11}}{(L_3^*)^6}\frac{\alpha_0^{23}}{\alpha_1^{12}}\frac{\left(1+\sqrt{1-\delta_2^2}\,\cos v_3'\right)^3}{2\left(3\frac{(\Gamma_2')^2}{L_1^2}-1\right)}\left[A_0\left(\gamma_3',v_3'\right)\cos 2\psi_1' + B_0\left(\gamma_3',v_3'\right)\sin 2\psi_1'\right] + \cdots \tag{56}$$

$$\hat{L}_3 = L_3' - \frac{L_2^4}{(L_3^*)^3}\frac{\alpha_0^{23}}{\alpha_{\mathrm{Kep}}}\left(\left(1+\sqrt{1-\delta_2^2}\,\cos v_3'\right)^3\left[\beta_0\,\sin^2\left(v_3'-\gamma_3'\right)+\beta_1\right]-\delta_2^3\left(\frac{\beta_0}{2}+\beta_1\right)\right) + \cdots \tag{57}$$

where the trigonometric polynomials A_0, B_0 are defined by (12), (12), the constants β_j are defined by

$$\beta_0 = \frac{1}{2}\left(9\,\delta_1^2 - 9\,\delta_3^2 + 15\,\frac{\delta_3^2}{\delta_1^2} - 15\right), \quad \beta_1 = \frac{1}{2}\left(5-3\,\delta_1^2\right), \tag{58}$$

and where v_3' is the true anomaly corresponding to the mean anomaly ℓ_3'.

To prove this lemma, we need the following estimate on the average of the secular Hamiltonian with respect to its fastest angles.

Lemma 19. *Consider the Hamiltonian F_{sec} defined in (3.3). Its average $\langle F_{\mathrm{sec}}\rangle_{(\tilde{\psi}_1,\tilde{\gamma}_2,\ell_3)}$ with respect to the angles $\tilde{\psi}_1,\tilde{\gamma}_2,\ell_3$ does not depend on $\tilde{\gamma}_3$ until the term of order $O\left(\frac{L_2^6}{L_3^{*8}}\right)$.*

The proof of this lemma is postponed to Appendix B.

Proof of Lemma 18. Since the frequencies of the angles all have different size the averaging procedure is rather standard, since there are no small divisors. We explicitly compute the first order of the symplectic coordinate transformation $\Phi^{(2)}$ in the actions Ψ_1', L_3' and approximate the order of the transformation in Γ_2'. In order to do this we search for a Hamiltonian J such that $\phi^1 = \Phi^{(2)}$ where ϕ^t is the Hamiltonian flow of J. In this case $\Phi^{(2)}$ is necessarily symplectic. First of all, notice that since F_{quad}^{12} does not depend on γ_2, and since the restriction to Λ of F_{oct}^{12} vanishes at first order, the first order term that could contain γ_2' is the term of order ε^9. Since the first order term in the expansion of F_{quad}^{12} is $\varepsilon^6 (\Gamma_2)'^2$ times a nontrivial constant, we can easily find a Hamiltonian J_0 of order ε^3 whose time-1 map gives a symplectic coordinate transformation that averages γ_2' from the term of order ε^9, hence (18).

Recall the first order terms containing $\tilde{\Psi}_1$, $\tilde{\psi}_1$, respectively, are H_1^{12}, H_0^{23}. Moreover the coefficients of H_1^{12}, H_0^{23} respectively in the expansion of the secular Hamiltonian are $\varepsilon^7\alpha_1^{12}$, $\varepsilon^2\mu^6\alpha_0^{23}$. Notice that, up to higher-order terms (and a multiplicative constant equal to L_1^2 in the case of h_0^{12}), the restrictions H_0^{12}, H_1^{12}, respectively, are equal to h_0^{12}, h_1^{12} where

$$h_0^{12} = (\Gamma_2')^2, \quad h_1^{12} = H_1^{12}|_\Lambda = \left(3\frac{(\Gamma_2')^2}{L_1^2}-1\right)\Psi_1' + 2\Gamma_2'.$$

Moreover the restriction to Λ of H_0^{23} is equal to H_0^{23}, up to higher order terms. Let $c_0 = L_1^{-2}\alpha_2^{23}$ and define the Hamiltonian $\hat{H}_0 = \varepsilon^6 c_0 h_0^{12} + \varepsilon^7\alpha_1^{12}h_1^{12} + \varepsilon^2\mu^6\alpha_0^{23}H_0^{23}$, a truncated version of $F_{\mathrm{sec}}|_\Lambda$. We now search for a Hamiltonian $J_1 = \varepsilon^{-5}\mu^6\frac{\alpha_0^{23}}{\alpha_1^{12}}\hat{J}_1$ such that, denoting by ϕ_1^t the Hamiltonian flow of J_1, the coordinate transformation ϕ_1^1 eliminates ψ_1' from H_0^{23}. Assuming J_1 does not depend on γ_2', we have

$$\begin{aligned}
\hat{H}_0 \circ \phi_1^{-1} &= \varepsilon^6 c_0 h_0^{12}\circ\phi_1^{-1} + \varepsilon^7\alpha_1^{12}h_1^{12}\circ\phi_1^{-1} + \varepsilon^2\mu^6\alpha_0^{23}H_0^{23}\circ\phi_1^{-1}\\
&= \varepsilon^6 c_0 h_0^{12} + \varepsilon^7\alpha_1^{12}\left(h_1^{12}-\{h_1^{12},J_1\}\right) + \varepsilon^2\mu^6\alpha_0^{23}H_0^{23}+\cdots\\
&= \varepsilon^6 c_0 h_0^{12} + \varepsilon^7\alpha_1^{12}h_1^{12} + \varepsilon^2\mu^6\alpha_0^{23}\left(H_0^{23}-\left\{h_1^{12},\hat{J}_1\right\}\right)+\cdots
\end{aligned} \tag{59}$$

24

Integrating (12), we see that

$$\left\langle H_0^{23}\right\rangle_{\psi_1'} = \frac{1}{2\pi}\int_{\mathbb{T}} H_0^{23}\, d\psi_1' = \left(1 + \sqrt{1-\delta_2^2}\,\cos v_3'\right)^3 \left[\frac{1}{2}A_0(\gamma_3', v_3') + C_0(\gamma_3', v_3')\right]. \tag{60}$$

We would like $\hat{J}_1$ to satisfy

$$\left\{h_1^{12}, \hat{J}_1\right\} = H_0^{23} - \left\langle H_0^{23}\right\rangle_{\psi_1'} = \frac{1}{2}\left(1 + \sqrt{1-\delta_2^2}\,\cos v_3'\right)^3 \left[A_0(\gamma_3', v_3')\cos 2\psi_1' + B_0(\gamma_3', v_3')\sin 2\psi_1'\right]. \tag{61}$$

Choose

$$\hat{J}_1 = \frac{1}{4}\left(3\frac{(\Gamma_2')^2}{L_1^2} - 1\right)^{-1}\left(1 + \sqrt{1-\delta_2^2}\,\cos v_3'\right)^3 \left[A_0(\gamma_3', v_3')\sin 2\psi_1' - B_0(\gamma_3', v_3')\cos 2\psi_1'\right].$$

Since $\hat{J}_1$ does not depend on γ_2', neither does J_1, and so (5) is satisfied; moreover, it can easily be checked that this choice of $\hat{J}_1$ solves (5). Hamilton's equations of motion imply that

$$\hat{\Psi}_1 = \Psi_1' - \frac{\partial J_1}{\partial \psi_1'} + \cdots = \Psi_1' - \frac{L_2^{11}}{(L_3^*)^6}\frac{\alpha_0^{23}}{\alpha_1^{12}}\frac{\partial \hat{J}_1}{\partial \psi_1'} + \cdots$$

Upon differentiating $\hat{J}_1$ with respect to ψ_1', we thus obtain (18).

The next step of the proof comprises the construction of a Hamiltonian J_2 such that the time-1 map of its flow ϕ_2^t averages the angle ℓ_3' from the secular Hamiltonian at first order by adjusting L_3'. The first appearance of L_3' in the secular Hamiltonian is the term $(L_3^*)^{-3}\alpha_{\mathrm{Kep}}\tilde{L}_3$ where the constant α_{Kep} is defined by (3.3), whereas ℓ_3' first appears via v_3' in the term $\varepsilon^2\mu^6\alpha_0^{23}\left\langle H_0^{23}\right\rangle_{\psi_1'}$, defined by (5). The average of $\left\langle H_0^{23}\right\rangle_{\psi_1'}$ with respect to ℓ_3' is

$$\left\langle H_0^{23}\right\rangle_{\psi_1', \ell_3'} = \frac{1}{2\pi}\int_{\mathbb{T}}\left\langle H_0^{23}\right\rangle_{\psi_1'} d\ell_3' = \frac{\delta_2^3}{2\pi}\int_{\mathbb{T}}\left(1 + \sqrt{1-\delta_2^2}\,\cos v_3'\right)\left[\beta_0\sin^2(v_3' - \gamma_3') + \beta_1\right]dv_3' + \cdots$$
$$= \delta_2^3\left(\frac{\beta_0}{2} + \beta_1\right) + \cdots$$

where the constants β_j are defined by (18), and where we have used the formula

$$d\ell_3' = d\ell_3 + \cdots = (1 + e_3\cos v_3)^{-2}\left(\frac{\Gamma_3}{L_3}\right)^3 dv_3 = \delta_2^3\left(1 + \sqrt{1-\delta_2^2}\,\cos v_3'\right)^{-2} dv_3' + O\left(\frac{1}{L_3^*}\right), \tag{62}$$

which comes from Kepler's second law. Therefore, writing $J_2 = (L_3^*)^3\varepsilon^2\mu^6\frac{\alpha_0^{23}}{\alpha_{\mathrm{Kep}}}\hat{J}_2$, and assuming that $\hat{J}_2$ does not depend on γ_2', by similar reasoning to (5), we search for a function $\hat{J}_2$ satisfying

$$\frac{\partial \hat{J}_2}{\partial \ell_3'} = \left\{L_3', \hat{J}_2\right\} = \left\langle H_0^{23}\right\rangle_{\psi_1'} - \left\langle H_0^{23}\right\rangle_{\psi_1', \ell_3'} + \cdots$$
$$= \left(1 + \sqrt{1-\delta_2^2}\,\cos v_3'\right)^3\left[\beta_0\sin^2(v_3' - \gamma_3') + \beta_1\right] - \delta_2^3\left(\frac{\beta_0}{2} + \beta_1\right) + \cdots$$

We could find $\hat{J}_2$ explicitly by integration; however, since we only need to know

$$\hat{L}_3 = L_3' - \frac{\partial J_2}{\partial \ell_3'} + \cdots = L_3' - \frac{L_2^4}{(L_3^*)^3}\frac{\alpha_0^{23}}{\alpha_{\mathrm{Kep}}}\frac{\partial \hat{J}_2}{\partial \ell_3'} + \cdots$$

we have already established the approximation (18).

Finally, we estimate the order of the coordinate transformation in Γ_3', by averaging γ_3' from the Hamiltonian. Suppose we have first averaged out the angles $\gamma_2', \psi_1', \ell_3'$. Observe that the Hamiltonian of this

25

coordinate transformation itself depends on γ_3', and so produces a transformation in Γ_3' that is order $\frac{L_2^4}{(L_3^*)^3}$ close to the identity (as in the derivation of (18)). After making this coordinate transformation, the resulting Hamiltonian F_{sec}' still depends on γ_3'. By Lemma 19, $\langle F_{\text{sec}}\rangle_{(\tilde\psi_1,\tilde\gamma_2,\ell_3)}$ does not depend on γ_3' until terms of order $O\left(\frac{L_2^6}{L_3^{*8}}\right)$. Since the frequency of γ_3' is of order $\frac{L_2^3}{\left(L_3^*\right)^6}$, the coordinate transformation that averages γ_3' from $\langle F_{\text{sec}}\rangle_{(\tilde\psi_1,\tilde\gamma_2,\ell_3)}$ is therefore $\frac{L_2^3}{\left(L_3^*\right)^2}$ close to the identity, as is the transformation that averages γ_3' from F_{sec}'. It follows that the coordinate transformation in Γ_3' is therefore $\max\left\{\frac{L_2^4}{(L_3^*)^3},\frac{L_2^3}{\left(L_3^*\right)^2}\right\}=\frac{L_2^3}{\left(L_3^*\right)^2}$ close to the identity, which completes the proof of the lemma.

$\square$

6   Computation of the scattering map

Now that we have analysed the inner dynamics on the normally hyperbolic invariant cylinder Λ, we must analyse its invariant manifolds, their transverse intersections and the corresponding dynamics. The next theorem, which describes the so-called outer dynamics associated to Λ and its invariant manifolds, is the main result of this section.

Theorem 20. *The stable and unstable invariant manifolds of the normally hyperbolic invariant manifold Λ intersect transversely along two homoclinic channels, giving rise to two scattering maps $S_\pm : \Lambda \to \Lambda'$ such that*

$$S_\pm : \left(\hat\gamma_2,\hat\Gamma_2,\hat\gamma_3,\hat\Gamma_3,\hat\psi_1,\hat\Psi_1,\hat\ell_3,\hat L_3\right) \longmapsto \left(\hat\gamma_2^*,\hat\Gamma_2^*,\hat\gamma_3^*,\hat\Gamma_3^*,\hat\psi_1^*,\hat\Psi_1^*,\hat\ell_3^*,\hat L_3^*\right)$$

with

$$\begin{cases} \hat\Psi_1^* = & \hat\Psi_1 + \dfrac{L_2^9}{(L_3^*)^6}\mathcal{S}_1^\pm\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3,\hat\Gamma_2\right) + \cdots \\[4mm] \hat\Gamma_3^* = & \hat\Gamma_3 + \dfrac{L_2^9}{(L_3^*)^6}\mathcal{S}_2^\pm\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3,\hat\Gamma_2\right) + \cdots \\[4mm] \hat L_3^* = & \hat L_3 + \dfrac{L_2^9}{(L_3^*)^6}\mathcal{S}_3^\pm\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3,\hat\Gamma_2\right) + \cdots \end{cases}$$

where

$$\mathcal{S}_1^\pm\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3,\hat\Gamma_2\right) = \mp\,\alpha_1^{23}\left(1+\sqrt{1-\delta_2^2}\cos\hat v_3\right)^3 \kappa\left(\frac{\pi\hat\Gamma_2}{A_2\,L_1^2}\right)\frac{\partial B_1}{\partial\hat\psi_1}\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3\right) + \frac{\alpha_0^{23}\,\alpha_2^{12}}{\alpha_1^{12}}\frac{L_1}{6}\sqrt{\frac{3}{2}}\hat\Gamma_2$$

$$\times \sqrt{1-\frac{5}{3}\frac{\hat\Gamma_2^2}{L_1^2}}\,\frac{\left(1+\sqrt{1-\delta_2^2}\cos\hat v_3\right)^3}{\left(3\frac{\hat\Gamma_2^2}{L_1^2}-1\right)}\left[A_0\left(\hat\gamma_3,\hat\ell_3\right)\sin 2\hat\psi_1 - B_0\left(\hat\gamma_3,\hat\ell_3\right)\cos 2\hat\psi_1\right] \tag{63}$$

$$\mathcal{S}_2^\pm\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3,\hat\Gamma_2\right) = \mp\,\alpha_1^{23}\left(1+\sqrt{1-\delta_2^2}\cos\hat v_3\right)^3 \kappa\left(\frac{\pi\hat\Gamma_2}{A_2\,L_1^2}\right)\frac{\partial B_1}{\partial\hat\gamma_3}\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3\right) \tag{64}$$

$$\mathcal{S}_3^\pm\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3,\hat\Gamma_2\right) = \mp\,\alpha_1^{23}\,\delta_2^{-3}\left(1+\sqrt{1-\delta_2^2}\cos\hat v_3\right)^4 \kappa\left(\frac{\pi\hat\Gamma_2}{A_2\,L_1^2}\right)\left[-3\sqrt{1-\delta_2^2}\sin\hat v_3\right.$$

$$\left.\times B_1\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3\right) + \left(1+\sqrt{1-\delta_2^2}\cos\hat v_3\right)\frac{\partial B_1}{\partial\hat v_3}\left(\hat\psi_1,\hat\gamma_3,\hat\ell_3\right)\right] \tag{65}$$

where the function κ is defined by

$$\kappa(x) = \sqrt{\frac{2}{3}}\frac{L_1^2}{\chi}\left[1-\frac{x}{\sinh x}\right] \tag{66}$$

and the trigonometric polynomials A_0, B_0, B_1 were introduced in Lemma 12.

26

In Section 4 we established the existence of two hyperbolic periodic orbits $Z^0_{\min}, Z^0_{\max}$ for the Hamiltonian H^{12}_0 (see (4)); moreover we found that the stable and unstable manifolds of these saddles coincide along a heteroclinic trajectory Z^0 introduced in Lemma 14. Furthermore, in the Poincaré variables ξ, η (see (4)), the two saddles are reduced to a single hyperbolic periodic orbit, which we denote by Z^0_*, and the heteroclinic connection becomes a homoclinic connection to Z^0_*; for convenience we denote by Z^0 this homoclinic connection. Recall in Section 4 we considered the phase space of H^{12}_0 to be of dimension 4, as the Hamiltonian H^{12}_0 depends only on the variables $\gamma_1, \Gamma_1, \tilde{\Gamma}_2$ (or equivalently on $\xi, \eta, \tilde{\Gamma}_2$). However we can lift the dynamics to the full secular phase space in the obvious natural way, by setting all remaining secular variables (i.e. $\tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3$) to be constant with respect to H^{12}_0. Considering Z^0_*, Z^0 as functions of all variables in this way, we obtain a parametrisation of the normally hyperbolic invariant manifold $\tilde{\Lambda}_0$ defined at the beginning of Section 5, and a parametrisation of the separatrix as in Lemma 14.

Write $\tilde{F}^{12}_{\text{quad}} = L^6_2 F^{12}_{\text{quad}}$. Then, the integrable Hamiltonian $\tilde{F}^{12}_{\text{quad}}$ possesses a hyperbolic periodic orbit Z^{quad}_* that is $O(L_2^{-1})$ close to Z^0_*; moreover there is a homoclinic orbit Z^{quad} to Z^{quad}_* that is $O(L_2^{-1})$ close to Z^0. Since $\tilde{F}^{12}_{\text{quad}}$ is integrable (see Lemma 10), the homoclinic trajectory Z^{quad} corresponds to a non-transverse homoclinic intersection of the stable and unstable manifolds of Z^{quad}_*.

Denote by $\bar{H} = L^6_2 F_{\text{sec}} - \tilde{F}^{12}_{\text{quad}}$ the Hamiltonian of the perturbation, and define the Poincaré-Melnikov potential by

$$\mathcal{L}\left(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3\right) = \int_{-\infty}^{\infty} \left[\bar{H}\left(Z^{\text{quad}}\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3\right)\right) \right.$$
$$\left. - \bar{H}\left(Z^{\text{quad}}_*\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3\right)\right) \right] dt. \tag{67}$$

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 $\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3$ appears in the expansion of $\mathcal{L}$.

Proposition 21. *The expansion of the Poincaré-Melnikov potential $\mathcal{L}$ satisfies the following properties, where the notation α^{ij}_k, H^{ij}_k is as in Proposition 7.*

1. *The first nontrivial term in the expansion of $\mathcal{L}$ is $\frac{1}{L^2_2}\alpha^{12}_2\mathcal{L}^{12}_2$ where*

$$\mathcal{L}^{12}_2\left(\tilde{\gamma}_2, \tilde{\Gamma}_2\right) = \int_{-\infty}^{\infty} \left(H^{12}_2\left(Z^0\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right)\right) \right.$$
$$\left. - H^{12}_2\left(Z^0_*\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right)\right) \right) dt$$
$$= \tilde{\mathcal{L}}^{12}_2\left(\tilde{\Gamma}_2\right)\sin\tilde{\gamma}_2$$

and where $\tilde{\mathcal{L}}^{12}_2$ is an analytic function of $\tilde{\Gamma}_2$ that is nonvanishing for $\tilde{\Gamma}_2 \in [\zeta_1, \zeta_2]$.

2. *The angles $\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3$ all appear for the first time in the expansion of $\mathcal{L}$ in the term $\frac{L^9_2}{(L^*_3)^6}\alpha^{23}_1\mathcal{L}^{23}_1$ where*

$$\mathcal{L}^{23}_1\left(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2\right) = \int_{-\infty}^{\infty} \left(H^{23}_1\left(Z^0\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right)\right) \right.$$
$$\left. - H^{23}_1\left(Z^0_*\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right)\right) \right) dt$$
$$= \left(1 + \sqrt{1 - \delta^2_2}\cos v_3\right)^3 \kappa\left(\frac{\pi\tilde{\Gamma}_2}{A_2 L^2_1}\right)$$
$$\times \left[A_1\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right)\cos\tilde{\gamma}_2 + B_1\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right)\sin\tilde{\gamma}_2\right]$$

where the function κ is defined by (20).

Proof. Part 1 of the proposition was proved in [27] (see also Proposition 27 and Appendix F of [13]). As for part 2, observe that, since $\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3$ are constant with respect to H_0^{12}, we can write

$$\mathcal{L}_1^{23}\left(\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2\right) = \left(1 + \sqrt{1-\delta_2^2}\,\cos v_3\right)^3 \left[A_1\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right)\mathcal{L}_1\left(\tilde{\gamma}_2, \tilde{\Gamma}_2\right) + B_1\left(\tilde{\gamma}_3, \tilde{\psi}_1, v_3\right)\mathcal{L}_2\left(\tilde{\gamma}_2, \tilde{\Gamma}_2\right)\right] \tag{68}$$

where

$$\mathcal{L}_j\left(\tilde{\gamma}_2, \tilde{\Gamma}_2\right) = \int_0^\infty \left(\mathcal{F}_j \circ Z^0\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right) - \mathcal{F}_j \circ Z_{\min}^0\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right)\right)dt+$$

$$+ \int_{-\infty}^0 \left(\mathcal{F}_j \circ Z^0\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right) - \mathcal{F}_j \circ Z_{\max}^0\left(t, \tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3\right)\right)dt,$$

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 $\tilde{\psi}_1, \tilde{\gamma}_3$. It was shown in Section 6.2 of [13] that we have

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

where κ is the function defined by (20). Combining these formulas with (6) completes the proof of the proposition. $\qquad\square$

Lemma 22. *The secular Hamiltonian has two homoclinic channels corresponding to the normally hyperbolic invariant manifold Λ, and there are two scattering maps defined globally on Λ by*

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

with

$$\tilde{\gamma}_2^* = \tilde{\gamma}_2 + \Delta_0^\pm\left(\tilde{\Gamma}_2; \cdots\right), \quad \tilde{\psi}_1^* = \tilde{\psi}_1 + \frac{1}{L_2^2}\Delta_1^\pm\left(\tilde{\Gamma}_2; \cdots\right),$$

$$\tilde{\gamma}_3^* = \tilde{\gamma}_3 + \frac{L_2^8}{(L_3^*)^6}\Delta_2^\pm\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right), \quad \tilde{\ell}_3^* = \tilde{\ell}_3 + \frac{L_2^9}{(L_3^*)^7}\Delta_3^\pm\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right),$$

$$\tilde{\Gamma}_2^* = \tilde{\Gamma}_2 + \frac{1}{L_2^3}\Theta_0^\pm\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right), \quad \tilde{\Psi}_1^* = \tilde{\Psi}_1 + \frac{L_2^9}{(L_3^*)^6}\Theta_1^\pm\left(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2; \cdots\right),$$

$$\tilde{\Gamma}_3^* = \tilde{\Gamma}_3 + \frac{L_2^9}{(L_3^*)^6}\Theta_2^\pm\left(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2; \cdots\right), \quad \tilde{L}_3^* = \tilde{L}_3 + \frac{L_2^9}{(L_3^*)^6}\Theta_3^\pm\left(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2; \cdots\right)$$

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

$$\Delta_0^\pm\left(\tilde{\Gamma}_2; \cdots\right) = 2\arctan\chi^{-1} + \cdots, \quad \Delta_1^\pm\left(\tilde{\Gamma}_2; \cdots\right) = \frac{L_1}{6}\sqrt{\frac{3}{2}}\,\alpha_2^{12}\tilde{\Gamma}_2\sqrt{1 - \frac{5}{3}\frac{\tilde{\Gamma}_2^2}{L_1^2}} + \cdots$$

$$\Delta_2^\pm\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right) = O(1), \quad \Delta_3^\pm\left(\tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\ell}_3, \tilde{L}_3\right) = O(1)$$

and

$$\Theta_0^\pm \left(\tilde\gamma_2, \tilde\Gamma_2, \tilde\gamma_3, \tilde\Gamma_3, \tilde\psi_1, \tilde\Psi_1, \tilde\ell_3, \tilde L_3 \right) = O(1),$$

$$\Theta_1^\pm \left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots \right) = \mp \alpha_1^{23} \left(1 + \sqrt{1 - \delta_2^2} \cos v_3 \right)^3 \kappa \left(\frac{\pi \tilde\Gamma_2}{A_2 \, L_1^2} \right) \frac{\partial B_1}{\partial \tilde\psi_1} \left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3 \right) + \cdots$$

$$\Theta_2^\pm \left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots \right) = \mp \alpha_1^{23} \left(1 + \sqrt{1 - \delta_2^2} \cos v_3 \right)^3 \kappa \left(\frac{\pi \tilde\Gamma_2}{A_2 \, L_1^2} \right) \frac{\partial B_1}{\partial \tilde\gamma_3} \left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3 \right) + \cdots,$$

$$\Theta_3^\pm \left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots \right) = \mp \alpha_1^{23} \delta_2^{-3} \left(1 + \sqrt{1 - \delta_2^2} \cos v_3 \right)^4 \kappa \left(\frac{\pi \tilde\Gamma_2}{A_2 \, L_1^2} \right) \left[-3 \sqrt{1 - \delta_2^2} \sin v_3 \right.$$
$$\left. \times B_1 \left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3 \right) + \left(1 + \sqrt{1 - \delta_2^2} \cos v_3 \right) \frac{\partial B_1}{\partial v_3} \left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3 \right) \right] + \cdots$$

Proof. Denote by $(\omega_0, \omega_1, \omega_2, \omega_3)$ the frequency vector of the angles $\left(\tilde\gamma_2, \tilde\psi_1, \tilde\gamma_3, \tilde\ell_3 \right)$ on a torus on Λ corresponding to fixed values of the actions $\tilde\Gamma_2, \tilde\Psi_1, \tilde\Gamma_3, \tilde L_3$ for the Hamiltonian $\tilde F = L_2^6 \left(F_{\text{quad}}^{12} + F_{\text{Kep}} - \tilde F_{\text{Kep}} \right)$. The computations of Section 3 therefore imply that

$$\omega_0 = 2\alpha_0^{12} \frac{\tilde\Gamma_2}{L_1^2} + O\left(\frac{1}{L_2} \right), \quad \omega_1 = O\left(\frac{1}{L_2} \right), \quad \omega_2 = 0, \quad \omega_3 = \frac{L_2^6}{(L_3^*)^3} \alpha_{\text{Kep}} + O\left(\frac{L_2^6}{(L_3^*)^4} \right).$$

Consider the function

$$\tau \longmapsto \mathcal{L} \left(\tilde\gamma_2 - \omega_0 \, \tau, \tilde\psi_1 - \omega_1 \, \tau, \tilde\gamma_3 - \omega_2 \, \tau, \tilde\ell_3 - \omega_3 \, \tau, \tilde\Gamma_2, \tilde\Psi_1, \tilde\Gamma_3, \tilde L_3 \right) \tag{69}$$

where $\mathcal{L}$ is the Poincaré-Melnikov potential defined by (6). Results of [18] imply that nondegenerate critical points of (6) correspond to transverse homoclinic intersections of the stable and unstable manifolds of Λ. Equation (6), Proposition 7, and Proposition 21 imply that

$$\mathcal{L} = \frac{1}{L_2^2} \alpha_2^{12} \mathcal{L}_2^{12} + \cdots$$

The function $\tau \mapsto \mathcal{L}_2^{12} \left(\tilde\gamma_2 - \omega_0 \, \tau, \tilde\Gamma_2 \right)$ has nondegenerate critical points $\tau_\pm$ where $\omega_0 \tau_\pm = \tilde\gamma_2 \pm \frac{\pi}{2}$. It follows that there are functions

$$\tau_\pm^* \left(\tilde\gamma_2, \tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2, \tilde\Psi_1, \tilde\Gamma_3, \tilde L_3 \right) = \frac{1}{\omega_0} \left(\tilde\gamma_2 \pm \frac{\pi}{2} \right) + \cdots$$

such that

$$\left. \frac{d}{d\tau} \right|_{\tau = \tau_\pm^*} \mathcal{L} \left(\tilde\gamma_2 - \omega_0 \, \tau, \tilde\psi_1 - \omega_1 \, \tau, \tilde\gamma_3 - \omega_2 \, \tau, \tilde\ell_3 - \omega_3 \, \tau, \tilde\Gamma_2, \tilde\Psi_1, \tilde\Gamma_3, \tilde L_3 \right) = 0.$$

We now introduce the reduced Poincaré-Melnikov potentials

$$\mathcal{L}_\pm^* \left(\tilde\gamma_2, \tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2, \tilde\Psi_1, \tilde\Gamma_3, \tilde L_3 \right) = \mathcal{L} \left(\tilde\gamma_2 - \omega_0 \, \tau_\pm^*, \tilde\psi_1 - \omega_1 \, \tau_\pm^*, \tilde\gamma_3 - \omega_2 \, \tau_\pm^*, \tilde\ell_3 - \omega_3 \, \tau_\pm^*, \tilde\Gamma_2, \tilde\Psi_1, \tilde\Gamma_3, \tilde L_3 \right).$$

Now, following again [18], the changes in the actions coming from the scattering maps $\tilde S_\pm$ are defined using the functions $\mathcal{L}_\pm^*$ via

$$\tilde\Gamma_2^* = \tilde\Gamma_2 + \frac{\partial \mathcal{L}_\pm^*}{\partial \tilde\gamma_2} + \cdots, \quad \tilde\Psi_1^* = \tilde\Psi_1 + \frac{\partial \mathcal{L}_\pm^*}{\partial \tilde\psi_1} + \cdots, \quad \tilde\Gamma_3^* = \tilde\Gamma_3 + \frac{\partial \mathcal{L}_\pm^*}{\partial \tilde\gamma_3} + \cdots, \quad \tilde L_3^* = \tilde L_3 + \frac{\partial \mathcal{L}_\pm^*}{\partial \tilde\ell_3} + \cdots$$

Note that the cylinder frequencies in the model in [18] 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 there are multiple time

29

scales and the angles $\tilde\psi_1$, $\tilde\gamma_3$, and $\tilde\ell_3$ appear only at higher order terms (see Proposition 7). Still, one can easily check that the statements in [18] are still valid in the present setting. The only difference is that the first order of the scattering maps in the actions $\tilde\Psi_1$, $\tilde\Gamma_3$, $\tilde L_3$ come from higher orders of the Melnikov potential.

By part 2 of Proposition 21, the first term in the expansion of the Poincaré-Melnikov potential depending on any of the angles $\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3$ is $\frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\mathcal{L}_1^{23}$, and in fact all three of those angles appear in this term. It follows that the first term in the expansion of the reduced Poincaré-Melnikov potentials depending on the angles $\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3$ is $\frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\left(\mathcal{L}_1^{23}\right)^*_\pm$ where

$$
\left(\mathcal{L}_1^{23}\right)^*_\pm = \mathcal{L}_1^{23}\left(\tilde\gamma_2 - \omega_0\,\tau_\pm^*, \tilde\psi_1 - \omega_1\,\tau_\pm^*, \tilde\gamma_3 - \omega_2\,\tau_\pm^*, \tilde\ell_3 - \omega_3\,\tau_\pm^*, \tilde\Gamma_2\right)
$$
$$
= \mathcal{L}_1^{23}\left(-\frac{\pi}{2}, \tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2\right) + \cdots
$$
$$
= \mp\left(1 + \sqrt{1-\delta_2^2}\cos v_3\right)^3 \kappa\left(\frac{\pi\tilde\Gamma_2}{A_2 L_1^2}\right) B_1\left(\tilde\gamma_3, \tilde\psi_1, \tilde\ell_3\right) + \cdots
$$

It follows that

$$
\frac{\partial\mathcal{L}^*_\pm}{\partial\tilde\psi_1} = \frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\frac{\partial\left(\mathcal{L}_1^{23}\right)^*_\pm}{\partial\tilde\psi_1} + \cdots = \mp\frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\left(1+\sqrt{1-\delta_2^2}\cos v_3\right)^3\kappa\left(\frac{\pi\tilde\Gamma_2}{A_2\,L_1^2}\right)\frac{\partial B_1}{\partial\tilde\psi_1}\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3\right) + \cdots,
$$
$$
\frac{\partial\mathcal{L}^*_\pm}{\partial\tilde\gamma_3} = \frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\frac{\partial\left(\mathcal{L}_1^{23}\right)^*_\pm}{\partial\tilde\gamma_3} + \cdots = \mp\frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\left(1+\sqrt{1-\delta_2^2}\cos v_3\right)^3\kappa\left(\frac{\pi\tilde\Gamma_2}{A_2\,L_1^2}\right)\frac{\partial B_1}{\partial\tilde\gamma_3}\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3\right) + \cdots,
$$

and

$$
\frac{\partial\mathcal{L}^*_\pm}{\partial\tilde\ell_3} = \frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\frac{\partial\left(\mathcal{L}_1^{23}\right)^*_\pm}{\partial\tilde\ell_3} + \cdots = \frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\frac{\partial\left(\mathcal{L}_1^{23}\right)^*_\pm}{\partial v_3}\frac{\partial v_3}{\partial\tilde\ell_3} + \cdots
$$
$$
= \mp\frac{L_2^9}{(L_3^*)^6}\alpha_1^{23}\delta_2^{-3}\left(1+\sqrt{1-\delta_2^2}\cos v_3\right)^4\kappa\left(\frac{\pi\tilde\Gamma_2}{A_2\,L_1^2}\right)
$$
$$
\times\left[-3\sqrt{1-\delta_2^2}\sin v_3\, B_1\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3\right) + \left(1+\sqrt{1-\delta_2^2}\cos v_3\right)\frac{\partial B_1}{\partial v_3}\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3\right)\right] + \cdots
$$

where we have used (5).

For the angles $\tilde\gamma_2, \tilde\psi_1, \tilde\gamma_3, \tilde\ell_3$, 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 shifts in the angles $\tilde\gamma_2, \tilde\psi_1$ come from F_{quad}^{12}, and are therefore the same as in [13] (see Lemma 30 and Appendix D). For the angles $\tilde\gamma_3$ and $\tilde\ell_3$, we simply estimate the order of the phase shift. Notice that the phase shift comes from the first appearance of the symplectic conjugate action in the Hamiltonian multiplying by functions of $\Gamma_1, \gamma_1, \tilde\gamma_2$, as these are the variables that behave differently on the separatrix and on Λ. Since $\tilde\Gamma_3$ first appears in the secular Hamiltonian in the term of order $\frac{L_2^3}{(L_3^*)^6}$, and since this term does not provide a phase shift in $\tilde\gamma_3$, the highest-order term that can potentially produce a phase shift has an additional factor of $\frac{1}{L_2}$. Since we normalise the entire secular Hamiltonian by L_2^6, we see that the phase shift in $\tilde\gamma_3$ is of order $O\left(\frac{L_2^3}{(L_3^*)^6}\frac{1}{L_2}L_2^6\right) = O\left(\frac{L_2^8}{(L_3^*)^6}\right)$. As for $\tilde\ell_3$, the phase shift comes from H_1^{23} upon expanding its coefficient

$$
\frac{L_2^3}{L_3^6} = \frac{L_2^3}{\left(L_3^* + \tilde L_3\right)^6} = \frac{L_2^3}{(L_3^*)^6} - 6\frac{L_2^3}{(L_3^*)^7}\tilde L_3 + \cdots
$$

in powers of $\frac{1}{\tilde L_3}$. Upon scaling this by L_2^6, we see that the first term that can provide a phase shift is of order $O\left(\frac{L_2^9}{(L_3^*)^7}\right)$, as required. $\qquad\square$

Lemma 23. *In the 'hat' coordinates, the scattering maps $S_\pm : \Lambda \to \Lambda'$, introduced in Lemma 22, are given by*

$$S_\pm : \left(\hat\gamma_2, \hat\Gamma_2, \hat\gamma_3, \hat\Gamma_3, \hat\psi_1, \hat\Psi_1, \hat\ell_3, \hat L_3\right) \longmapsto \left(\hat\gamma_2^*, \hat\Gamma_2^*, \hat\gamma_3^*, \hat\Gamma_3^*, \hat\psi_1^*, \hat\Psi_1^*, \hat\ell_3^*, \hat L_3^*\right)$$

with

$$\begin{cases} \hat\Psi_1^* = \hat\Psi_1 + \dfrac{L_2^9}{(L_3^*)^6}\mathcal{S}_1^\pm\left(\hat\psi_1, \hat\gamma_3, \hat\ell_3, \hat\Gamma_2\right) + \cdots \\[2ex] \hat\Gamma_3^* = \hat\Gamma_3 + \dfrac{L_2^9}{(L_3^*)^6}\mathcal{S}_2^\pm\left(\hat\psi_1, \hat\gamma_3, \hat\ell_3, \hat\Gamma_2\right) + \cdots \\[2ex] \hat L_3^* = \hat L_3 + \dfrac{L_2^9}{(L_3^*)^6}\mathcal{S}_3^\pm\left(\hat\psi_1, \hat\gamma_3, \hat\ell_3, \hat\Gamma_2\right) + \cdots \end{cases}$$

where $\mathcal{S}_j^\pm$ are given by (20), (20), and (20) respectively.

Proof. Recall we denote by Φ the coordinate transformation on Λ from 'tilde' coordinates to 'hat' coordinates constructed in Theorem 15. Moreover, by Lemma 18, we can write

$$\hat\Psi_1 = \tilde\Psi_1 + \frac{L_2^{11}}{(L_3^*)^6}\Phi_1\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots\right), \quad \hat\Gamma_3 = \tilde\Gamma_3 + \frac{L_2^3}{(L_3^*)^2}\Phi_2\left(\tilde\gamma_3, \tilde\ell_3; \cdots\right) \quad \hat L_3 = \tilde L_3 + \frac{L_2^4}{(L_3^*)^3}\Phi_3\left(\tilde\gamma_3, \tilde\ell_3; \cdots\right)$$

with

$$\Phi_1\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots\right) = -\frac{\alpha_0^{23}}{\alpha_1^{12}}\frac{\left(1 + \sqrt{1 - \delta_2^2}\,\cos\tilde v_3\right)^3}{2\left(3\frac{(\tilde\Gamma_2)^2}{L_1^2} - 1\right)}\left[A_0\left(\tilde\gamma_3, \tilde v_3\right)\cos 2\tilde\psi_1 + B_0\left(\tilde\gamma_3, \tilde v_3\right)\sin 2\tilde\psi_1\right] + \cdots$$

$$\Phi_3\left(\tilde\gamma_3, \tilde\ell_3; \cdots\right) = -\frac{\alpha_0^{23}}{\alpha_{\mathrm{Kep}}}\left(\left(1 + \sqrt{1 - \delta_2^2}\,\cos\tilde v_3\right)^3\left[\beta_0\,\sin^2\left(\tilde v_3 - \tilde\gamma_3\right) + \beta_1\right] - \delta_2^3\left(\frac{\beta_0}{2} + \beta_1\right)\right) + \cdots$$

and some function $\Phi_2\left(\tilde\gamma_3, \tilde\ell_3; \cdots\right)$, where the constants β_j are defined by (18). Therefore, using the notation of Lemma 22, we have

$$\hat\Psi_1^* = \tilde\Psi_1^* + \frac{L_2^{11}}{(L_3^*)^6}\Phi_1\left(\tilde\psi_1^*, \tilde\gamma_3^*, \tilde\ell_3^*, \tilde\Gamma_2^*; \cdots\right)$$

$$= \tilde\Psi_1 + \frac{L_2^9}{(L_3^*)^6}\Theta_1^\pm\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots\right)$$

$$\quad + \frac{L_2^{11}}{(L_3^*)^6}\Phi_1\left(\tilde\psi_1 + \frac{1}{L_2^2}\Delta_1^\pm\left(\tilde\Gamma_2; \cdots\right), \tilde\gamma_3 + O\left(\frac{L_2^8}{(L_3^*)^6}\right), \tilde\ell_3 + O\left(\frac{L_2^9}{(L_3^*)^7}\right), \tilde\Gamma_2 + O\left(\frac{1}{L_2^3}\right); \cdots\right)$$

$$= \hat\Psi_1 + \frac{L_2^9}{(L_3^*)^6}\left[\Theta_1^\pm\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots\right) + \partial_{\tilde\psi_1}\Phi_1\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots\right)\Delta_1^\pm\left(\tilde\Gamma_2; \cdots\right)\right] + \cdots$$

$$= \hat\Psi_1 + \frac{L_2^9}{(L_3^*)^6}\left[\Theta_1^\pm\left(\hat\psi_1, \hat\gamma_3, \hat\ell_3, \hat\Gamma_2; \cdots\right) + \partial_{\hat\psi_1}\Phi_1\left(\hat\psi_1, \hat\gamma_3, \hat\ell_3, \hat\Gamma_2; \cdots\right)\Delta_1^\pm\left(\hat\Gamma_2; \cdots\right)\right] + \cdots, \tag{70}$$

$$\hat\Gamma_3^* = \tilde\Gamma_3^* + \frac{L_2^3}{(L_3^*)^2}\Phi_2\left(\tilde\gamma_3^*, \tilde\ell_3^*; \cdots\right)$$

$$= \tilde\Gamma_3 + \frac{L_2^9}{(L_3^*)^6}\Theta_2^\pm\left(\tilde\psi_1, \tilde\gamma_3, \tilde\ell_3, \tilde\Gamma_2; \cdots\right) + \frac{L_2^3}{(L_3^*)^2}\Phi_2\left(\tilde\gamma_3 + O\left(\frac{L_2^8}{(L_3^*)^6}\right), \tilde\ell_3 + O\left(\frac{L_2^9}{(L_3^*)^7}\right); \cdots\right)$$

$$= \hat\Gamma_3 + \frac{L_2^9}{(L_3^*)^6}\Theta_2^\pm\left(\hat\psi_1, \hat\gamma_3, \hat\ell_3, \hat\Gamma_2; \cdots\right) + \cdots, \tag{71}$$

and

$$\hat{L}_3^* = \tilde{L}_3^* + \frac{L_2^4}{\left(L_3^*\right)^3} \Phi_3\left(\tilde{\gamma}_3^*, \tilde{\ell}_3^*; \cdots\right)$$

$$= \tilde{L}_3 + \frac{L_2^9}{\left(L_3^*\right)^6} \Theta_3^\pm\left(\tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3, \tilde{\Gamma}_2; \cdots\right) + \frac{L_2^4}{\left(L_3^*\right)^3} \Phi_3\left(\tilde{\gamma}_3 + O\left(\frac{L_2^8}{\left(L_3^*\right)^6}\right), \tilde{\ell}_3 + O\left(\frac{L_2^9}{\left(L_3^*\right)^7}\right); \cdots\right)$$

$$= \hat{L}_3 + \frac{L_2^9}{\left(L_3^*\right)^6} \Theta_3^\pm\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Gamma}_2; \cdots\right) + \cdots \tag{72}$$

Combining (6) and (6) with the formulas given for $\Theta_2^\pm$ and $\Theta_3^\pm$ in Lemma 22 already gives the expressions (20) and (20). From (6), the formula for $\Theta_1^\pm$ given in Lemma 22 and the formula above for Φ_1 we find that

$$\mathcal{S}_1^\pm\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Gamma}_2\right) = \Theta_1^\pm\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Gamma}_2\right) + \partial_{\hat{\psi}_1} \Phi_1\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Gamma}_2\right) \Delta_1^\pm\left(\hat{\Gamma}_2\right)$$

$$= \mp \alpha_1^{23}\left(1 + \sqrt{1 - \delta_2^2}\cos\hat{v}_3\right)^3 \kappa\left(\frac{\pi\hat{\Gamma}_2}{A_2 L_1^2}\right)\frac{\partial B_1}{\partial\hat{\psi}_1}\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right) + \frac{\alpha_0^{23}\,\alpha_2^{12}}{\alpha_1^{12}}\frac{L_1}{6}\sqrt{\frac{3}{2}}\hat{\Gamma}_2$$

$$\times \sqrt{1 - \frac{5}{3}\frac{\hat{\Gamma}_2^2}{L_1^2}}\frac{\left(1 + \sqrt{1 - \delta_2^2}\cos\hat{v}_3\right)^3}{\left(3\frac{(\hat{\Gamma}_2)^2}{L_1^2} - 1\right)}\left[A_0\left(\hat{\gamma}_3, \hat{\ell}_3\right)\sin 2\hat{\psi}_1 - B_0\left(\hat{\gamma}_3, \hat{\ell}_3\right)\cos 2\hat{\psi}_1\right],$$

which completes the proof of the lemma. $\qquad\square$

7   Shadowing in a Poincaré Section

In this section we prove Theorem 2 by proving the existence of suitable transition chains of almost invariant tori, and applying the shadowing results of [14] (which are summarised in Appendix D) to obtain orbits of the full four-body problem that visit arbitrarily small neighbourhoods of these tori in the prescribed order. As the shadowing results of [14] are proved for mappings, we first choose a suitable Poincaré section, and analyse the corresponding first return map; this is done is Section 7.1. In Section 7.2 we prove that we can find pseudo-orbits (that is orbits of the iterated function system consisting of the Poincaré map and the scattering maps) that connect (up to small errors) any given sequence of almost invariant tori on the normally hyperbolic cylinder. Finally in Section 7.3 we apply the results of [14] to complete the proof of Theorem 2.

7.1   Reduction to a Poincaré Map

Denote by $\tilde{\mathcal{D}}$ the region of the phase space where the coordinates $\left(\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3\right)$ take the following values: ξ, η live in the open ball in $\mathbb{R}^2$ centred at the origin of radius $\sqrt{2L_1}$; the angles $\tilde{\gamma}_2, \tilde{\psi}_1, \tilde{\gamma}_3, \tilde{\ell}_3 \in \mathbb{T}$; and the actions satisfy $\tilde{\Gamma}_2 \in [\zeta_1, \zeta_2]$ and $\tilde{\Psi}_1, \tilde{\Gamma}_3, \tilde{L}_3 \in [-1, 1]$, where the constants ζ_1, ζ_2 were defined in (4). We now make the following further refinement to these constants:

$$0 < \zeta_1 < \zeta_2 < \frac{L_1}{\sqrt{3}}. \tag{73}$$

This refinement guarantees that our inner map satisfies a twist condition; see Lemma 26 below. By slightly shrinking the region $\tilde{\mathcal{D}}$ if necessary, we obtain a region $\mathcal{D}$ in which there is a (not necessarily symplectic) near-identity coordinate transformation

$$\Upsilon : \left(\xi, \eta, \tilde{\gamma}_2, \tilde{\Gamma}_2, \tilde{\psi}_1, \tilde{\Psi}_1, \tilde{\gamma}_3, \tilde{\Gamma}_3, \tilde{\ell}_3, \tilde{L}_3\right) \longmapsto \left(\xi, \eta, \hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3, \hat{\ell}_3, \hat{L}_3\right)$$

where $\left(\hat{\gamma}_2, \hat{\Gamma}_2, \hat{\psi}_1, \hat{\Psi}_1, \hat{\gamma}_3, \hat{\Gamma}_3, \hat{\ell}_3, \hat{L}_3\right)$ are the coordinates provided by Theorem 15. We consider the range of energies $\mathcal{E} = \{F_{\text{sec}}(z) : z \in \mathcal{D}\} \subset \mathbb{R}$.

Theorem 24. *Choose $E_0 \in \mathcal{E}$ and consider the Poincaré section $M = \{\hat{\gamma}_2 = 0\} \cap \{F_{\mathrm{sec}} = E_0\} \cap \mathcal{D}$. We have the following.*

1. *The flow of F_{sec} gives rise to a well-defined Poincaré map $F : M \to M$, and the set $\hat{\Lambda} = \Lambda \cap M$ is a normally hyperbolic invariant manifold for F.*

2. *The variables $\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right)$ define coordinates on $\hat{\Lambda}$ and the inner map $f = F|_{\hat{\Lambda}}$ is of the form*

$$
f : \begin{cases} \left(\hat{\psi}_1^*, \hat{\gamma}_3^*, \hat{\ell}_3^*\right) & = \left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right) + g\left(\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right) + O\left(\varepsilon^{k_1 - 6} \mu^{k_2}\right) \\ \left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right) & = \left(\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right) + O\left(\varepsilon^{k_1 - 6} \mu^{k_2}\right) \end{cases} \tag{74}
$$

where $\varepsilon = \frac{1}{L_2}$, $\mu = \frac{L_2}{L_3}$, where k_1, k_2 come from part 2 of Theorem 15, and where

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

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

3. *There are two scattering maps $\hat{S}_{\pm} : \hat{\Lambda} \to \hat{\Lambda}'$ where $\hat{\Lambda}'$ is an open cylinder in $\mathbb{T}^3 \times \mathbb{R}^3$ containing $\hat{\Lambda}$. Moreover with $\left(\hat{\psi}_1^*, \hat{\gamma}_3^*, \hat{\ell}_3^*, \hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right) = \hat{S}_{\pm}\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right)$ we have*

$$
\begin{cases} \hat{\Psi}_1^* = & \hat{\Psi}_1 + \dfrac{L_2^9}{(L_3^*)^6} \mathcal{S}_1^{\pm}\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Gamma}_2\right) + \cdots \\[2mm] \hat{\Gamma}_3^* = & \hat{\Gamma}_3 + \dfrac{L_2^9}{(L_3^*)^6} \mathcal{S}_2^{\pm}\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Gamma}_2\right) + \cdots \\[2mm] \hat{L}_3^* = & \hat{L}_3 + \dfrac{L_2^9}{(L_3^*)^6} \mathcal{S}_3^{\pm}\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Gamma}_2\right) + \cdots \end{cases}
$$

where $\mathcal{S}_j^{\pm}$ are given by (20), (20), and (20) respectively.

Remark 25. *The inequality (7.1) gives us the true range of values of the mutual inclination i_{12} along the diffusive orbits we have found (see also Remark 13). Indeed, the computation $\cos i_{12} = \frac{\tilde{\Gamma}_2}{\tilde{\Gamma}_1} + O\left(L_2^{-1}\right)$ that we made in Remark 13 combined with the fact that $\tilde{\Gamma}_2 \in [\zeta_1, \zeta_2]$ where ζ_1, ζ_2 satisfy (7.1) implies that, on the circular ellipse $\{\Gamma_1 = L_1\}$, we have $|\cos i_{12}| < \sqrt{\frac{1}{3}} + O\left(L_2^{-1}\right)$, which means that i_{12} is more than roughly $55°$.*

The proof of parts 1 and 3 of Theorem 24 is equivalent to the proof of Lemmas 33 and 36 of [13], so we do not repeat it here. The proof of part 2 of the theorem is contained in Lemma 26 below.

Lemma 26. *The inner map $f = F|_{\hat{\Lambda}}$ has the form*

$$
f : \begin{cases} \left(\hat{\psi}_1^*, \hat{\gamma}_3^*, \hat{\ell}_3^*\right) & = \left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right) + g\left(\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right) + O\left(\varepsilon^{k_1 - 6} \mu^{k_2}\right) \\ \left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right) & = \left(\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right) + O\left(\varepsilon^{k_1 - 6} \mu^{k_2}\right) \end{cases} \tag{75}
$$

where $\varepsilon = \frac{1}{L_2}$, $\mu = \frac{L_2}{L_3}$, where k_1, k_2 come from part 2 of Theorem 15, and where

$$
\det Dg\left(\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right) \neq 0 \tag{76}
$$

as long as

$$
\hat{\Gamma}_2 \notin \left\{0, \frac{L_1}{\sqrt{3}}\right\}. \tag{77}
$$

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

Proof. The first order term of the normalised Hamiltonian $L_2^6\hat{F}$ is $c_0\hat{\Gamma}_2^2$ by Theorem 15, so the frequency of $\hat{\gamma}_2$ is $2c_0\hat{\Gamma}_2 + O(\varepsilon) = O(1)$. It follows that the return time to the Poincaré section $\{\hat{\gamma}_2 = 0\}$ of the flow of the Hamiltonian function $L_2^6\hat{F}$ is of order 1. Since we have averaged the angles on the cylinder from the Hamiltonian $L_2^6\hat{F}$ up to terms of order $O\left(\varepsilon^{k_1-6}\mu^{k_2}\right)$ (see Theorem 15), the Poincaré map f has the form (26).

Now, define

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

for $i = 0, 1, 2, 3$ where $P_0 = \hat{\Gamma}_2$, $P_1 = \hat{\Psi}_1$, $P_2 = \hat{\Gamma}_3$, $P_3 = \hat{L}_3$, and $P = (P_1, P_2, P_3)$. Restricting to an energy level $\{\hat{F}_0 = E_0\}$ with $E_0 \in \mathcal{E}$, the implicit function theorem implies that we can write $P_0 = \alpha(P)$ where

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

where we have defined $\hat{\omega}_i(P) = \omega_i(\alpha(P), P)$. It is not hard to see that $g_i(P) = \hat{\omega}_0(P)^{-1}\hat{\omega}_i(P)$. In the following computation, borrowed from Lemma 34 of [13], we suppress dependence of all functions on P and $P_0 = \alpha(P)$ for convenience of notation:

$$
\begin{aligned}
\hat{\omega}_0^3\left(Dg\right)_{ij} = \hat{\omega}_0^3\frac{\partial g_i}{\partial P_j} &= \hat{\omega}_0^2\left(\frac{\partial\omega_i}{\partial P_0}\frac{\partial\alpha}{\partial P_j} + \frac{\partial\omega_i}{\partial P_j}\right) - \hat{\omega}_0\,\hat{\omega}_i\left(\frac{\partial\omega_0}{\partial P_0}\frac{\partial\alpha}{\partial P_j} + \frac{\partial\omega_0}{\partial P_j}\right)\\
&= -\hat{\omega}_0\,\hat{\omega}_j\frac{\partial\omega_i}{\partial P_0} + \hat{\omega}_0^2\frac{\partial\omega_i}{\partial P_j} + \hat{\omega}_i\,\hat{\omega}_j\frac{\partial\omega_0}{\partial P_0} - \hat{\omega}_0\,\hat{\omega}_i\frac{\partial\omega_0}{\partial P_j}\\
&= -\frac{\partial\hat{F}_0}{\partial P_0}\frac{\partial\hat{F}_0}{\partial P_j}\frac{\partial^2\hat{F}_0}{\partial P_0\partial P_i} + \left(\frac{\partial\hat{F}_0}{\partial P_0}\right)^2\frac{\partial^2\hat{F}_0}{\partial P_j\partial P_i} + \frac{\partial\hat{F}_0}{\partial P_i}\frac{\partial\hat{F}_0}{\partial P_j}\frac{\partial^2\hat{F}_0}{\partial P_0^2} - \frac{\partial\hat{F}_0}{\partial P_0}\frac{\partial\hat{F}_0}{\partial P_i}\frac{\partial^2\hat{F}_0}{\partial P_j\partial P_0}.
\end{aligned}
$$

Inserting the formulas for the derivatives of $\hat{F}_0$ given in Appendix E into this formula, we see that

$$
\begin{array}{lll}
\hat{\omega}_0^3\left(Dg\right)_{11} = \varepsilon^{20}\hat{A}_{11}, & \hat{\omega}_0^3\left(Dg\right)_{12} = \varepsilon^{16}\mu^6\hat{A}_{12}, & \hat{\omega}_0^3\left(Dg\right)_{21} = \varepsilon^{16}\mu^6\hat{A}_{21},\\[4pt]
\hat{\omega}_0^3\left(Dg\right)_{22} = \varepsilon^{16}\mu^6\hat{A}_{22}, & \hat{\omega}_0^3\left(Dg\right)_{13} = \varepsilon^{16}\mu^3\hat{A}_{13}, & \hat{\omega}_0^3\left(Dg\right)_{23} = \varepsilon^{16}\mu^7\hat{A}_{23},\\[4pt]
\hat{\omega}_0^3\left(Dg\right)_{31} = \varepsilon^{16}\mu^3\hat{A}_{31}, & \omega_0^3\left(Dg\right)_{32} = \varepsilon^{16}\mu^7\hat{A}_{32}, & \omega_0^3\left(Dg\right)_{33} = \varepsilon^{16}\mu^4\hat{A}_{33}
\end{array}
$$

where $\hat{A}_{ij} = O(1)$ for each i, j, and where

$$\hat{A}_{11} = C_{12}^3\,\frac{54\left(L_1^2 - 3\hat{\Gamma}_2^2\right)\left(L_1^2 + \hat{\Gamma}_2^2\right)}{L_1^6\,\delta_1^{11}} + \cdots$$

$$\hat{A}_{22} = -C_{12}^2\,C_{23}\,\frac{36\,\hat{\Gamma}_2^2\left(12\,\delta_1^2 - 20\right)}{L_1^4\,\delta_1^8\,\delta_2^3} + \cdots$$

$$\hat{A}_{33} = -C_{12}^2\,\alpha_{\text{Kep}}\,\frac{108\,\hat{\Gamma}_2^2}{L_1^4\,\delta_1^6} + \cdots$$

Noticing that $\hat{\omega}_0^3 = O\left(\varepsilon^{18}\right)$, we define $A_{ij} = \varepsilon^{18}\hat{\omega}_0^{-3}\hat{A}_{ij}$. We thus obtain

$$Dg = \begin{pmatrix} \varepsilon^2 A_{11} & \frac{\mu^6}{\varepsilon^2}A_{12} & \frac{\mu^3}{\varepsilon^2}A_{13}\\[6pt] \frac{\mu^6}{\varepsilon^2}A_{21} & \frac{\mu^6}{\varepsilon^2}A_{22} & \frac{\mu^7}{\varepsilon^2}A_{23}\\[6pt] \frac{\mu^3}{\varepsilon^2}A_{31} & \frac{\mu^7}{\varepsilon^2}A_{32} & \frac{\mu^4}{\varepsilon^2}A_{33} \end{pmatrix}.$$

We must find a condition under which the determinant of Dg is nonzero, and in addition we must estimate the order of its bottom eigenvalue. We compute

$$\det Dg = \frac{\mu^{10}}{\varepsilon^2}A_{11}A_{22}A_{33} + \cdots$$

34

where we have used (2.2). It is not hard to see that the first-order approximations of the three eigenvalues of Dg are therefore given by the entries on the diagonal. Comparing their respective sizes using assumption (2.2), we thus see that the bottom eigenvalue is indeed of order $\frac{\mu^6}{\varepsilon^2}$. In addition we see that the nondegeneracy condition (26) is satisfied as long as

$$0 \neq \hat{A}_{11} \hat{A}_{22} \hat{A}_{33} = 209952\, C_{12}^6\, C_{23}\, \alpha_{\text{Kep}}\, \frac{\hat{\Gamma}_2^4}{L_1^{14}\, \delta_1^{25}\, \delta_2^3} \left(L_1^2 - 3\hat{\Gamma}_2^2\right)\left(L_1^2 + \hat{\Gamma}_2^2\right)\left(12\,\delta_1^2 - 20\right) + \cdots$$

which yields (26). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

7.2 Transition chains of almost invariant tori

We define a foliation of the normally hyperbolic manifold $\hat{\Lambda}$ via the leaves

$$\mathcal{L}\left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right) = \left\{\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right) \in \hat{\Lambda} : \hat{\Psi}_1 = \hat{\Psi}_1^*, \hat{\Gamma}_3 = \hat{\Gamma}_3^*, \hat{L}_3 = \hat{L}_3^*\right\} \tag{78}$$

for each $\left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right) \in [-1,1]^3$. Observe that each leaf $\mathcal{L}\left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right)$ is almost invariant under the inner map f, as it takes the form (2). A *transition chain* is a sequence $\{\mathcal{L}_j\}$ of such leaves such that for each j there is $\beta_j \in \{+,-\}$ and $z \in \mathcal{L}_j$ such that $\hat{S}_{\beta_j}(z) \in \mathcal{L}_{j+1}$ and

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

This is an essential condition of the shadowing theorems of [14] (see Appendix D).

Recall that the first order terms $\mathcal{S}_1^\pm, \mathcal{S}_2^\pm, \mathcal{S}_3^\pm$ of the images of the actions of a point $\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right)$ under the scattering maps $\hat{S}_\pm : \hat{\Lambda} \to \hat{\Lambda}'$ are given by (20), (20), (20) respectively as a result of Theorem 24. Let $z = \left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right) \in \mathbb{T}^3$. In what follows we fix some values of the actions, and suppress the dependence of the functions $\mathcal{S}_1^\pm, \mathcal{S}_2^\pm, \mathcal{S}_3^\pm$ on these actions for convenience of notation. Define the matrices $A^\pm(z)$ by

$$A^\pm(z) = \begin{pmatrix} \partial_{\hat{\psi}_1}\mathcal{S}_1^\pm(z) & \partial_{\hat{\gamma}_3}\mathcal{S}_1^\pm(z) & \partial_{\hat{\ell}_3}\mathcal{S}_1^\pm(z) \\ \partial_{\hat{\psi}_1}\mathcal{S}_2^\pm(z) & \partial_{\hat{\gamma}_3}\mathcal{S}_2^\pm(z) & \partial_{\hat{\ell}_3}\mathcal{S}_2^\pm(z) \\ \partial_{\hat{\psi}_1}\mathcal{S}_3^\pm(z) & \partial_{\hat{\gamma}_3}\mathcal{S}_3^\pm(z) & \partial_{\hat{\ell}_3}\mathcal{S}_3^\pm(z) \end{pmatrix}. \tag{79}$$

Lemma 27. *There are constants $\nu_j > 0$ and $C > 0$ for $j = 1,2,3$ such that for sufficiently large $L_2 \ll L_3^{\frac{1}{3}}$ and any leaf $\mathcal{L}^* = \mathcal{L}\left(\hat{\Psi}_1^0, \hat{\Gamma}_3^0, \hat{L}_3^0\right)$ of the foliation of $\hat{\Lambda}$, there is $\sigma \in \{+,-\}$ and there are open sets $U_j \subset \mathcal{L}^* \simeq \mathbb{T}^3$ for $j = 1,\ldots,8$ such that*

$$\det A^\sigma|_{\overline{U_j}} \neq 0 \tag{80}$$

where the matrices $A^\pm$ are defined by (7.2), $\mu(U_j) > C$ where μ is the Lebesgue measure on $\mathbb{T}^3$ and moreover:

1. *For all $z \in U_1$ we have $\mathcal{S}_1^\sigma(z) > \nu_1$, $\mathcal{S}_2^\sigma(z) > \nu_2$, and $\mathcal{S}_3^\sigma(z) > \nu_3$.*

2. *For all $z \in U_2$ we have $\mathcal{S}_1^\sigma(z) > \nu_1$, $\mathcal{S}_2^\sigma(z) > \nu_2$, and $\mathcal{S}_3^\sigma(z) < -\nu_3$.*

3. *For all $z \in U_3$ we have $\mathcal{S}_1^\sigma(z) > \nu_1$, $\mathcal{S}_2^\sigma(z) < -\nu_2$, and $\mathcal{S}_3^\sigma(z) > \nu_3$.*

4. *For all $z \in U_4$ we have $\mathcal{S}_1^\sigma(z) > \nu_1$, $\mathcal{S}_2^\sigma(z) < -\nu_2$, and $\mathcal{S}_3^\sigma(z) < -\nu_3$.*

5. *For all $z \in U_5$ we have $\mathcal{S}_1^\sigma(z) < -\nu_1$, $\mathcal{S}_2^\sigma(z) > \nu_2$, and $\mathcal{S}_3^\sigma(z) > \nu_3$.*

6. *For all $z \in U_6$ we have $\mathcal{S}_1^\sigma(z) < -\nu_1$, $\mathcal{S}_2^\sigma(z) > \nu_2$, and $\mathcal{S}_3^\sigma(z) < -\nu_3$.*

7. *For all $z \in U_7$ we have $\mathcal{S}_1^\sigma(z) < -\nu_1$, $\mathcal{S}_2^\sigma(z) < -\nu_2$, and $\mathcal{S}_3^\sigma(z) > \nu_3$.*

8. *For all $z \in U_8$ we have $\mathcal{S}_1^\sigma(z) < -\nu_1$, $\mathcal{S}_2^\sigma(z) < -\nu_2$, and $\mathcal{S}_3^\sigma(z) < -\nu_3$.*

Proof. Fix any leaf $\mathcal{L}^* = \mathcal{L}\left(\hat{\Psi}_1^0, \hat{\Gamma}_3^0, \hat{L}_3^0\right)$ of the foliation of $\hat{\Lambda}$, and for $z = \left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right) \in \mathbb{T}^3$, consider the map $\mathcal{S}^\pm : \mathbb{T}^3 \to \mathbb{R}^3$ given by $\mathcal{S}^\pm(z) = \left(\mathcal{S}_1^\pm(z), \mathcal{S}_2^\pm(z), \mathcal{S}_3^\pm(z)\right)$.

We prove Lemma 27 in two steps. The first step is to prove that, for each $(\sigma_1, \sigma_2, \sigma_3) \in \{+, -\}^3$, there exist points $z \in \mathcal{L}^*$ such that for $j = 1, 2, 3$, $S_j^\pm(z)$ is non zero and of sign σ_j. By continuity the same happens for open sets V_i, $i = 1 \ldots 8$ containing these points. Then, the second step is to prove that there exist open subsets $U_i \subset V_i$ where (27) is satisfied.

Step 1: We choose $\hat{\gamma}_3 = 0$ and $\hat{\ell}_3$ such that $\hat{v}_3 \in \{\pm\frac{\pi}{4}, \pm\frac{3\pi}{4}\}$, so that $\cos^2 \hat{v}_3 = \sin^2 \hat{v}_3$. Then, we obtain

$$\mathcal{S}_1^\pm = C_{1,1}^\pm \cos \hat{\psi}_1 + C_{1,2}^\pm \sin \hat{\psi}_1 \cos \hat{v}_3 \sin \hat{v}_3 + C_{1,3}^\pm \sin(2\hat{\psi}_1) + C_{1,4}^\pm \cos(2\hat{\psi}_1) \cos \hat{v}_3 \sin \hat{v}_3$$

$$\mathcal{S}_2^\pm = C_{2,1}^\pm \sin \hat{\psi}_1 \cos \hat{v}_3 \sin \hat{v}_3$$

$$\mathcal{S}_3^\pm = C_{3,1}^\pm \sin \hat{\psi}_1 \sin \hat{v}_3 + C_{3,2}^\pm \cos \hat{\psi}_1 \cos \hat{v}_3 + C_{3,3}^\pm \sin \hat{\psi}_1 \cos \hat{v}_3 \sin \hat{v}_3,$$

where $C_{i,j}^\pm$ for $(i,j) \neq (1,3)$ are non zero bounded functions whose sign does not depend on $\hat{\psi}_1$ and $\hat{v}_3$ and which are far from zero uniformly in $\hat{\psi}_1$ and $\hat{v}_3$ chosen as above. The function $C_{1,3}^\pm$ is bounded but can be zero.

We study these expressions when $\hat{\psi}_1 \to 0^\pm$ or $\pi^\pm$, and then $\sin \hat{\psi}_1$ is close to zero and of either sign. We have

$$\mathcal{S}_1^\pm = C_{1,1}^\pm \cos \hat{\psi}_1 + C_{1,4}^\pm \cos \hat{v}_3 \sin \hat{v}_3 + \text{h.o.t}$$

$$\mathcal{S}_2^\pm = C_2^\pm \sin \hat{\psi}_1 \cos \hat{v}_3 \sin \hat{v}_3 + +\text{h.o.t}$$

$$\mathcal{S}_3^\pm = C_{3,2}^\pm \cos \hat{\psi}_1 \cos \hat{v}_3 + +\text{h.o.t.}$$

Then, we can chose the sign of $\cos \hat{\psi}_1$ so that $C_{1,1}^\pm \cos \hat{\psi}_1$ is of sign σ_1. Then, we can chose the sign of $\mathcal{S}_3^\pm$, that is σ_3, by choosing the sign of $\cos \hat{v}_3$. After that, we chose the sign of $\sin \hat{v}_3$ to make sure that the second term in $\mathcal{S}_1^\pm$ has also sign σ_1. Finally, we can choose the sign of $\mathcal{S}_2^\pm$ by choosing the sign of $\sin \hat{\psi}_1$, which completes Step 1.

Step 2: In the second step, we prove that $\det A^\pm$ is non zero almost everywhere in $\mathcal{L}^*$ and, therefore, the sets $U_j \subset V_j$ introduced in Lemma 27 are nonempty . For fixed values of the actions, $\det A^\pm$ is an analytic function of the angles. Then, we just have to check that it is not zero for a particular choice of the angles.

To this end, we prove that the Fourier cofficient of the harmonic in $e^{i4\hat{\psi}_1}$ of the function $\det A^\pm$ (seen as a function of $\hat{\gamma}_3, \hat{v}_3$) is non zero almost everywhere. To this end, it is enough to find a point $(\hat{\gamma}_3, \hat{v}_3)$ on which this Fourier coefficient is non zero. We prove this for points with $\hat{v}_3 = 0$. Indeed, with this restriction, one can easily check that the wanted coefficient (which now only depends on $\hat{\gamma}_3$) is equal to a nonzero function of the actions times the Fourier coefficient associated to $e^{i4\hat{\psi}_1}$ of the determinant of the matrix

$$\begin{pmatrix} 2A_0 \cos 2\hat{\psi}_1 + 2B_0 \sin 2\hat{\psi}_1 & \partial_{\hat{\gamma}_3} A_0 \sin 2\hat{\psi}_1 - \partial_{\hat{\gamma}_3} B_0 \cos 2\hat{\psi}_1 & 0 \\ \frac{\partial^2 B_1}{\partial \hat{\psi}_1 \partial \hat{\gamma}_3} & \frac{\partial^2 B_1}{\partial \hat{\gamma}_3^2} & 0 \\ \frac{\partial^2 B_1}{\partial \hat{\psi}_1 \partial \hat{v}_3} & \frac{\partial^2 B_1}{\partial \hat{\gamma}_3 \partial \hat{v}_3} & B_1 \end{pmatrix}.$$

Therefore, it is enough to find a value of $\hat{\gamma}_3$ where the Fourier coefficient of the harmonic $e^{i4\hat{\psi}_1}$ of the function

$$B_1 \left[(2A_0 \cos 2\hat{\psi}_1 + 2B_0 \sin 2\hat{\psi}_1) \frac{\partial^2 B_1}{\partial \hat{\gamma}_3^2} - (\partial_{\hat{\gamma}_3} A_0 \sin 2\hat{\psi}_1 - \partial_{\hat{\gamma}_3} B_0 \cos 2\hat{\psi}_1) \frac{\partial^2 B_1}{\partial \hat{\psi}_1 \partial \hat{\gamma}_3} \right]$$

is non zero. One can check that this harmonic is equal to

$$h(\hat{\gamma}_3) = \left[-\frac{1}{2i} \beta \sin^2 \hat{\gamma}_3 + \frac{\delta_1 \beta}{2\delta_3} \cos \hat{\gamma}_3 \sin \hat{\gamma}_3 \right] \times$$

$$\left[\left(A_0 + \frac{B_0}{i} \right) \left(-\frac{\beta}{i}(\cos^2 \hat{\gamma}_3 - \sin^2 \hat{\gamma}_3) - \frac{2\delta_1 \beta}{\delta_3} \cos \hat{\gamma}_3 \sin \hat{\gamma}_3 \right) \right.$$

$$\left. - \left(-\frac{\eta}{i} \cos \hat{\gamma}_3 \sin \hat{\gamma}_3 - \frac{15\delta_3}{\delta_1}(1 - \delta_1^2)(\sin^2 \hat{\gamma}_3 - \cos^2 \hat{\gamma}_3) \right) \left(-\beta \cos \hat{\gamma}_3 \sin \hat{\gamma}_3 + \frac{\delta_1 \beta}{2i\delta_3}(\sin^2 \hat{\gamma}_3 - \cos^2 \hat{\gamma}_3) \right) \right]$$

where
$$\beta = \frac{\delta_3}{\delta_1} \left(5 - 4\delta_1^2\right) > 0 \qquad \text{and} \qquad \eta = 15\delta_1^2 \left(1 - \delta_1^2\right) \left(\delta_1^2 + \delta_3^2\right) > 0.$$

When $\hat{\gamma}_3 \to 0$, we have
$$h(\hat{\gamma}_3) \sim -\frac{15\delta_1}{4i\delta_3}(1 - \delta_1^2)\beta^2\hat{\gamma}_3 + O\left(\hat{\gamma}_3^2\right),$$

which is non zero for $\hat{\gamma}_3 \neq$ small enough. This completes the proof of Step 2.

$\square$

Lemma 28. *Suppose $z \in \mathcal{L}_0 \cap U_j$ and $z^* = \hat{S}_\sigma(z) \in \mathcal{L}_1$ where $\sigma \in \{+, -\}$ and the open sets U_j were found in Lemma 27, and where $\mathcal{L}_0$, $\mathcal{L}_1$ are leaves of the foliation of $\hat{\Lambda}$. Then $\hat{S}_\sigma$ maps $\mathcal{L}_0$ transversely across $\mathcal{L}_1$ at the point $z^* = \hat{S}_\sigma(z)$ in the sense that*

$$T_{z^*}\hat{\Lambda} = T_{z^*}\left(\hat{S}_\sigma\left(\mathcal{L}_0\right)\right) \oplus T_{z^*}\mathcal{L}_1.$$

Moreover the order of transversality (in the sense of Definition 39 in Appendix D) is $\frac{L_2^9}{\left(L_3^*\right)^6}$.

Proof. Suppose $z \in U_j$ for some $j = 1, \ldots, 8$. The tangent space $T_{z^*}\hat{\Lambda}$ is a real vector space of dimension 6, and its elements are of the form $v = (Q, P)$ where $Q \in \mathbb{R}^3$ represents tangents in the $(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3)$ directions, and $P \in \mathbb{R}^3$ represents tangents in the $(\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3)$ directions. The scattering map $\hat{S}_\sigma$ is smooth, and so $D_z\hat{S}_\sigma\left(T_z\mathcal{L}_0\right) = T_{z^*}\hat{S}_\sigma\left(\mathcal{L}_0\right)$. With $z = \left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3, \hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right)$ and $z^* = \left(\hat{\psi}_1^*, \hat{\gamma}_3^*, \hat{\ell}_3^*, \hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right)$ we can write

$$D_z\hat{S}_\sigma = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$$

where

$$A_1 = \frac{\partial\left(\hat{\psi}_1^*, \hat{\gamma}_3^*, \hat{\ell}_3^*\right)}{\partial\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right)}, \quad A_2 = \frac{\partial\left(\hat{\psi}_1^*, \hat{\gamma}_3^*, \hat{\ell}_3^*\right)}{\partial\left(\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right)}, \quad A_3 = \frac{\partial\left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right)}{\partial\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right)}, \quad A_4 = \frac{\partial\left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right)}{\partial\left(\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3\right)}.$$

Therefore

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \cdots, \quad A_3 = A^\sigma\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right) + \cdots$$

where the matrices $A^\pm$ are defined by (7.2).

Let $v_0 \in T_z\mathcal{L}_0$ and $v_1 \in T_{z^*}\mathcal{L}_1$. Since the leaves are defined as tori with constant actions, we have $v_j = (Q_j, 0)$ with $Q_j \in \mathbb{R}^3$ for $j = 0, 1$. Therefore we have

$$\begin{pmatrix} \bar{Q} \\ \bar{P} \end{pmatrix} = D_z\hat{S}_\sigma(v_0) + v_1 = \begin{pmatrix} Q_0 + Q_1 \\ A^\sigma\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right)Q_0 \end{pmatrix} + \cdots$$

By Lemma 27, the matrix $A^\sigma\left(\hat{\psi}_1, \hat{\gamma}_3, \hat{\ell}_3\right)$ is nonsingular. Thus by varying $v_0 \in T_z\mathcal{L}_0$ and $v_1 \in T_{z^*}\mathcal{L}_1$ we can obtain any tangent vector in $T_{z^*}\hat{\Lambda}$, which is precisely the transversality we want to prove. The fact that the order of transversality is $\frac{L_2^9}{\left(L_3^*\right)^6}$ follows from the fact that the order of the jumps in the scattering maps in the direction of each of the actions is $\frac{L_2^9}{\left(L_3^*\right)^6}$.

$\square$

7.3 Orbits of the four-body problem shadowing the transition chains

In this section we apply the shadowing theorems of [14] (see also Appendix D) to the secular Hamiltonian and to the Hamiltonian of the full four-body problem in order to complete the proof of Theorem 2. In fact, our first observation is that we have already proved that the Poincaré map $F : M \to M$ constructed in Theorem 24 satisfies assumptions [A1-3] of Theorem 40. Indeed, Theorem 24 implies that F has a normally hyperbolic invariant manifold $\hat{\Lambda} \simeq \mathbb{T}^3 \times [0,1]^3$, and that the inner map $f = F|_{\hat{\Lambda}}$ is a near-integrable twist map satisfying a non-uniform twist condition of order $\frac{L_2^8}{(L_3^*)^6}$ as per Definitions 37 and 38. In addition the stable and unstable manifolds of $\hat{\Lambda}$ have a transverse homoclinic intersection along (at least) two homoclinic channels, and the order of the transversality of the invariant manifolds is $\frac{1}{L_2^2}$. These homoclinic channels give rise to two scattering maps $\hat{S} : \hat{\Lambda} \to \hat{\Lambda}'$ due to Theorems 20 and 24. In Section 7.2 we constructed a foliation of $\hat{\Lambda}'$, the leaves of which are given by constant values of the actions, and are almost invariant under the inner map. In addition Lemmas 27 and 28 imply that, by iterating one of the scattering maps in appropriate neighbourhoods in the leaves of the foliation, we may move in any direction, connecting leaves that are $O\left(\frac{L_2^9}{(L_3^*)^6}\right)$ apart, and mapping leaves transversely across leaves.

In Section 3.3 (see Remark 6 in particular) we made a symplectic change of coordinates (3.3) to the 'tilde' variables, which were the basis for all further analysis in the paper. However this coordinate transformation is local, whereas the drift in eccentricity, inclination, and the semimajor axis described in Theorem 1 is global. In order to define these coordinates we introduced constants L_3^* and δ_j for $j = 1, 2, 3$. Here δ_2 is the coefficient of total angular momentum, and is therefore fixed for the secular system. The constants $L_3^*, \delta_1, \delta_3$ 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 $L_3^* \gg L_2^3$, $\delta_1 \in (0,1)$, and $\delta_3 \in (-1,1)$. Denote by $\tilde{C}_{L_3^*, \delta_1, \delta_3}$ the system of 'tilde' coordinates corresponding to the values $L_3^*, \delta_1, \delta_3$. Then the results of Section 4 apply in $\tilde{C}_{L_3^*, \delta_1, \delta_3}$ coordinates for each relevant value of $L_3^*, \delta_1, \delta_3$, so we have a normally hyperbolic invariant manifold $\Lambda_{L_3^*, \delta_1, \delta_3}$ in each such system of coordinates. Moreover, since the cylinder depends smoothly on the parameters $L_3^*, \delta_1, \delta_3$, this construction allows us to obtain one large normally hyperbolic invariant cylinder Λ^*.

Observe that the contents of Sections 7.1 and 7.2 apply equally in each system of coordinates $\tilde{C}_{L_3^*, \delta_1, \delta_3}$. Furthermore, since the $\hat{\gamma}_2$ variable does not depend on $L_3^*, \delta_1, \delta_3$, the Poincaré section is global, and so we obtain a large 6-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 $\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3$ drift by an amount of order L_3^* along the chain (note that the values of L_3 that we choose belong to some bounded set, as in (2.2)), and choose some sequence $\{\tilde{C}_{L_3^{*,k}, \delta_1^k, \delta_3^k}\}_{k=1,\dots,K}$ 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 $\tilde{C}_{L_3^{*,k}, \delta_1^k, \delta_3^k}$. Note that the shadowing results of [14] apply equally well using the many different coordinate systems, as the coordinates used in that paper are purely local. Thus the assumptions of Theorem 40 apply to the global transition chain on the cylinder $\hat{\Lambda}^*$.

Denoting by $\{\mathcal{L}_i\}$ the global transition chain, Theorem 40 implies that, for any $\eta > 0$ and sufficiently large L_2, L_3, there is a sequence $\{z_j\}_{j \in \mathbb{N}_0}$ in the secular phase space and times $t_j > 0$ such that

$$z_{j+1} = \phi_{\sec}^{t_j}(z_j), \quad d(z_j, \mathcal{L}_j) < \eta$$

for each $j \in \mathbb{N}$ where $\phi_{\sec}^t$ is the flow of the secular Hamiltonian. Moreover, the time to move a distance of order $L_3^{*,1}$ in the L_3 variable and a distance of order L_2 in the $\hat{\Psi}_1, \hat{\Gamma}_3$ directions is of order

$$L_3 L_2^6 \frac{L_3^{12}}{L_2^{18}} \frac{L_3^6}{L_2^8} \frac{L_3^6}{L_2^9} = \frac{L_3^{25}}{L_2^{29}}. \tag{81}$$

This follows from formula (40) and the following facts: we move a distance of order L_3 in the actions $\hat{\Psi}_1$, $\hat{\Gamma}_3, \hat{L}_3$; the return times to the Poincaré section are themselves of order L_2^6 as this is the reciprocal of the order of the frequency of $\hat{\gamma}_2$; the order of the splitting is $\frac{1}{L_2^2}$; the size of the jumps in the scattering map are of order $\frac{L_2^9}{L_3^6}$; and the order of the twist condition is $\frac{L_2^8}{L_3^6}$.

Now, consider the Hamiltonian F of the full four-body problem after averaging the mean anomalies ℓ_1, ℓ_2, defined by (3.2). Fix some $L_1^{\pm}, L_2^{\pm} \in \mathbb{R}$ with $0 < L_1^- < L_1^+ \ll L_2^- < L_2^+$, so that if $(L_1^0, L_2^0) \in [L_1^-, L_1^+] \times [L_2^-, L_2^+]$ then there exist initial conditions $L_3^0 \in \mathbb{R}$ such that (L_1^0, L_2^0, L_3^0) satisfies our assumption (2.2). Let $\Sigma = \mathbb{T}^2 \times [L_1^-, L_1^+] \times [L_2^-, L_2^+]$, and recall from the beginning of Section 7.1 the definition of the subset $\mathcal{D}$ of the secular phase space. We consider the values of energy of the full four-body problem belonging to the set $\mathcal{E}_{4\text{bp}} = \{F(z, \ell_1, \ell_2, L_1, L_2) : z \in \mathcal{D}, (\ell_1, \ell_2, L_1, L_2) \in \Sigma\}$. Fix $E_1 \in \mathcal{E}_{4\text{bp}}$, and denote by Ψ the Poincaré map of the flow of F to the section $\widehat{M} = (\mathcal{D} \times \Sigma) \cap \{\hat{\gamma}_2 = 0\} \cap \{F = E_1\}$. Continuing to denote by z a point in $\mathcal{D}$, we write $(\bar{z}, \bar{\ell}_1, \bar{\ell}_2, \bar{L}_1, \bar{L}_2) = \Psi(z, \ell_1, \ell_2, L_1, L_2)$ with $\bar{z} = G(z, \ell_1, \ell_2, L_1, L_2)$ and $(\bar{\ell}_1, \bar{\ell}_2, \bar{L}_1, \bar{L}_2) = \phi(z, \ell_1, \ell_2, L_1, L_2)$. Since the Hamiltonian F is obtained by averaging the mean anomalies ℓ_1, ℓ_2, there are $\hat{k}_1, \hat{k}_2 \in \mathbb{N}$ such that the variables ℓ_1, ℓ_2 do not appear in F until terms of order $\varepsilon^{\hat{k}_1} \mu^{\hat{k}_2}$ where $\varepsilon = \frac{1}{L_2}$ and $\mu = \frac{L_2}{L_3}$, and moreover we can choose $\hat{k}_1, \hat{k}_2$ to be as large as we like. Therefore the map G takes the form $G(z, \ell_1, \ell_2, L_1, L_2) = \widetilde{G}(z; L_1, L_2) + O\left(\varepsilon^{\hat{k}_1-6} \mu^{\hat{k}_2}\right)$ where the higher-order terms are uniformly bounded in the C^r topology for any $r \in \mathbb{N}$, and where for any fixed values of L_1, L_2, the map $z \mapsto \widetilde{G}(z; L_1, L_2)$ is a Poincaré map of the type constructed in Theorem 24, and thus satisfies the assumptions [A1-3] of Theorem 40. Consequently the map Ψ satisfies assumption [B1] of Theorem 41. In addition, if we write $\phi(z, \ell_1, \ell_2, L_1, L_2) = (\phi_1(z, \ell_1, \ell_2, L_1, L_2), \phi_2(z, \ell_1, \ell_2, L_1, L_2))$ so that $(\bar{\ell}_1, \bar{\ell}_2) = \phi_1(z, \ell_1, \ell_2, L_1, L_2)$ and $(\bar{L}_1, \bar{L}_2) = \phi_2(z, \ell_1, \ell_2, L_1, L_2)$ then we have $\phi_2(z, \ell_1, \ell_2, L_1, L_2) = (L_1, L_2) + O\left(\varepsilon^{\hat{k}_1-6} \mu^{\hat{k}_2}\right)$ with the higher-order terms uniformly C^r-bounded, and so assumption [B2] of Theorem 41 is also satisfied. As explained in Appendix D, a consequence of results of [19] is that the map Ψ has a normally hyperbolic locally invariant manifold $\widetilde{\Lambda}$ that is close to $\hat{\Lambda} \times \Sigma$. Therefore we can use the variables $(w, \ell_1, \ell_2, L_1, L_2)$ as coordinates on $\widetilde{\Lambda}$ where w are coordinates on $\hat{\Lambda}$, and construct a foliation of $\widetilde{\Lambda}$ by the leaves

$$\tilde{\mathcal{L}}\left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_1^*, \hat{L}_2^*, \hat{L}_3^*\right) = \left\{(w, \ell_1, \ell_2, L_1, L_2) : w \in \mathcal{L}\left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right), L_1 = \hat{L}_1^*, L_2 = \hat{L}_2^*\right\}$$

where $\mathcal{L}\left(\hat{\Psi}_1^*, \hat{\Gamma}_3^*, \hat{L}_3^*\right)$ is the leaf of the foliation of $\hat{\Lambda}$ defined by (7.2). Fix $\eta > 0$ and $K_1, K_2 \in \mathbb{N}$, and choose some initial values L_1^1, L_2^1 of the variables L_1, L_2 so that $(\ell_1, \ell_2, L_1^1, L_2^1) \in \text{Int}\,\Sigma$ for any $(\ell_1, \ell_2) \in \mathbb{T}^2$. Choose $N \leq \varepsilon^{-K_1} \mu^{-K_2}$, and values $P_*^1, \ldots, P_*^N$ of the actions $\hat{\Psi}_1, \hat{\Gamma}_3, \hat{L}_3$ such that the leaves $\mathcal{L}\left(P_*^j\right)$ of the foliation of $\hat{\Lambda}$ are connected by one of the scattering maps of the secular Hamiltonian with $L_1 = L_1^1$, $L_2 = L_2^1$ in the sense of Lemma 27. Then by Theorem 41, there are $(L_1^2, L_2^2), \ldots, (L_1^N, L_2^N) \in \left[L_1^-, L_1^+\right] \times \left[L_2^-, L_2^+\right]$ such that, with $\tilde{\mathcal{L}}_j = \tilde{\mathcal{L}}\left(P_*^j, L_1^j, L_2^j\right)$, there are points $(z^1, \ell_1^1, \ell_2^1, L_1^1, L_2^1), \ldots, (z^N, \ell_1^N, \ell_2^N, L_1^N, L_2^N)$ in the phase space of the full four-body problem and times $t_j^* > 0$ such that

$$(z^{i+1}, \ell_1^{i+1}, \ell_2^{i+1}, L_1^{i+1}, L_2^{i+1}) = \phi_{4\text{bp}}^{t_j^*}(z^i, \ell_1^i, \ell_2^i, L_1^i, L_2^i), \quad d\left((z^i, \ell_1^i, \ell_2^i, L_1^i, L_2^i), \tilde{\mathcal{L}}_i\right) < \eta$$

where $\phi_{4\text{bp}}^t$ is the flow of the full four-body problem. Moreover the time estimate (7.3) also holds in this case as the order of the time required to move a distance of order L_3 in the L_3 variable and a distance of order L_2 in the $\hat{\Psi}_1, \hat{\Gamma}_3$ directions. This completes the proof of Theorem 2.

8 Proof of Theorems 1 and 3

In Section 7 the proof of Theorem 2 was completed; the purpose of this section is to show that the analysis of Sections 3 - 7 extends to a complete proof of Theorems 1 and 3.

8.1 The planetary regime: Proof of Theorem 3

We consider now the planetary regime where the masses of bodies 1,2, and 3 are small with respect to body 0. However, in order to make this work we will see that the semimajor axis a_3 depends on the small mass parameter.

Up to now we have investigated what we have called the strongly hierarchical regime, where the semimajor axes satisfy

$$O(1) = a_1 \ll a_2 \to \infty \qquad \text{and} \qquad a_2^3 \ll a_3$$

('strong' meaning that the ratio between a_2 and a_3 has to be rather large).

Now we assume that
$$m_1, m_2, m_3 \sim \rho \to 0.$$

That is, three masses are small and of comparable size. We scale the Deprit actions via

$$L = \rho \check{L}, \quad \Gamma = \rho \check{\Gamma}, \quad \Psi = \rho \check{\Psi}$$

where $L = (L_1, L_2, L_3)$, $\Gamma = (\Gamma_1, \Gamma_2, \Gamma_3)$, and $\Psi = (\Psi_1, \Psi_2)$.

Proposition 29. *The instability mechanism which we have shown to exist in the hierarchical regime in Sections 3-7 continues in the planetary regime (i.e. as ρ tends to 0) as long as the scaled total angular momentum $\check{\Psi}_2$ satisfies*
$$\check{\Psi}_2 \gtrsim \rho^{-1/3}.$$
Moreover, the instability time is of the order of $\rho^{-31/3}$.

Proof. Write $m_j = \rho \bar{m}_j$, $j = 1, 2, 3$, so that, when $\rho \ll 1$,

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

Consider the Keplerian Hamiltonian of the third planet (see (3.2)),

$$F_{\text{Kep},3} = -\frac{\mu_3^3 M_3^2}{2 L_3^2}.$$

If we express it in term of the scaled actions (2.2), we obtain

$$F_{\text{Kep},3} \sim \frac{\rho}{\check{L}_3^2}. \qquad \text{and} \qquad \dot{\ell}_3 \sim \frac{1}{\check{L}_3^3} \tag{82}$$

Proceeding analogously for the perturbing function, one can see that it scales differently in ρ. Indeed, consider for instance the part of the perturbing function F_{per}^{12} regarding planets 1 and 2 (first introduced in equation (3.1)), written in terms of Legendre polynomials. that is

$$F_{\text{per}}^{12} = -\frac{\mu_1 m_2}{\|q_2\|} \sum_{n=2}^{\infty} \tilde{\sigma}_{1,n} P_n(\cos \zeta_1) \left(\frac{\|q_1\|}{\|q_2\|} \right)^n \sim \rho^2 \frac{\check{L}_1^4}{\check{L}_2^6}$$

where

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

Note that if one considers $\|q_1\|$, $\|q_2\|$ independent of ρ, this Hamiltonian has size of order ρ^2. One obtains the same behavior for F_{per}^{23}.

Now, for $\check{L}_3$ independent of ρ: if one lets $\rho \to 0$, the frequency of $\dot{\ell}_3$ becomes much faster than the other secular variables. This alters the hierarchy of time scales considered in the proof of Theorem 2.

To keep the hierarchy it is enough to choose $\check{\Psi}_2 \sim \check{L}_3$ large enough depending on ρ. Indeed, note that in Theorem 2, the only condition on Ψ_2 is a lower bound and therefore one can take $1 \sim \check{L}_1^0 \ll \check{L}_2^0$ independent of ρ and $\check{\Psi}_2 \gtrsim \rho^{-1/3}$, where the exponent of $-\frac{1}{3}$ comes from (8.1). Scaling time by a factor ρ^2 and applying the change of coordinates (3.3), one obtains a secular Hamiltonian whose associated Keplerian (for planet 3), quadrupolar, octupolar Hamiltonians have first orders independent of ρ. Then, the proof of Theorem 2 applies. For this scaled model, the estimate (7.3) implies the innstability time is

$$\check{T} \lesssim \frac{\check{L}_3^{25}}{\check{L}_2^{29}} \lesssim \check{L}_3^{25} \rho^{-29/3}.$$

Scaling back time by the factor ρ^2, one obtains

$$T \sim \rho^{-35/3}.$$

□

8.2 Proof of Theorem 1

Theorem 1 and the comments stated after it are a direct consequence of the proof of Theorems 2 and 3. Indeed, to prove Theorem 1 it is enough to deduce the evolution of the orbital elements and the normalised angular momentum vector from that of the Deprit variables in Theorem 2 and their definitions in Section 2.1 (see, in particular (2.1), (2.1) and (2.1)). The normalised angular momentum vector $\tilde{C}_2$ is determined by the eccentricity e_2, the mutual inclination θ_{23} of bodies 2 and 3, and the longitude h_2 of the node of planet 2. It can easily be checked that Theorems 2 and 3 imply that we can vary the eccentricity e_2, and the mutual inclination θ_{23} in any way we like up to some small error term by varying Γ_2, Γ_3 (see (2.1) and (2.1)). The angle h_2 is determined by angles on the normally hyperbolic cylinder. Although not stated explicitly in Theorems 2 and 3 or in Section 7, it is clear that the shadowing methods of [14] also allow us to shadow values of the angles on the cylinder by time shifts of the flow, and so we can control h_2 in the same way that we control the actions. See Section 9 of [13] for an elaboration of this discussion. The time estimates in (1.2) are also provided by Theorem 2. Finally the analysis of the planetary regime and the time estimates in (1.2) are a consequence of Theorem 3.

A   A refinement of Moser's trick

In Section 5 (see Lemma 17), we prove that there is a near-identity coordinate transformation that straightens the restriction to the normally hyperbolic invariant manifold of the symplectic form. If we were to proceed with this proof, for example, as in Lemma 23 of [13], we would see that the coordinate transformation in the action $\tilde{\Psi}_1$ would dominate the transformation coming from the averaging (i.e. Lemma 18). This would be problematic, as the computation of the coordinate transformation that straightens the symplectic form is significantly more complicated than that of the coordinate transformation that averages the inner angles. Instead, we notice that the first two terms in the Taylor expansion of the parametrisation of the normally hyperbolic manifold do not depend on the angle $\tilde{\psi}_1$ (see Lemma 16); from a symplectic point of view, it seems natural therefore that the coordinate transformation straightening the symplectic form should not alter its symplectic conjugate $\tilde{\Psi}_1$. However, this coordinate transformation (provided by Moser's trick from his proof of Darboux's theorem) is obviously not symplectic, and so it is not clear a priori that this is the case. Below we provide a proof of this fact.

Denote by M a symplectic manifold of dimension 6, and by $(q_0, q_1, q_2, p_0, p_1, p_2)$ coordinates on M such that the symplectic form is

$$\Omega = d\lambda = \sum_{i=0}^{2} dq_i \wedge dp_i \quad \text{where} \quad \lambda = -\sum_{i=0}^{2} p_i \, dq_i.$$

Write $q = (q_1, q_2)$, $p = (p_1, p_2)$, and suppose $\Lambda \subset M$ is a submanifold of M that can be written as the graph of a function $\rho = (\rho_q, \rho_p)$ in the sense that

$$\Lambda = \{(q_0, q, p_0, p) \in M : q_0 = \rho_q(q, p), \ p_0 = \rho_p(q, p), \ (q, p) \in U\}$$

for some domain U. Suppose moreover that $\rho = O(\varepsilon)$ does not depend on q_2; this can be expressed as $\rho = \varepsilon \, \rho'(q_1, p)$, where $\rho' = (\rho'_q, \rho'_p)$. Let $\hat{\varepsilon} = \varepsilon^2$, and write

$$\Omega' = \Omega|_\Lambda = \hat{\varepsilon} \, d\rho'_q \wedge d\rho'_p + \Omega_0, \quad \lambda' = \lambda|_\Lambda = -\hat{\varepsilon} \, \rho'_p \, d\rho'_q + \lambda_0$$

where

$$\Omega_0 = \sum_{i=1}^{2} dq_i \wedge dp_i, \quad \lambda_0 = -\sum_{i=1}^{2} p_i \, dq_i.$$

The goal is to find a coordinate transformation h on Λ that is $O(\hat{\varepsilon})$-close to the identity such that $h^*\Omega' = \Omega_0$. Moser's trick (see the proof of Lemma 17 for the precise construction) is to construct h as the time-$\hat{\varepsilon}$ map of a nonautonomous vector field X_t satisfying

$$i_{X_{\hat{\varepsilon}}}\Omega' = -\frac{d}{d\hat{\varepsilon}}\lambda'. \tag{83}$$

41

Lemma 30. *Denote by h the time-$\hat{\varepsilon}$ map of the vector field X_t defined by (A). Then h has the form*

$$h : \begin{cases} \bar{q}_1 = & q_1 + \hat{\varepsilon}\, f_1(q_1, p) \\ \bar{q}_2 = & q_2 + \hat{\varepsilon}\, f_2(q_1, p) \\ \bar{p}_1 = & p_1 + \hat{\varepsilon}\, f_3(q_1, p) \\ \bar{p}_2 = & p_2. \end{cases}$$

Proof. Notice we can write

$$d\rho'_q \wedge d\rho'_p = A_{13}\, dq_1 \wedge dp_1 + A_{14}\, dq_1 \wedge dp_2 + A_{34}\, dp_1 \wedge dp_2$$

where

$$\begin{cases} A_{13} = & \partial_{q_1}\rho'_q\, \partial_{p_1}\rho'_p - \partial_{q_1}\rho'_p\, \partial_{p_1}\rho'_q \\ A_{14} = & \partial_{q_1}\rho'_q\, \partial_{p_2}\rho'_p - \partial_{q_1}\rho'_p\, \partial_{p_2}\rho'_q \\ A_{34} = & \partial_{p_1}\rho'_q\, \partial_{p_2}\rho'_p - \partial_{p_1}\rho'_p\, \partial_{p_2}\rho'_q. \end{cases}$$

Writing Ω' in matrix form, we have

$$\Omega' = \Omega_0 + \hat{\varepsilon}\, A \quad \text{where} \quad \Omega_0 = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{84}$$

and

$$A = \begin{pmatrix} 0 & 0 & A_{13} & A_{14} \\ 0 & 0 & 0 & 0 \\ -A_{13} & 0 & 0 & A_{34} \\ -A_{14} & 0 & -A_{34} & 0 \end{pmatrix}. \tag{85}$$

On the other hand we can write

$$\rho'_p\, d\rho'_q = a_1\, dq_1 + a_3\, dp_1 + a_4\, dp_2$$

where

$$\begin{cases} a_1 = & \rho'_p\, \partial_{q_1}\rho'_q \\ a_3 = & \rho'_p\, \partial_{p_1}\rho'_q \\ a_4 = & \rho'_p\, \partial_{p_2}\rho'_q \end{cases}$$

so we can write λ' in vector form as

$$\lambda' = \lambda_0 + \hat{\varepsilon}\, a \quad \text{where} \quad \lambda_0 = -\begin{pmatrix} p_1 \\ p_2 \\ 0 \\ 0 \end{pmatrix}, \quad a = -\begin{pmatrix} a_1 \\ 0 \\ a_3 \\ a_4 \end{pmatrix}. \tag{86}$$

Now, let $X_{\hat{\varepsilon}} = (\dot{q}, \dot{p})$. Then from (A) and (A), we see that the left-hand side of (A) is

$$i_{X_{\hat{\varepsilon}}}\Omega' = X_{\hat{\varepsilon}}^T\, \Omega_0 + \hat{\varepsilon}\, X_{\hat{\varepsilon}}^T\, A = \begin{pmatrix} -\dot{p}_1 \\ -\dot{p}_2 \\ \dot{q}_1 \\ \dot{q}_2 \end{pmatrix} + \hat{\varepsilon} \begin{pmatrix} -A_{13}\,\dot{p}_1 - A_{14}\,\dot{p}_2 \\ 0 \\ A_{13}\,\dot{q}_1 - A_{34}\,\dot{p}_2 \\ A_{14}\,\dot{q}_1 + A_{34}\,\dot{p}_1 \end{pmatrix}. \tag{87}$$

Meanwhile (A) implies that the right-hand side of (A) is

$$-\frac{d}{d\hat{\varepsilon}}\lambda' = \begin{pmatrix} a_1 + \hat{\varepsilon}\partial_{\hat{\varepsilon}}a_1 \\ 0 \\ a_3 + \hat{\varepsilon}\partial_{\hat{\varepsilon}}a_3 \\ a_4 + \hat{\varepsilon}\partial_{\hat{\varepsilon}}a_4 \end{pmatrix}. \tag{88}$$

Combining (A) and (A) yields

$$\begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{pmatrix} = \begin{pmatrix} a_3 + \hat{\varepsilon}\partial_{\hat{\varepsilon}}a_3 - \hat{\varepsilon}\,(A_{13}\,\dot{q}_1 - A_{34}\,\dot{p}_2) \\ a_4 + \hat{\varepsilon}\partial_{\hat{\varepsilon}}a_4 - \hat{\varepsilon}\,(A_{14}\,\dot{q}_1 + A_{34}\,\dot{p}_1) \\ -a_1 - \hat{\varepsilon}\partial_{\hat{\varepsilon}}a_1 + \hat{\varepsilon}\,(A_{13}\,\dot{p}_1 + A_{14}\,\dot{p}_2) \\ 0 \end{pmatrix}$$

which completes the proof of the lemma. $\qquad\square$

B The dependence on $\tilde{\gamma}_3$: Proof of Lemma 19

The first step to prove Lemma 19 is to analyse more carefully the error terms in the expansion of the Hamiltonian F_{per} (see (2.1)) given in (3.1).

To this end, recall that F_{per} can be written as $F_{\mathrm{per}} = F_{\mathrm{per}}^{12} + R$ with

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

$$R = \frac{\mu_3 M_3}{||q_3||} - \frac{m_0 m_3}{||q_3 + \sigma_{22} q_2 + \sigma_{11} q_1||} - \frac{m_1 m_3}{||q_3 + \sigma_{22} q_2 + (\sigma_{11} - 1) q_1||} - \frac{m_2 m_3}{||q_3 + (\sigma_{22} - 1) q_2||}$$

where $M_j = \sum_{i=0}^{j} m_i$, $\sigma_{ij} = \frac{m_i}{M_j}$ and $\mu_j^{-1} = M_{j-1}^{-1} + m_j^{-1}$.

Lemma 31. *The Hamiltonian R can be written as $R = F_{\mathrm{per}}^{23} + F_{\mathrm{per}}^{13}$ with*

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

$$F_{\mathrm{per}}^{13} = -\frac{\mu_1 m_3}{||q_3 + \sigma_{22} q_2||} \sum_{n=2}^{+\infty} \tilde{\sigma}_{3n} P_n(\cos(\alpha(q_3 + \sigma_{22} q_2, q_1))) \left(\frac{||q_1||}{||q_3 + \sigma_{22} q_2||} \right)^n$$

where P_n is the Legendre polynomial of degree n, $\alpha(x, y)$ denotes the angle between the vectors x and y, $\tilde{\sigma}_{2n} = (\sigma_{02} + \sigma_{12})^{n-1} + (-1)^n \sigma_{22}^{n-1}$ and $\tilde{\sigma}_{3n} = \sigma_{01}^{n-1} + (-1)^n \sigma_{11}^{n-1}$.

The proof of this lemma is just a direct consequence of expanding the terms in R using the Legendre polynomials and taking advantage of certain cancellations in the first order terms.

After averaging the angles ℓ_1 and ℓ_2, we have that the Hamiltonian $\tilde{F}_{\mathrm{sec}}$ in (3.2) can be written as $\tilde{F}_{\mathrm{sec}} = F_{\mathrm{sec}}^{12} + F_{\mathrm{sec}}^{23} + F_{\mathrm{sec}}^{13}$ where F_{sec}^{12} and F_{sec}^{23} are the functions introduced in (3.2) and (3.2) respectively and

$$F_{\mathrm{sec}}^{13} = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} F_{\mathrm{per}}^{13} d\ell_1 d\ell_2.$$

Recall that we have

$$F_{\mathrm{sec}} = (F_{\mathrm{Kep}} - \tilde{F}_{\mathrm{Kep}}) + \tilde{F}_{\mathrm{sec}} = (F_{\mathrm{Kep}} - \tilde{F}_{\mathrm{Kep}}) + F_{\mathrm{sec}}^{12} + F_{\mathrm{sec}}^{23} + F_{\mathrm{sec}}^{13}.$$

Now we analyse the $\tilde{\gamma}_3$-dependence of each of these terms.

The terms in F_{sec} depending on $\tilde{\gamma}_3$ are F_{sec}^{23} and F_{sec}^{13}. We conclude the proof of Lemma 19 using the two following lemmas.

Lemma 32. *The angle $\tilde{\gamma}_3$ does not appear in $\left\langle F_{\mathrm{sec}}^{23} \right\rangle_{(\tilde{\psi}_1, \tilde{\gamma}_2, \ell_3)}$ until the term of order $O\left(\frac{L_2^6}{L_3^{*8}} \right)$.*

Proof. Recall that $F_{\mathrm{sec}}^{23} = -\frac{\mu_2 m_3}{2\pi} \left(F_{\mathrm{quad}}^{23} + O\left(\frac{a_2^3}{a_3^4} \right) \right) = -\frac{\mu_2 m_3}{2\pi} \left(F_{\mathrm{quad}}^{23} + O\left(\frac{L_2^6}{L_3^{*8}} \right) \right)$, and we already know (see Proposition 7) that $\left\langle F_{\mathrm{quad}}^{23} \right\rangle_{(\tilde{\psi}_1, \tilde{\gamma}_2, l_3)}$ does not depend on $\tilde{\gamma}_3$, which completes the proof. $\square$

Lemma 33. *The angle $\tilde{\gamma}_3$ does not appear in $\left\langle F_{\mathrm{sec}}^{13} \right\rangle_{(\tilde{\psi}_1, \tilde{\gamma}_2, \ell_3)}$ until the term of order $O\left(\frac{L_2^2}{L_3^{*8}} \right)$.*

Proof. Keeping just the first order term in the expression of F_{sec}^{13}, we obtain

$$F_{\mathrm{sec}}^{13} = -\frac{\mu_1 m_3}{2\pi} \int_{\mathbb{T}} P_2(\cos(\alpha(q_3 + \sigma_{22} q_2, q_1))) \frac{||q_1||^2}{||q_3 + \sigma_{22} q_2||^3} d\ell_1 + O\left(\frac{1}{||q_3||^4} \right).$$

Expanding the integrand in powers of $\frac{||q_2||}{||q_3||}$, we have

$$F_{\mathrm{sec}}^{13} = -\frac{\mu_1 m_3}{2\pi} F_{\mathrm{quad}}^{13} + O\left(\frac{||q_2||}{||q_3||^4} \right),$$

43

where

$$F_{\text{quad}}^{13} = \int_{\mathbb{T}} P_2(\cos(\alpha(q_3, q_1))) \frac{||q_1||^2}{||q_3||^3} d\ell_1.$$

As $O\left(\frac{||q_2||}{||q_3||^4}\right) = O\left(\frac{L_2^2}{L_3^{*8}}\right)$, we just need to prove that $\tilde{\gamma}_3$ does not appear in $\left\langle F_{\text{quad}}^{13} \right\rangle_{(\tilde{\psi}_1, \tilde{\gamma}_2, \ell_3)}$ before order $O\left(\frac{L_2^6}{L_3^{*8}}\right)$ to conclude.

We first expand $\cos(\alpha(q_3, q_1))$. Following the notation of Section 3.3, we have

$$\cos(\alpha(q_3, q_1)) = \bar{q}_1.\bar{q}_3 = \mathcal{R}_x(\tilde{i}_2)\mathcal{R}_z(\psi_1)\mathcal{R}_x(i_1)\bar{Q}_1.I_z\mathcal{R}_x(i_3)\bar{Q}_3.$$

We know that $\cos i_3 = 1 + O\left(\frac{L_2^2}{L_3^{*2}}\right)$ and $\sin i_3 = A\frac{L_2}{L_3^*} + O\left(\frac{L_2^2}{L_3^{*2}}\right)$, where A, of which we do not need the explicit expression, is a function of some actions. Therefore, we can write $\mathcal{R}_x(i_3) = I + \frac{L_2}{L_3^*}I_1 + O\left(\frac{L_2^2}{L_3^{*2}}\right)$, where

$$I_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -A \\ 0 & A & 0 \end{pmatrix}.$$

Then, we obtain $\cos(\alpha(q_3, q_1)) = W_0 + \frac{L_2}{L_3^*}W_1 + O\left(\frac{L_2^2}{L_3^{*2}}\right)$, where

$$W_0 = \mathcal{R}_x(\tilde{i}_2)\mathcal{R}_z(\psi_1)\mathcal{R}_x(i_1)\bar{Q}_1 \cdot I_z\bar{Q}_3$$
$$W_1 = \mathcal{R}_x(\tilde{i}_2)\mathcal{R}_z(\psi_1)\mathcal{R}_x(i_1)\bar{Q}_1 \cdot I_z I_1\bar{Q}_3.$$

We then obtain

$$P_2(\cos(\alpha(q_3, q_1))) = P_2\left(W_0 + \frac{L_2}{L_3^*}W_1 + O\left(\frac{L_2^2}{L_3^{*2}}\right)\right) = P_2(W_0) + \frac{L_2}{L_3^*}P_2'(W_0)W_1 + O\left(\frac{L_2^2}{L_3^{*2}}\right)$$

$$= P_2(W_0) + \frac{L_2}{L_3^*}3W_0W_1 + O\left(\frac{L_2^2}{L_3^{*2}}\right)$$

and therefore

$$F_{\text{quad}}^{13} = \int_{\mathbb{T}} P_2(W_0)\frac{||q_1||^2}{||q_3||^3}d\ell_1 + \frac{L_2}{L_3^*}\int_{\mathbb{T}} 3W_0W_1\frac{||q_1||^2}{||q_3||^3}d\ell_1 + O\left(\frac{L_2^2}{L_3^{*8}}\right).$$

Thus, to conclude, we just need to prove that the two integrals in the expression above, averaged with respect to $\tilde{\psi}_1, \tilde{\gamma}_2, \ell_3$ do not depend on $\tilde{\gamma}_3$. We compute their average using the explicit expression of each term and integrating over the angle u_1 when averaging with respect to ℓ_1 and over the angle v_3 when averaging with respect to ℓ_3. Then, it is easy to check that these explicit expressions do not depend on $\tilde{\gamma}_3$, which completes the proof. $\square$

C The scattering map of a normally hyperbolic invariant manifold

In this section we denote by M a C^r smooth manifold, and by $\phi^t : M \to M$ a smooth flow with $\frac{d}{dt}\big|_{t=0}\phi^t = X$ where $X \in C^r(M, TM)$. Let $\Lambda \subset M$ be a compact ϕ^t-invariant submanifold, possibly with boundary. By ϕ^t-invariant we mean that X is tangent to Λ, but that orbits can escape through the boundary (a concept sometimes referred to as *local* invariance).

Definition 34. *We call Λ a* normally hyperbolic invariant manifold *for ϕ^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:

$$\|D\phi^t|_{E^s}\| \le C\lambda^t \text{ for all } t \ge 0,$$
$$\|D\phi^t|_{E^u}\| \le C\lambda^{-t} \text{ for all } t \le 0,$$
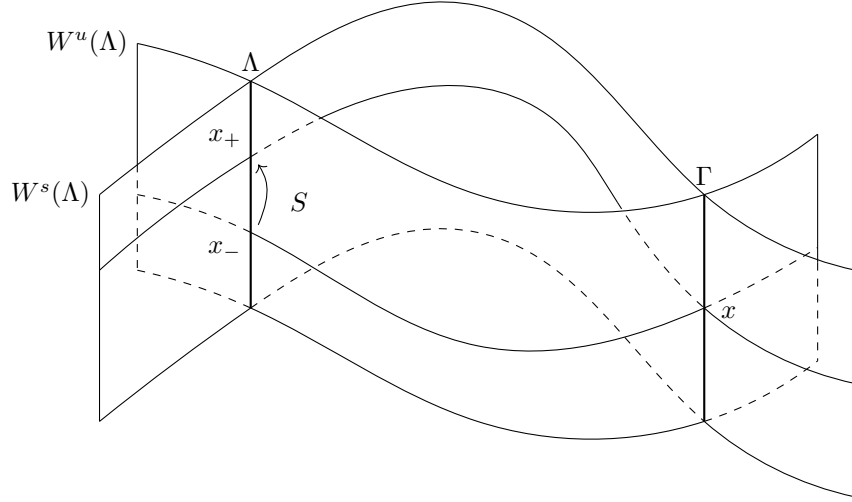$$\|D\phi^t|_{T\Lambda}\| \le C\mu^{|t|} \text{ for all } t \in \mathbb{R}.$$

Figure 1: 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 Γ, and from there follows the unique leaf of the strong stable foliation passing through x to the point x_+ on Λ.

Moreover, Λ is called an r-normally hyperbolic invariant manifold if it is C^r smooth, and

$$0 < \lambda < \mu^{-r} < 1 \tag{89}$$

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 $W_{\mathrm{loc}}^s(\Lambda)$ is the set of points in a small neighbourhood of Λ whose forward orbits never leave the neighbourhood, and tend exponentially to Λ. The local unstable manifold $W_{\mathrm{loc}}^u(\Lambda)$ is the set of points in the neighbourhood whose backward orbits stay in the neighbourhood and tend exponentially to Λ. We then define

$$W^s(\Lambda) = \bigcup_{t \geq 0}^{\infty} \phi^{-t}\left(W_{\mathrm{loc}}^s(\Lambda)\right), \quad W^u(\Lambda) = \bigcup_{t \geq 0}^{\infty} \phi^t\left(W_{\mathrm{loc}}^u(\Lambda)\right).$$

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 $\phi^t\left(W^s(x)\right) = W^s\left(\phi^t(x)\right)$ and $\phi^t\left(W^u(x)\right) = W^u\left(\phi^t(x)\right)$ for each $x \in \Lambda$ and $t \in \mathbb{R}$. We thus define the *holonomy maps* $\pi^{s,u} : W^{s,u}(\Lambda) \to \Lambda$ 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 $x_+ \in \Lambda$ 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 $x \in \left(W^s(\Lambda) \pitchfork W^u(\Lambda)\right) \setminus \Lambda$ 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

$$\begin{aligned}
T_x W^s(x_+) \oplus T_x\left(W^s(\Lambda) \cap W^u(\Lambda)\right) &= T_x W^s(\Lambda), \\
T_x W^u(x_-) \oplus T_x\left(W^s(\Lambda) \cap W^u(\Lambda)\right) &= T_x W^u(\Lambda).
\end{aligned} \tag{90}$$

In this case we can take a sufficiently small neighbourhood Γ of x in $W^s(\Lambda) \cap W^u(\Lambda)$ so that (C) holds at each point of Γ, and the restrictions to Γ of the holonomy maps are bijections onto their images. We call Γ a *homoclinic channel* (see Figure 1). We can then define the scattering map as follows [19].

Definition 35. *Let $y_- \in \pi^u(\Gamma)$, let $y = \left(\pi^u|_\Gamma\right)^{-1}(y_-)$, and let $y_+ = \pi^s(y)$. The* scattering map *$S : \pi^u(\Gamma) \to \pi^s(\Gamma)$ is defined by*

$$S = \pi^s \circ (\pi^u)^{-1} : y_- \longmapsto y_+.$$

Suppose now that the smoothness r of M and X is at least 2, suppose the normally hyperbolic invariant manifold Λ is a C^r submanifold of M, and suppose the large spectral gap condition (34) holds. This implies C^{r-1} regularity of the strong stable and strong unstable foliations [42], which in turn implies that the scattering map S is C^{r-1} [19].

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.

D   A general shadowing argument

We follow the notation and exposition of [14]. Let M be a C^r manifold of dimension $d = 2(m+n)$ where $r \geq 4$. Let $F \in \mathrm{Diff}^4(M)$, and assume F depends smoothly on a small parameter ϵ, 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 (34); suppose moreover that Λ is diffeomorphic to $\mathbb{T}^n \times [0,1]^n$. 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 Γ for all sufficiently small $\epsilon > 0$. We allow for the possibility that the angle between $W^{s,u}(\Lambda)$ along the homoclinic channel Γ goes to 0 as $\epsilon \to 0$. 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 36. *For $\sigma \geq 0$, we say that the angle of the splitting along Γ is of order ϵ^σ if there is a positive constant C (independent of ϵ) such that*

$$\alpha_\Gamma(x) \geq C\epsilon^\sigma \qquad \text{for all} \quad x \in \Gamma.$$

Recall we have assumed that the normally hyperbolic invariant manifold Λ is diffeomorphic to $\mathbb{T}^n \times [0,1]^n$, and denote by $(q, p) \in \mathbb{T}^n \times [0,1]^n$ smooth coordinates on Λ. Define $f := F|_\Lambda$, which also depends on the small parameter ϵ.

Definition 37. *We say that $f : \Lambda \to \Lambda$ is a* near-integrable twist map *if there is some $k \in \mathbb{N}$ 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 $f : \Lambda \to \Lambda$ is a near-integrable twist map, then there exist twist parameters $T_+ > \widetilde{T}_- > 0$ such that

$$\widetilde{T}_- \|v\| \leq \|Dg(p)v\| \leq T_+ \|v\|$$

for all $p \in [0,1]^n$ and all $v \in \mathbb{R}^n$. We can always choose T_+ to be independent of ϵ. Our formulation of the problem allows the parameter $\widetilde{T}_-$ to depend on ϵ: there is $\tau \in \mathbb{N}_0$ and a strictly positive constant T_- (independent of ϵ) such that $\widetilde{T}_- = \epsilon^\tau T_-$.

Definition 38. *Suppose $f : \Lambda \to \Lambda$ is a near-integrable twist map. Denote by $T_+ > \widetilde{T}_- = \epsilon^\tau T_- > 0$ the twist parameters. We say that f satisfies:*

- *A* uniform twist condition *if $\tau = 0$;*

- *A non-uniform twist condition (of order ϵ^τ) if $\tau > 0$, and the order ϵ^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 Λ, the leaves of which are given by

$$\Lambda(p^*) = \{(q, p) \in \Lambda : p = p^*\}. \tag{91}$$

If $f : \Lambda \to \Lambda$ is a near-integrable twist map in the sense of Definition 37, then each leaf of the foliation is almost invariant under f, up to terms of order ϵ^k. Denote by $U \subset \Lambda$ the domain of definition of the scattering map S.

Definition 39. *We say that the scattering map S is* transverse to leaves along leaves, *and that the angle of transversality is of order ϵ^v (with respect to the leaves (D) of the foliation of Λ) if there are $c, C > 0$ such that for all $p_0^* \in [0,1]^n$ and all $p^* \in [0,1]^n$ satisfying $\|p^* - p_0^*\| < c\,\epsilon^v$ we have*

$$S\left(\Lambda(p_0^*) \cap U\right) \pitchfork \Lambda(p^*) \neq \emptyset$$

and there is $x \in S\left(\Lambda(p_0^*) \cap U\right) \pitchfork \Lambda(p^*)$ such that

$$\inf \alpha(v_0, v) \geq C\epsilon^v$$

where the infimum is taken over all $v_0 \in T_x S\left(\Lambda(p_0^*) \cap U\right)$ and $v \in T_x \Lambda(p^*)$ 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 (3.3) to prove the existence of drifting orbits in the secular subsystem.

[A1] The stable and unstable manifolds $W^{s,u}(\Lambda)$ have a strongly transverse homoclinic intersection along a homoclinic channel Γ, and so we have an open set $U \subseteq \Lambda$ and a scattering map $S : U \to \Lambda$. The angle of the splitting along Γ is of order ϵ^σ.

[A2] The inner map $f = F|_\Lambda$ is a near-integrable twist map with error terms of order ϵ^k satisfying a non-uniform (or uniform) twist condition of order ϵ^τ.

[A3] The scattering map S is transverse to leaves along leaves (with respect to the leaves (D) of the foliation of Λ), and the angle of transversality is of order ϵ^v.

Theorem 40. *Fix $\eta > 0$, let $\epsilon > 0$ be sufficiently small, and suppose $k \geq 2(\rho + \tau) + 1$ where $\rho = \max\{2\sigma, 2v, \tau\}$. Choose $\{p_j^*\}_{j=1}^\infty \subset [0,1]^n$ such that*

$$S\left(\Lambda_j \cap U\right) \cap \Lambda_{j+1} \neq \emptyset,$$

and $S\left(\Lambda_j \cap U\right)$ is transverse to Λ_{j+1}, where $\Lambda_j = \Lambda(p_j^*)$. Suppose the distance between Λ_j and Λ_{j+1} is of order ϵ^v for each j. Then there are $\{z_i\}_{i=1}^\infty \subset M$ and $n_i \in \mathbb{N}$ such that $z_{i+1} = F^{n_i}(z_i)$ and

$$d(z_i, \Lambda_i) < \eta.$$

Moreover, the time to move a distance of order 1 in the p-direction is bounded from above by a term of order

$$\epsilon^{-\rho - \tau - v}. \tag{92}$$

Observe that Theorem 40 cannot be applied to (3.2). Indeed, a crucial assumption in Theorem 40 is that the scattering map S is transverse to leaves along leaves. For (3.2), 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 (3.2). Theorem 41 below generalises Theorem 40 to settings where transversality is only known in some directions, and thus allows us to complete the proof of Theorem 2.

To state Theorem 41 we consider, as before, a C^r manifold M of dimension $2(m + n)$ where $r \geq 4$ and $m, n \in \mathbb{N}$. Let $\Sigma = \mathbb{T}^{\ell_1} \times [0,1]^{\ell_2}$ for some $\ell_1, \ell_2 \in \mathbb{N}_0$, and denote by $(\theta, \xi) \in \mathbb{T}^{\ell_1} \times [0,1]^{\ell_2}$ coordinates on Σ. Write $\widetilde{M} = M \times \Sigma$. Suppose $\Psi \in \text{Diff}^4\left(\widetilde{M}\right)$ such that

$$\Psi(z, \theta, \xi) = (G(z, \theta, \xi), \phi(z, \theta, \xi))$$

where $z \in M$, $G \in C^4\left(\widetilde{M}, M\right)$, and $\phi \in C^4\left(\widetilde{M}, \Sigma\right)$. Suppose Ψ depends on a small parameter ϵ. We make the following assumptions on Ψ.

[B1] There is some $L \in \mathbb{N}$ such that
$$G(z, \theta, \xi) = \widetilde{G}(z; \xi) + O\left(\epsilon^L\right)$$
where the higher order terms are uniformly bounded in the C^4 topology, and for each $\xi \in [0,1]^{\ell_2}$ the map
$$\widetilde{G}(\cdot; \xi) : z \in M \longmapsto \widetilde{G}(z; \xi) \in M$$
satisfies the assumptions [A1-3] of Theorem 40.

[B2] Moreover, the map ϕ has the form

$$\phi : \begin{cases} \bar{\theta} = \phi_1(z, \theta, \xi) \\ \bar{\xi} = \phi_2(z, \theta, \xi) = \xi + O\left(\epsilon^L\right) \end{cases}$$

where the higher order terms are uniformly bounded in the C^4 topology.

Results from [19] imply that Ψ has a normally hyperbolic invariant manifold $\widetilde{\Lambda}$ that is $O\left(\epsilon^L\right)$ close in the C^4 topology to $\Lambda \times \Sigma$ where $\Lambda \subset M$ is the normally hyperbolic invariant manifold of $\widetilde{G}(\cdot; \xi)$. Moreover there is an open set $\widetilde{U} \subset \widetilde{\Lambda}$ and a scattering map $\widetilde{S} : \widetilde{U} \to \widetilde{\Lambda}$ such that the z-component of $\widetilde{S}(z, \theta, \xi)$ is $O\left(\epsilon^L\right)$ close in the C^3 topology to $S(z; \xi)$ where $S(\cdot; \xi) : U \to \Lambda$ is the scattering map corresponding to $\widetilde{G}(\cdot; \xi)$.

We use the coordinates (q, p, θ, ξ) on $\widetilde{\Lambda}$ where (q, p) are the coordinates on Λ and (θ, ξ) are the coordinates on Σ. Notice that the sets

$$\widetilde{\Lambda}(p^*, \xi^*) = \left\{ (q, p, \theta, \xi) \in \widetilde{\Lambda} : p = p^*, \xi = \xi^* \right\} = \Lambda(p^*) \times \mathbb{T}^{\ell_1} \times \{\xi^*\}$$

for $p^* \in [0,1]^n$ and $\xi^* \in [0,1]^{\ell_2}$ define the leaves of a foliation of $\widetilde{\Lambda}$, where $\Lambda(p^*)$ are the leaves of the foliation of Λ defined by (D).

Theorem 41. *Fix $\eta > 0$ and $K \in \mathbb{N}$ and let $\epsilon > 0$ be sufficiently small. Choose $N \in \mathbb{N}$ satisfying*

$$N \leq \frac{1}{\epsilon^K}$$

$\xi_1^* \in \text{Int}\left([0,1]^{\ell_2}\right)$ so that $\widetilde{G}(\cdot; \xi_1^*)$ satisfies assumptions [A1-3], and $p_1^*, \ldots, p_N^* \in [0,1]^n$ as in Theorem 40 such that
$$S\left(\Lambda_j \cap U; \xi_1^*\right) \cap \Lambda_{j+1} \neq \emptyset$$
and $S\left(\Lambda_j \cap U; \xi_1^*\right)$ is transverse to Λ_{j+1}, where $\Lambda_j = \Lambda(p_j^*)$. Suppose the distance between Λ_j and Λ_{j+1} is of order ϵ^{υ} for each j, and $L > 0$ is sufficiently large, depending on K. Then there are $\xi_2^*, \ldots, \xi_N^* \in [0,1]^{\ell_2}$ such that, with $\widetilde{\Lambda}_j = \widetilde{\Lambda}\left(p_j^*, \xi_j^*\right)$, there are $w_1, \ldots, w_N \in \widetilde{M}$ and $n_i \in \mathbb{N}$ such that the ξ component of w_1 is ξ_1^*,
$$w_{i+1} = \Psi^{n_i}(w_i),$$
and
$$d\left(w_i, \widetilde{\Lambda}_i\right) < \eta$$

where ρ, σ, τ are as in the statement of Theorem 40. Moreover, the time to move a distance of order 1 in the p-direction is of order $\epsilon^{-\rho - \tau - \upsilon}$.

Note that the transition chain obtained in Theorem 41 is only of finite length, while the one obtained in Theorem 40 may be infinite.

E   Derivatives of the inner Hamiltonian

Recall the definition of the integrable part $\hat{F}_0$ of the inner secular Hamiltonian, after straightening the symplectic form and averaging the inner angles, as constructed in Theorem 15. In order to prove in Lemma 26 that the Poincaré map satisfies a twist condition, we need to compute the first and second partial derivatives of $\hat{F}_0$ with respect to the inner actions, or in some cases simply estimate the order. This information is provided in the following lemma.

Lemma 42. *The first and second-order partial derivatives of $\hat{F}_0$ with respect to the actions $\hat{\Gamma}_2$, $\hat{\Psi}_1$, $\hat{\Gamma}_3$, and $\hat{L}_3$ are as follows:*

$$\frac{\partial \hat{F}_0}{\partial \hat{\Gamma}_2} = \varepsilon^6 \, C_{12} \, \frac{6\hat{\Gamma}_2}{L_1^2 \, \delta_1^3} + \cdots , \qquad\qquad \frac{\partial \hat{F}_0}{\partial \hat{\Psi}_1} = 3\,\varepsilon^7 \, C_{12} \, \frac{L_1^2 - 3\,\hat{\Gamma}_2^2}{L_1^2 \, \delta_1^4} + \cdots ,$$

$$\frac{\partial \hat{F}_0}{\partial \hat{\Gamma}_3} = \varepsilon^3 \, \mu^6 C_{23} \, \frac{(20 - 12\,\delta_1^2)\,\delta_3}{\delta_1^2 \, \delta_2^3} + \cdots , \qquad \frac{\partial \hat{F}_0}{\partial \hat{L}_3} = \varepsilon^3 \, \mu^3 \, \alpha_{\text{Kep}} + \cdots ,$$

$$\frac{\partial^2 \hat{F}_0}{\partial \hat{\Gamma}_2^2} = \varepsilon^6 \, C_{12} \, \frac{6}{L_1^2 \, \delta_1^3} + \cdots , \qquad\qquad \frac{\partial^2 \hat{F}_0}{\partial \hat{\Gamma}_2 \, \partial \hat{\Psi}_1} = -\,\varepsilon^7 \, C_{12} \, \frac{18\,\hat{\Gamma}_2}{L_1^2 \, \delta_1^4} + \cdots ,$$

$$\frac{\partial^2 \hat{F}_0}{\partial \hat{\Gamma}_2 \, \partial \hat{\Gamma}_3} = \varepsilon^4 \, \mu^6 C_{23} \, \frac{24\,\delta_3}{\delta_1 \, \delta_2^3} + \cdots , \qquad\qquad \frac{\partial^2 \hat{F}_0}{\partial \hat{\Psi}_1^2} = 12\,\varepsilon^8 \, C_{12} \, \frac{3\,\hat{\Gamma}_2^2 - L_1^2}{L_1^2 \, \delta_1^5} + \cdots ,$$

$$\frac{\partial^2 \hat{F}_0}{\partial \hat{\Psi}_1 \, \partial \hat{\Gamma}_3} = -\,\varepsilon^4 \, \mu^6 C_{23} \, \frac{40\,\delta_3}{\delta_1^3 \, \delta_2^3} + \cdots , \qquad\qquad \frac{\partial^2 \hat{F}_0}{\partial \hat{\Gamma}_3^2} = \varepsilon^4 \, \mu^6 C_{23} \, \frac{20 - 12\,\delta_1^2}{\delta_1^2 \, \delta_2^3} + \cdots ,$$

$$\frac{\partial^2 \hat{F}_0}{\partial \hat{L}_3^2} = -\,3\,\varepsilon^4 \, \mu^4 \, \alpha_{\text{Kep}} + \cdots , \qquad\qquad \frac{\partial^2 \hat{F}_0}{\partial \hat{L}_3 \, \partial \hat{\Gamma}_2} = O\left(\varepsilon^4 \, \mu^7\right) ,$$

$$\frac{\partial^2 \hat{F}_0}{\partial \hat{L}_3 \, \partial \hat{\Psi}_1} = O\left(\varepsilon^4 \, \mu^7\right) , \qquad\qquad \frac{\partial^2 \hat{F}_0}{\partial \hat{L}_3 \, \partial \hat{\Gamma}_3} = O\left(\varepsilon^4 \, \mu^7\right) ,$$

where $\varepsilon = \frac{1}{L_2}$, $\mu = \frac{L_2}{L_3^*}$, where C_{12}, C_{23} are nonzero constants independent of L_2 and L_3^* coming from F_{quad}^{12}, F_{quad}^{23} respectively, and where the nonzero constant α_{Kep} is defined by (3.3).

Proof. Observe that $F_{\text{quad}}^{12}\big|_\Lambda$ and $F_{\text{quad}}^{23}\big|_\Lambda$ are the same, after we average the inner angles, as the analogous objects in [13] up to higher order terms depending also on $\hat{L}_3$. Therefore all of the derivatives taken with respect to the variables $\hat{\Gamma}_2$, $\hat{\Psi}_1$, $\hat{\Gamma}_3$ are given at first order by Lemma 25 of [13].

From (3.3), (3.3), and (3.3), we see that

$$\frac{\partial \hat{F}_0}{\partial \hat{L}_3} = \frac{\partial F_{\text{Kep}}}{\partial \hat{L}_3} + \cdots = \frac{1}{(L_3^*)^3} \, \alpha_{\text{Kep}} + \cdots , \quad \frac{\partial^2 \hat{F}_0}{\partial \hat{L}_3^2} = \frac{\partial^2 F_{\text{Kep}}}{\partial \hat{L}_3^2} + \cdots = -3 \, \frac{1}{(L_3^*)^4} \, \alpha_{\text{Kep}} + \cdots .$$

As for the mixed second partial derivatives with respect to $\hat{L}_3$ and the other actions, we estimate the order as follows. Products of $\hat{L}_3$ and the other actions come, at first order, in the expansion of F_{quad}^{23}. We can find these by normalising F_{quad}^{23} to obtain

$$\tilde{F}_{\text{quad}}^{23} = \frac{L_3^6}{L_2^4} \, \frac{1}{(2\,\pi)^4} \int_{\mathbb{T}^4} F_{\text{quad}}^{23} \, d\tilde{\gamma}_2 \, d\tilde{\psi}_1 \, d\tilde{\gamma}_3 \, d\tilde{\ell}_3,$$

expanding the coefficient

$$\frac{L_2^4}{L_3^6} = \frac{L_2^4}{\left(L_3^* + \tilde{L}_3\right)^6} = \frac{L_2^4}{(L_3^*)^6} - 6 \, \frac{L_2^4}{(L_3^*)^7} \, \tilde{L}_3 + O\left(\frac{L_2^4}{(L_3^*)^8}\right), \tag{93}$$

and noticing that the first appearance of the actions $\tilde{\Gamma}_2$, $\tilde{\Psi}_1$, and $\tilde{\Gamma}_3$ in the expansion of $\tilde{F}_{\text{quad}}^{23}$ can be estimated by

$$\frac{\partial \tilde{F}_{\text{quad}}^{23}}{\partial \left(\tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3 \right)} = O\left(\frac{1}{L_2} \right) \tag{94}$$

because the first order term in the expansion of F_{quad}^{23} (see H_0^{23}, defined by (12)) does not depend on any of the actions. Combining (E) and (E), and defining $\left\langle F_{\text{quad}}^{23} \right\rangle = \frac{L_2^4}{L_3^6} \tilde{F}_{\text{quad}}^{23}$ yields

$$\frac{\partial^2 \left\langle F_{\text{quad}}^{23} \right\rangle}{\partial \tilde{L}_3 \, \partial \left(\tilde{\Gamma}_2, \tilde{\Psi}_1, \tilde{\Gamma}_3 \right)} = O\left(\frac{L_2^3}{\left(L_3^* \right)^7} \right) = O\left(\varepsilon^4 \, \mu^7 \right),$$

and so

$$\frac{\partial^2 \left\langle F_{\text{quad}}^{23} \right\rangle}{\partial \hat{L}_3 \, \partial \left(\hat{\Gamma}_2, \hat{\Psi}_1, \hat{\Gamma}_3 \right)} = O\left(\frac{L_2^3}{\left(L_3^* \right)^7} \right) = O\left(\varepsilon^4 \, \mu^7 \right) \tag{95}$$

since the transformation from the 'tilde' to the 'hat' coordinates is close to the identity in C^r. To see that this implies the estimates given in the statement of the lemma, it only remains to check that restricting F_{sec} to the cylinder Λ and then taking derivatives does not spoil the estimates. Indeed, recall the first order of the Hamiltonian that depends on the Poincaré variables ξ, η is of order $\frac{1}{L_2^6}$ (i.e. the first order term H_0^{12} in the expansion of the secular Hamiltonian, defined by (4)). Moreover the first order term of the graph ρ defining Λ that depends on the variable $\tilde{L}_3$ is of order $\frac{L_2^{10}}{\left(L_3^* \right)^7}$ (see Lemma (16)). Since H_0^{12} depends quadratically on ξ and η (see (4)), the term H_1^{12} is the lowest-order term that could contain products of the form $\tilde{L}_3 P$ for $P \in \{ \tilde{\Gamma}_3, \tilde{\Psi}_1, \tilde{\Gamma}_3 \}$; since the coefficient of H_1^{12} is of order $\frac{1}{L_2^7}$, the order of such terms is $\frac{1}{L_2^7} \frac{L_2^{10}}{\left(L_3^* \right)^7} = \varepsilon^4 \, \mu^7$. Therefore the estimates (E) imply the estimates on the mixed partial derivatives of $\hat{F}_0$ with respect to $\hat{L}_3$ and $\hat{\Gamma}_2$, $\hat{\Psi}_1$, $\hat{\Gamma}_3$ as required. $\qquad\square$

Acknowledgments

References

[1] V. I. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. *Uspehi Mat. Nauk*, 18(6 (114)):91–192, 1963.

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

[3] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. *Mathematical aspects of classical and celestial mechanics*, volume 3 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, third edition, 2006.

[4] P. Bernard. The dynamics of pseudographs in convex Hamiltonian systems. *J. Amer. Math. Soc.*, 21(3):615–669, 2008.

[5] P. Bernard, V. Kaloshin, and K. Zhang. Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders. *Acta Math.*, 217(1):1–79, 2016.

[6] S. Bolotin and D. Treschev. Unbounded growth of energy in nonautonomous Hamiltonian systems. *Nonlinearity*, 12(2):365–388, 1999.

[7] M. J. Capiński and M. Gidea. Arnold diffusion, quantitative estimates, and stochastic behavior in the three-body problem. *Communications on Pure and Applied Mathematics*, 2021.

[8] 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(1):329–360, 2017.

[9] C. Cheng. Dynamics around the double resonance. *Camb. J. Math.*, 5(2):153–228, 2017.

[10] C. Cheng and J. Yan. Existence of diffusion orbits in a priori unstable Hamiltonian systems. *J. Differential Geom.*, 67(3):457–517, 2004.

[11] L. Chierchia and G. Pinzari. Deprit's reduction of the nodes revisited. *Celestial Mechanics and Dynamical Astronomy*, 109(3):285–301, 2011.

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

[13] A. Clarke, J. Fejoz, and M. Guardia. Why are inner planets not inclined? *arXiv:2210.11311*, 2022.

[14] A. Clarke, J. Fejoz, and M. Guàrdia. Topological shadowing methods in arnold diffusion: weak torsion and multiple time scales. *Nonlinearity*, 36(1):426 – 457, 2023. Cited by: 0; All Open Access, Green Open Access.

[15] A. Clarke and D. Turaev. Arnold diffusion in multi-dimensional convex billiards. *Duke Mathematical Journal*, to appear, 2022.

[16] 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 $\mathbb{T}^2$. *Comm. Math. Phys.*, 209(2):353–392, 2000.

[17] A. Delshams, R. de la Llave, and T. Seara. Orbits of unbounded energy in quasi-periodic perturbations of geodesic flows. *Adv. Math.*, 202(1):64–188, 2006.

[18] 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. Amer. Math. Soc.*, 179(844):viii+141, 2006.

[19] A. Delshams, R. De La Llave, and T. M. Seara. Geometric properties of the scattering map of a normally hyperbolic invariant manifold. *Advances in Mathematics*, 217(3):1096–1153, 2008.

[20] A. Delshams, R. de la Llave, and T. M. Seara. Instability of high dimensional Hamiltonian systems: multiple resonances do not impede diffusion. *Adv. Math.*, 294:689–755, 2016.

[21] A. Delshams and G. Huguet. Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems. *Nonlinearity*, 22(8):1997–2077, 2009.

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

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

[24] M. Farhat, P. Auclair-Desrotour, G. Boué , and J. Laskar. The resonant tidal evolution of the earth-moon distance. *Astronomy & Astrophysics*, 2022.

[25] J. Fejoz. Quasiperiodic motions in the planar three-body problem. *Journal of Differential Equations*, 183(2):303–341, 2002.

[26] J. Féjoz. Démonstration du 'théorème d'Arnold' sur la stabilité du système planétaire (d'après Herman). *Ergodic Theory Dynam. Systems*, 24(5):1521–1582, 2004.

[27] J. Fejoz and M. Guardia. Secular instability in the three-body problem. *Archive for rational mechanics and analysis*, 221(1):335–362, 2016.

[28] J. Féjoz, M. Guàrdia, V. Kaloshin, and P. Roldán. Kirkwood gaps and diffusion along mean motion resonances in the restricted planar three-body problem. *J. Eur. Math. Soc. (JEMS)*, 18(10):2315–2403, 2016.

[29] J. Féjoz and S. Serfaty. *Deux cents ans après Lagrange*. Journée Annuelle, Paris, le 28 juin 2013. Société Mathématique de France, 2013.

[30] N. Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana University Mathematics Journal*, 21(3):193–226, 1971.

[31] N. Fenichel. Asymptotic stability with rate conditions. *Indiana University Mathematics Journal*, 23(12):1109–1137, 1974.

[32] N. Fenichel. Asymptotic stability with rate conditions, ii. *Indiana University Mathematics Journal*, 26(1):81–93, 1977.

[33] V. Gelfreich and D. Turaev. Unbounded energy growth in Hamiltonian systems with a slowly varying parameter. *Comm. Math. Phys.*, 283(3):769–794, 2008.

[34] V. Gelfreich and D. Turaev. Arnold diffusion in a priori chaotic symplectic maps. *Comm. Math. Phys.*, 353(2):507–547, 2017.

[35] M. Gidea and R. de la Llave. Topological methods in the instability problem of hamiltonian systems. *Discrete & Continuous Dynamical Systems*, 14(2):295, 2006.

[36] M. Gidea, R. de la Llave, and T. M-Seara. A general mechanism of diffusion in Hamiltonian systems: qualitative results. *Comm. Pure Appl. Math.*, 73(1):150–209, 2020.

[37] R. Gomes, A. Morbidelli, and H. F. Levison. Planetary migration in a planetesimal disk: why did neptune stop at 30 au? *Icarus*, 170:492–507, 2004.

[38] M. Guardia, J. Paradela, and T. Seara. A degenerate Arnold diffusion mechanism in the Restricted 3 Body Problem. Preprint, available at https://arxiv.org/abs/2302.06973, 2023.

[39] M. Guzzo, E. Lega, and C. Froeschlé. First numerical investigation of a conjecture by N. N. Nekhoroshev about stability in quasi-integrable systems. *Chaos*, 21(3):033101, 12, 2011.

[40] R. S. Harrington. Dynamical evolution of triple stars. *Astronom. J.*, pages 190–194, 1968.

[41] M. Herman. Some open problems in dynamical systems. In *Proceedings of the International Congress of Mathematicians (Berlin, 1998)*, volume Extra Vol. II, pages 797–808 (electronic), 1998.

[42] M. W. Hirsch, C. C. Pugh, and M. Shub. Invariant manifolds. *Bulletin of the American Mathematical Society*, 76(5), 1970.

[43] V. Kaloshin and K. Zhang. *Arnold diffusion for smooth systems of two and a half degrees of freedom*, volume 208 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2020.

[44] J. Lagrange and J. Bertrand. *Mécanique analytique*. Mallet-Bachelier, 1853.

[45] J.-L. Lagrange. Mémoire sur la théorie des variations des éléments de planètes et en particulier des variations des grands axes de leurs orbites. *Œuvres, t. VI*, pages 713–768, 1808.

[46] P.-S. Laplace. Sur le principe de la gravitation universelle et sur les inégalités séculaires des planètes qui en dépendent. Mémoire de l'Académie des sciences de Paris, Savants étrangers, année 1773. *Œuvres, t. VIII*, page 201, 1776.

[47] J. Laskar. Le système solaire est-il stable? *Séminaire Poincaré*, pages 221–246, 2010.

[48] E. Mathieu. Mémoire sur les inégalités séculaires des grands axes des orbites des planètes. *J. Reine Angew. Math.*, 80:97–127, 1875.

[49] R. Moeckel. Generic drift on Cantor sets of annuli. In *Celestial mechanics (Evanston, IL, 1999)*, volume 292 of *Contemp. Math.*, pages 163–171. Amer. Math. Soc., Providence, RI, 2002.

[50] N. N. Nehorošev. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. *Uspehi Mat. Nauk*, 32(6(198)):5–66, 287, 1977.

[51] L. Niederman. Stability over exponentially long times in the planetary problem. *Nonlinearity*, 9(6):1703–1751, 1996.

[52] G. Pinzari. *On the Kolmogorov set for many-body problems*. PhD thesis, Università degli Studi di Roma Tre, 2009.

[53] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste*. Gauthier-Villars, 1892.

[54] S. Poisson. Mémoire sur les inégalités séculaires des moyens mouvements des planètes. *J. École Polytechnique*, 8:1–56, 1809.

[55] D. Treschev. Evolution of slow variables in a priori unstable hamiltonian systems. *Nonlinearity*, 17(5):1803–1841, 2004.

[56] D. Treschev. Arnold diffusion far from strong resonances in multidimensional *a priori* unstable Hamiltonian systems. *Nonlinearity*, 25(9):2717–2757, 2012.

[57] J. Xue. Arnold diffusion in a restricted planar four-body problem. *Nonlinearity*, 27(12):2887–2908, 2014.

Andrew Clarke,
Departament de Matemàtiques, Universitat Politècnica de Catalunya, Av. Eduard Maristany, 16 08019 Barcelona, Spain
E-mail address: andrew.michael.clarke@upc.edu

Jacques Fejoz,
Université Paris Dauphine PSL, CEREMADE, Place du Maréchal de Lattre de Tassigny, 75016 Paris, France
Observatoire de Paris PSL, IMCCE, 77 avenue Denfert Rochereau, 75014 Paris, France
E-mail address: jacques.fejoz@dauphine.psl.eu

Marcel Guardia,
Departament de Matemàtiques i Informàtica, Universitat de Barcelona, Gran Via, 585, 08007 Barcelona, Spain
Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, 08193 Bellaterra, Spain
E-mail address: guardia@ub.edu