Stability in the plane planetary three-body problem

Thibaut Castan

Université Pierre et Marie Curie

École Doctorale d'Astronomie et d'Astrophysique d'île-de-France

Institut de mécanique céleste et de calcul des éphémérides

Thèse de Doctorat de l'Observatoire de Paris

Discipline : Mathématiques appliquées et applications des mathématiques

présentée par

Thibaut Castan

Stability in the plane planetary three-body problem

---

sous la direction de

Jacques Féjoz

Thèse soutenue le 21 avril 2017 devant le jury composée de :

M. Alain Chenciner	Professeur émérite	Examinateur
M. Luigi Chierchia	Professeur	Rapporteur
M. Jacques Féjoz	Professeur	Directeur de thèse
M. Vadim Kaloshin	Professeur	Examinateur
M. Jean-Pierre Marco	Maître de conférence HDR	Examinateur
M. Anatoly Neishtadt	Professeur	Rapporteur
M. Laurent Niederman	Maître de conférence HDR	Rapporteur
M. Eric Séré	Professeur	Examinateur

IMCCE
Observatoire de Paris
61, avenue de l'Observatoire
75 014 Paris
France

UPMC
4 place Jussieu
75005 Paris
France

Remerciements

Remerciements...c'est un mot qui sonne un peu formel pour finalement une occasion presque unique dans la vie de dire merci à ceux qui nous ont soutenus, ceux qui nous ont formés, à ceux qui nous ont accompagnés le long de ma thèse (qui a quand-même duré trois ans et demi, donc ça en fait du monde), mais aussi le long de ma vie. Sans toutes ses rencontres, sans toutes cette union de personne, je n'en serais probablement pas là aujourd'hui, et il est important que je puisse le dire ici, avec le coeur.
J'oublie sans doute des gens, et donc j'invite tout oublié à ne pas se formaliser et à me demander des remerciements de vive voix !

Le rituel veut que l'on débute par son directeur de thèse, cette personne qui vous a accompagné pendant toute la thèse (voire avant), soutenu, corrigé vos erreurs, su vous remettre sur la bonne route afin de pouvoir boucler sa thèse en temps (presque-)imparti. Alors merci, merci Jacques Féjoz. Merci de m'avoir donné cette opportunité de repartir dans les maths après beaucoup (trop?) de physique. Merci de m'avoir lancé dans un domaine passionant, dans lequel je pourrais m'épanouir tout au long de ma vie. Merci pour ta disponibilité quand j'en avais vraiment besoin, et merci de m'avoir laissé mon autonomie quand j'en avais besoin aussi. Merci pour tes bons conseils tout au long de ces années, pour ton énorme travail de relecture sur la fin de thèse, pour l'enrichissement que tu m'as apporté. Je vais m'arrêter là parce que sinon les remerciements seront plus long que la thèse, mais merci pour tout.
Continuons avec le jury (et par ordre alphabétique). Merci à Alain Chenciner d'avoir accepté d'être examinateur pour ma thèse. C'est un honneur d'avoir un professeur si reconnu dans son jury, qui est aussi un des leaders charismatique de l'équipe ASD. Merci pour tes conseils, j'espère ne pas t'avoir trop dérangé lors de mes discussions avec Jacques. Thank you Luigi Chierchia for accepting to review my thesis. I am honored and proud to have you in my jury, especially in this subject in which you are such a knowledgeable person. Thank you Vadim Kaloshin, to find the time to be present for my defense. Thank you as well for these Summer and Winter school you organized, for your courses that were so instructive. I am glad we can work together soon on some celestial mechanics problems. Merci à Jean-Pierre Marco d'avoir accepté d'être d'abord mon rapporteur, puis examinateur. Merci pour ta présence, et pour ta sympathie. Thank you Anatoly Neishtadt to have accepted to read this thesis and review it, as well as giving me some more references. Merci particulièrement à Laurent Niederman, qui n'a pas hésité une seule seconde à me "sauver" quand je n'avais plus qu'un rapporteur, et d'avoir accepté de relire mon manuscrit. Merci aussi pour tout tes conseils, pour tes idées, et peut-être une future collaboration sur Nekhoroshev. Merci à Eric Séré d'avoir accepté d'être examinateur (malgré l'heure de la soutenance), j'espère avoir l'occasion de faire plus ample connaissance et de m'instruire dans un domaine que je ne connais encore que trop peu.
De manière général, merci à toute l'équipe ASD, à son accueil, à sa sympathie. Merci à Jacques Laskar, qui est la première personne à m'avoir appris les coordonnées de Delaunay (à l'ENS), ça ne m'intéressait pas trop à l'époque, mais si j'avais su... Merci à mon Alain Albouy, avec qui je partage mon bureau, pour les réponses que tu as pu m'apporter qu'en j'en avais besoin, pour ces moments de travail paisibles au long de ces années. Merci à Philippe Robutel pour son aide précieuse sur des problèmes rencontrés lors de ma thèse. Merci à Hervé Manche et Gwenaël

Boué de me proposer de venir manger avec l'équipe malgré le fait que je ne mange pas souvent avec celle-ci, au moins je sais que je peux me joindre sans problème, et ça me fait plaisir. Merci à Mickaël Gastineau de toujours m'aider avec mes problèmes informatiques (merci de m'avoir aidé la fois où je n'avais plus internet, et que c'était simplement le câble ethernet qui n'était pas branché...). Merci à Abed Bounemoura pour tes conseils, et pour m'avoir aidé sur les nombres diophantiens depuis l'autre bout de la Terre. Merci à Agnès Patu d'avoir été si serviable pour les missions, et pour sa gentillesse. Merci aussi à tous les autres.

L'ordre dans lequel on doit faire des remerciements est finalement un vrai débat, j'ai choisi de suivre la flèche du temps, comme ça on commence avec la famille. Merci Papa, Maman, de m'avoir soutenu dans la vie, d'être à mes côtés, de m'avoir laissé faire mes erreurs, de m'avoir conseillé et aidé si je devais m'en sortir. Bref merci pour votre amour. Merci à ma soeurette de m'avoir enseigné les mathématiques et de m'avoir donné que des 20 pour éviter mes pleurs, tout ce que je sais c'est toi qui me l'as appris...Plus sérieusement, merci de ton soutien et de ton écoute. Merci à mes grands-parents, d'avoir apporté cette deuxième couche d'éducation qui est nécessaire. Dans la vie, j'ai l'impression qu'on essaie de faire du mieux possible non pas pour impressionner ses parents, mais plus pour rendre fier ses grands-parents. Alors merci à vous d'avoir jouer ce rôle, et d'une belle manière. Merci à Papi Yves de m'avoir donné ce goût pour les maths et les "égnimes", merci à Mamie Colette pour son soutien et son amour, merci à Papi Guy pour toute sa connaissance sur plein de choses (le jardinage, l'histoire, les petits villages où il fallait tourner à droite et pas à gauche), merci à Mamie Suzanne pour s'être occupée de moi quand j'étais petit. Merci à Tata l'Hélène et Tonton Dominique pour m'avoir comblé en jeu pendant ma jeunesse. Merci à Tonton Pierre-Yves pour ces discussions bizarres sur la science, mais aussi pour ces parties d'échecs (que je perds toujours), et pour ta bonne humeur. Merci à tous mes cousins pour tous ces jeux qu'on a pu faire pendant notre enfance ou après. Merci à la famille de Nantes, qui sont toujours très chaleureux, merci à la famille de Mur-De-Barrez, pour tous ces bons moments passés dans ma jeunesse. Merci aux autres aussi, que je vois moins souvent. Beaucoup de gens n'ont que peu de famille avec qui bien s'entendre, il est vrai que dans mon cas c'est l'inverse, je suis comblé par cette grande famille où tout le monde s'entraide, tout le monde est chaleureux. Merci d'avoir été là.

Pour ce qui est des amis, je me dois de commencer par le plus ancien: Quentin. On a fait tellement de choses ensemble, depuis qu'on a 5 ans, des barrages, des jeux, des bosses pour le vélo, des bosses pour la luge, des vacances inoubliables, et j'en passe. Si on est influencé par ses amitiés, c'est de toi que j'ai eu le plus d'influence, et je t'en remercie. Ensuite, je dois remercier toute la bande du BPC, avec qui j'ai passé toute mon enfance et mon adolescence. Merci bien sûr à Recep, Julian, Fifi, Ronan, Adrien, Jojo, Pierre, Nico, Fanou, Cyril, Le Chaux, Mus, Cuadro, Romain, ainsi qu'à leur compagne Manu, Chloé, Coco, Sarah, Betul, et j'en passe. Merci pour tous ces moments passés avec vous, ces parties de foot, ces barbecues à Aubusson, ces soirées à Borbe...

Merci à mes amis de collèges qui m'ont beaucoup marqué: Jérémy, Sylvain, Alexandre, Martin, Antony, Quentin, Ozkan.

Merci à mes amis de lycée, avec qui on a passé de très bonnes années. Merci tout particulièrement à Anaëlle, qui était aussi en prépa avec moi, et qui m'a souvent aidé à passer les épreuves par une simple conversation. Merci à Anne, Pierre, Béné, Ivan, Mathilde.

Merci à mes potes de prépas, qui m'ont aidé à me relâcher pendant ces années, et m'ont permis de les prendre "à la cool". Merci à Gilles, Mohammed, Mathieu, Marc, Anthony.

Bon, alors là on va rentrer à l'ENS et ça devient compliqué, une foultitude de nom, je vais en oublier pour sûr, donc il ne faut pas m'en vouloir.

Merci à mes premiers potes de l'ENS pour m'avoir aidé à développer mes théories dans la bonne humeur. Merci Robert pour m'avoir enseigné l'anglais de manière détournée lors de tes conversations avec Nofer (que je remercie pour m'avoir appris la vraie "Lad'attitude"). Merci à

Pierre et Jonathan de n'être qu'un, et d'avoir une vision aussi pure sur le future des maths, qui passe évidemment par l'invention du serpent Brownien. Merci à Hélène pour ses gâteaux (et aussi de ne pas m'avoir violenter). Merci à Ségolène pour sa motivation extrême, pour son livre sur le zéro déchet (que je n'ai toujours pas lu). Merci à Elsa pour sa sportivité, l'athlète du groupe, avec Pierre-Alain. Merci à Brice et Nelly, pour chanter avec moi au fond de l'église lors des mariages des autres. Merci à Ruben et Clothilde, pour tant de moments arméniens et ne pas m'en vouloir quand j'oublie des anniversaires. Merci à Xavier, pour sa connaissance de la religion, qui ne m'a pas encore servi, mais qui a levé le voile sur des choses que je ne comprenais pas. Merci à Marine pour sa spontanéité.

Merci à Antoine Hocher, qui a été mon premier pote de physique, avec qui j'ai passé énormément de bons moments, qui m'a permis de voir qu'on faisait pas n'importe quoi en biologie, et que même c'était très intéressant. Merci à Raphaël, avec qui on a des discussions passionantes sur les systèmes intégrables, et avec qui j'espère on réussira à bosser ensemble sur un projet. Merci à Rémi, à ses précieux conseils, à sa connaissance des maths, et à son support. Merci à Hatchu, pour les bons moments passés ensemble. Merci à Cyrus pour sa folie, et sa folie des maths. Merci à Garreau, pour avoir réussi à m'intéresser à des problèmes de statistiques alors que bon, c'est pas gagné.

Merci à Louise, pour les weekends inoubliables à Mareuil. Merci à Nil et Willy, pour les pendaisons de crémaillères à Londres. Merci à John, pour son enseignement sur la culture rastafari. Merci à Léo, pour son entrain et son humour. Merci à Gab, Léa etc pour les échanges qu'on a pu avoir.

Merci à l'équipe K-Fêt, pour les soirées, les conversations existentielles, les après-midi en courô, en vrac: Romain, Oscar, Joran, Dréa, Edith, Fermé, Marie x2, Axelle, Zaïd, Lepot, Trousse...

Merci aux autres groupes de potes que j'ai croisé et avec qui je peux toujours compter pour passer de bonnes soirées, je pense en particulier à Cheisson, Sarah, Anne-Laure, Francès, Cyril, mais aussi Laurie, ou encore Séb, et bien d'autres mais j'en oublie encore.

Je dois ensuite passer à mes deux premières colocataires, Pélagie et Barbara. Pélagie, merci pour ton support, pour tes rires, pour tes invitations aux soirées d'halloween où finalement je suis venu, merci pour nos soirées passées à se raconter notre vie plutôt que de travailler. Barbara, merci pour ta bonne humeur, pour t'occuper des chats, pour tes invitations au match de basket.

Merci à mes potes de l'observatoire, qui ont égayé ma vie ici, et que j'espère bien revoir après nos post-docs. Merci à Adrien x2, merci à César (doublement pour en plus m'aider avec le pôt de thèse !), merci à Pierre.

Un très grand merci à Jessica, qui m'a beaucoup aidé en maths quand j'en avais besoin, et qui m'a aussi beaucoup permis de comprendre des choses qui n'était pas facile. Je fais écho à ta thèse en disant que je n'étais pas présent quand tu stressais, mais là tu n'es pas présente non plus avant ma thèse, donc un partout. Mais merci de m'aider quand j'en ai besoin, et de réclamer de l'aide quand tu en as besoin. Merci à Alex de m'avoir fait découvrir le "Loir-et-Cher", parce que je ne sais pas ce que j'aurai pu faire sans cette musique.

Merci à Mathieu Baron, de m'avoir fait découvrir qu'on pouvait faire des soirées autrement, même à Paris, et de s'être motivé pour monter nous-même notre association. Merci donc à la Mangrove, pour ses moments de stress mais de purs plaisirs, et d'avoir créer un vrai groupe qui vit pour la même chose, merci en vrac (et il en manque) à Latiss', Galloche, Vince, Rémy, Thomas-Jacques, Lucie, Akima, MHF, Yoshi, Mathieu...Merci à notre cher cuisinier aussi de s'être proposé de m'aider pour le pôt, merci Prosper.

Merci à Basile d'être devenu mon pote avec qui je vais voir des matchs de rugby, parce que cette année sera la bonne pour l'ASM.

Merci quand-même à Géraldine.

Contents

Résumé

Arnold a démontré l'existence de solutions quasipériodiques dans le problème planétaire à trois corps plan, sous réserve que la masse de deux des corps, les planètes, soit petite par rapport à celle du troisième, le Soleil. Cette condition de petitesse dépend de façon cachée de la largeur d'analyticité de l'hamiltonien du problème, dans des coordonnées transcendantes. Hénon explicita un rapport de masses minimal nécessaire à l'application du théorème de Arnold. L'objectif de cette thèse sera de donner une condition suffisante sur les rapports de masses. Une première partie de mon travail consiste à estimer cette largeur d'analyticité, ce qui passe par l'étude précise de l'équation de Kepler dans le complexe, ainsi que celle des singularités complexes de la fonction perturbatrice. Une deuxième partie consiste à mettre l'hamiltonien sous forme normale, dans l'optique d'une application du théorème KAM (du nom de Kolmogorov-Arnold-Moser). Il est nécessaire d'étudier le hamiltonien séculaire pour le mettre sous une forme normale adéquate. On peut alors quantifier la non-dégénérescence de l'hamiltonien séculaire, ainsi qu'estimer la perturbation. Enfin, il faut démontrer une version quantitative fine du théorème KAM, inspirée de Pöschel, avec des constantes explicites. A l'issue de ce travail, il est montré que le théorème KAM peut être appliqué pour des rapports de masses entre planètes et étoile de l'ordre de 10^{-85}.

Abstract

Arnold showed the existence of quasi-periodic solutions in the plane planetary three-body problem, provided that the mass of two of the bodies, the planets, is small compared to the mass of the third one, the Sun. This smallness condition depends in a sensitive way on the analyticity widths of the Hamiltonian of the three-body problem, expressed with the help of some transcendental coordinates. Hénon gave a minimal ratio of masses necessary to the application of Arnold's theorem. The main objective of this thesis is to determine a sufficient condition on this ratio. A first part of this work consists in estimating these analyticity widths, which requires a precise study of the complex Kepler equation, as well as the complex singularities of the disturbing function. A second part consists in reworking the Hamiltonian to put it under normal form, in order to apply the KAM theorem (KAM standing for Kolmogorov-Arnold-Moser). In this aim, it is essential to work with the secular Hamiltonian to put it under a suitable normal form. We can then quantify the non-degeneracy of the secular Hamiltonian, as well as estimate the perturbation. Finally, it is necessary to derive a quantitative version of the KAM theorem, in order to identify the hypotheses necessary for its application to the plane three-body problem. After this work, it is shown that the KAM theorem can be applied for a ratio of masses that is close to 10^{-85} between the planets and the star.

Introduction

The Universal Attraction of bodies masterly explained Kepler's Laws of planetary motion in our Solar System. But Newton himself understood thaat the mutual attraction of planets could destroy the striking regularity that can be observed over a short period of time. This thesis aims at contributing to this question of the stability of motion of planets. The main goal is the application of the KAM theorem (which stands for Kolmogorov-Arnold-Moser) to the three-body problem, that gives the stability of this system over an infinite time. Arguably, this theorem is not directly relevant in Astronomy because the hypotheses (a much simplified model of our Solar System) and the conclusions (infinite time stability for some initial conditions) are too strong. Two main interests actually derive from the first one. First, again from a physical point of view, we extract some important computation on the complex collisions of the plane planetary three-body problem. This work, necessary to apply the KAM theorem, will allow one to apply stability theorems over finite time, which could lead to meaningful stability time in some stellar systems. The second interest is the precise study and statements of the theorems that are needed to carry out this study, as well as a complete construction of the secular and perturbed system related to the three-body problem. The tools derived here can come in handy when studying a system with similarities with the present system.

Let us give the historical interest of this work (consult [27, 39, 22, 5, 6] to have a more precise introduction on the subject). The precise astronomical observations performed in the 16^{th} and 17^{th} century led to empirical laws known as Kepler's laws of planetary motions. Bypassing the beliefs of his time, Hooke and Newton [53] managed to give an analytic explanation of these laws, by correlating the gravity (in its Latin meaning) occurring on Earth, and the motion of the planets: the Newton's laws of motion were born. The equation of motions given by Newton's laws take the form of a differential equation, and it was a trait of genius of Newton to solve these equations for the two-body problem. Yet, the deterministic approach, that consists in determining the state of the system in the future using the state of it in a near past showed its limit. The equation of motion of the three-body problem, which is no more an integrable system, is far from being completely understood, given the difficulty of the computations involved. Laplace, who believed in the deterministic approach [38], and Lagrange developed a secular solution (see [16] and [17]) of this problem to fit the observations as much as possible. This secular solution gave good results, and allowed Adams [2] and Le Verrier [42] to discover the existence of Neptune. Nevertheless, the secular solution does not give an exact solution in time. Poincaré and his ideas completely turned the problem around, changing the nature of the question. Instead of studying the solution in time, one can study the motion of the system qualitatively. It would be difficult to enumerate all the advances he made on celestial mechanics or in mathematics related to it, though some decisive work can be found in [58, 59, 60, 61]. In particular, he showed that it was not possible to determine an analytical solution over an infinite time using only integrals and algebraic expressions. Sundman, in 1909 in [68], found an analytic solution of the three-body problem, but as a slowly converging time series, and hence of no theoretical use. The study of the general behavior of non-integrable systems, in addition to the Poincaré recurrence lemma, could let think that the ergodic hypothesis was holding in the solar system. In this case, any attempt to prove some kind of stability in the solar system would have been vain. This idea was completely shoved by Kolmogorov in 1954 with his theorem [36] which showed the subsistence of some tori when perturbing an integrable Hamiltonian. In 1963 [4], Arnold proved that this theorem could be applied to the plane planetary three-body problem, hence showing that some stability exists under some conditions in solar systems. This theorem was adapted to smooth Hamiltonian by Moser [31], to finally give the KAM theorem its famous name. Although Hénon, in a letter to Arnold [29] in 1966, noticed that the KAM theorem as it was proven by Arnold could not be applied to the Solar System, the idea of a possible stability resurfaced for a wide class of Hamiltonian systems nearly integrable. Nekhoroshev proved in 1977 [52] a theorem of stability in finite time. Lots of results were then derived from this theorem, so as to prove different cases of stability in the solar system, see for example [46, 47, 26]. An attempt of application in the case of a resonant three-body problem was done by Niederman [54] in 1995, to show some stability under a period equivalent to the lifespan of the solar system, though the masses of the planets still needed to be very small. Besides, some computer-assisted proofs on the application of the KAM theorem to some systems composed of three bodies have been done by Chierchia, Celletti, Giorgilli, Locatelli and Robutel ([11, 12, 10, 45, 63]). Finally, it was recently shown by Herman and Féjoz [20] that the KAM theorem could be applied to the N-body problem, this result being proven differently by Chierchia and Pinzari in 2011 [15], following more closely Arnold initial strategy. Yet, at this time, no attempt was made to apply in a quantitative way the KAM theorem to the three-body problem, or more precisely to find some physical constraints to have stability over an infinite time for a system. One of the requirement to apply the KAM theorem to the plane planetary three-body problem is the precise study of the perturbation, which is as well necessary to apply Nekhoroshev-type theorems. Hence, this work could lead to better stability time in this case, approaching even more the real condition of stability in the solar system.

Let us consider the planetary regime of the three-body problem in the plane; "planetary" means that two masses (planets) are small with respect to the third one (Sun), and that each planet revolves around the Sun. These assumptions emanate from our Solar System, where the mass of Jupiter is 10^{-3} the mass of the Sun. In this regime, the two planets mainly undergo the attraction of the Sun, so the system reduces to two uncoupled Kepler problems. Newton stupendously solved such Kepler problems. In particular, bounded orbits are periodic, surprisingly, whereas one would expect, in general, quasi-periodic motions with two frequencies. This well-known but mysterious *proper degeneracy* (in the terminology of Arnold), specific to the Newtonian and elastic potentials, according to Bertrand's theorem, is a source of difficulties: one cannot expect quasi-periodic perturbed motions. If they existed, they would have the same number of frequencies. Thus the problem is a singular perturbation problem.
The full planetary three-body problem, when one takes into account the mutual attraction of planets, is not integrable, so we are reduced to indirect descriptions of its solutions. Outside low order Keplerian resonances, the next order approximation of the dynamics is described by the Hamiltonian obtained by averaging the initial Hamiltonian along the Keplerian 2-tori; averaging here looks like spreading the potential of mutual attraction along the Kepler ellipses of the planets. The so obtained Hamiltonian, called *secular* is integrable, as counting dimensions and taking into account the invariance of the problem by rotations show. The secular dynamics has an elliptic fixed point located at circular Kepler ellipses, which Arnold calls a *limit degeneracy*. Lagrange and Laplace have studied the eigenmodes of this fixed point. We ourselves localize our study in the neighborhood of the fixed point, for both astronomical reasons and the sake of simplicity. In order to simplify even further the complications, we focus on the sub-region of the phase space where the semi-major axis ratio is small. It is then relevant to expand the secular Hamiltonian in averaged Legendre polynomials, truncate the series, and compute its Birkhoff normal form.
When all small parameters (planet masses, semi-major axis ratios, eccentricities) are small,

the size of the remainder of the normal form tends to zero. This goes with the competing phenomenon that the Birkhoff invariants of the normal form go to zero too, the first invariant being the secular frequencies, and the second the so-called torsion. But in the KAM theorem the size of the allowed perturbation (on some complex extension of the phase space, for a result in the analytic class) depends crucially on the first two Birkhoff invariants. Hence, the whole construction, and the used version of the KAM theorem, must say quantitatively how the invariants degenerate as a function of the smallness of the perturbation.

The main result of this thesis is theorem 6.1, which is a quantitative estimate of the mass ratio such that the KAM theorem can be applied to the plane planetary three-body problem. Roughly, we show that the KAM theorem can be applied if the ratio of mass between the planets and the star is close to 10^{-85}. The full conditions are given in chapter 6. First, we localize our study where the ratio of distances between the planets is very large, guaranteeing that our development of the secular Hamiltonian is relevant to the problem. Secondly, the eccentricities are very small, so that the terms of order more than 4 in eccentricities are part of the perturbation. Finally, we ask the frequencies to verify some Diophantine condition. We recall here the definition of the Diophantine set for constants γ and τ:

$$D(\gamma,\tau) = \left\{\omega \in \mathbb{R}^n : \ \forall k \in \mathbb{Z}^n, \quad |k \cdot \omega| \geq \frac{\gamma}{|k|_1^\tau}\right\},$$

where $|.|_1$ is the l_1-norm. We ask that the two frequencies related to the fast angles verify an optimal Diophantine condition in γ with $\tau = 2$ (this optimal condition is made clear in chapter 5). Secondly, the full frequency vector (of four dimensions) must verify a Diophantine condition with $\tau = 4$ and $\gamma = 10^{-26} \min(\omega_i)$. Under the conditions we enumerated, the perturbation of the three-body problem can be "absorbed" by the secular Hamiltonian, and some quasi-periodic solutions exist.

The first chapter of the thesis is dedicated to the study of the complex collisions, as well as a computation of the norm of the perturbation on a complex neighborhood of the initial conditions in Poincaré variables. The change of variables between the eccentric and the mean anomaly requires the study of Kepler's equation. The latter equation defines a diffeomorphism between the two different angles. Hence, it is fundamental to find a complex neighborhood on which this equation continues to define an analytic diffeomorphism. The study of the singular points of this change of variable is essential in the three-body problem, since they correspond as well to complex collisions between the planets and the star. With this computation, it is then possible to find a dominating series that approaches the series expansion of the perturbation well. We give a simple upper bound of the perturbation in a complex domain depending in the width of this domain, so as to know the rate of growth of the perturbation close to the real domain. In a second chapter, we state and proof quantitative theorems that will allow us to split the Hamiltonian into an unperturbed part, that is not degenerate, plus a perturbation. We focus on two different classical transformations. First, we are interested in an explicit Birkhoff normal form (BNF) theorem which puts the averaged Hamiltonian under a suitable normal form. Secondly, to be able to apply the KAM theorem, the frequencies of the final perturbation need to be small compared to the frequencies of the unperturbed part, which is not the case at first, since the angles of the periapses are given by the initial perturbation and are small. It is therefore necessary to perform a transformation that keeps the small frequencies and pushes to a higher order the perturbing frequencies. In a third chapter, we will give a quantitative statement of the KAM theorem we will use for the application. Several statements of the KAM theorem exists, with their own characteristics. Here we choose the version of Pöschel [62], that is a KAM theorem with parameters. This choice is a personal choice, after the comparison with other KAM theorems (such as the one developed by Féjoz in [24], which consists in performing a local inversion theorem), which are less fit to quantitative applications. This KAM theorem relies on letting the frequencies become a parameter, which will change at every step. In this configuration, it is sufficient that the unperturbed Hamiltonian, depending on the action and the new parameter, is only linear in the actions (the torsion being included in the computation of the parameter). One can then derive a KAM theorem, with explicit constants and hypotheses. We also gather the information on the different sizes of the transformations involved along the proof, so as to know the distance between the initial geometry of the system and the position of the torus where the motion takes place in the new variables. We can then apply precisely the KAM theorem in the fourth chapter. Before the straight application of the KAM theorem, we need compute the secular Hamiltonian. First, one needs to compute the expansion of the perturbation explicitly as well as its average over the fast angles (the mean longitudes). Then, with the help of some transformations and the theorems of the second chapter, we put the Hamiltonian into a suitable form to be able to apply the KAM theorem. All these transformations need to be quantitatively explicit, so as to compute the loss of analyticity width related to them. Keeping track of the remainders of each operations and adding them to the perturbation, we compute a bound on the norm of the final perturbation. Finally, we express a way to determine the analyticity width in frequencies given by the unperturbed Hamiltonian, which is an important value for the application of the KAM theorem. In a fifth chapter, we are interested in the Diophantine condition. We recall some basics of this condition, such as the measure of the vectors satisfying it, as a reminder of the importance that had the KAM theory in the field of dynamical systems. We then follow with the study of Diophantine numbers, which are the irrational that are badly approximated by rationals, this condition being well described with the help of the continued fraction theory. We introduce a visual tool to understand better how to work with these numbers, and that proves to be useful when trying to determine if an irrational is Diophantine or not. With the help of the previous work, we justify the choice of Diophantine condition we chose to apply the KAM theorem, introducing some optimality property regarding this condition. Finally, to understand better where the KAM theorem can be applied, we study the presence of Diophantine numbers and their constants in the interval $[0, 1]$, which showcases the importance of the resonances in the three-body problem, as the existence of the Kirkwood gaps shows. In the last and sixth chapter, we introduce the final results of this application. There exist lots of subtleties in this application, and the various choices made are discussed. Moreover, we get interested in the possible improvements, such as the possibility of removing some restrictions that have been made in the computation. Indeed, we give some clues, so as to weaken the hypotheses we made. Several easy points to improve this application are given, as well as other relevant points that would be more difficult to implement.

Chapter 1

Analyticity, singularities, bound on the Hamiltonian norm

This chapter is dedicated to the study of the perturbation in the plane planetary three-body problem. Although a lot of theorems, from the KAM theorem to theorems of stability in finite time, use the analyticity of the perturbation of an integrable Hamiltonian, only few attempts were made to compute the actual analyticity widths in the three-body problem. Regarding the perturbation, another problem kept the attention of researchers for years: the convergence of the expansion in power of the eccentricities of the disturbing function. This matter is closely related to Kepler's equation, which associates the mean anomaly to the eccentric one, and in particular to the convergence of the inverse Kepler equation. This study started with the work of mathematicians studying the celestial mechanics such as Lagrange and Laplace. Later, Poincaré provided a necessary condition on the elliptic variables for this expansion to converge (see [57]). This work was followed by the statement of a sufficient condition by Silva [66] that was proven by Sundman to be as well a necessary condition [69]. Wintner showed in [71] some conditions on complex eccentricities this time in Kepler's equation: he showed that although the inverse converges for real eccentricities strictly smaller than 1, the size of the disk of convergence for complex eccentricities is given by Laplace's limit ~ 0.66274). Some works regarding the series expansion and the complex singularities have been developed by Petrovskaya in 1970 [56]. Sokolov [67] went further in his research by looking not only at complex eccentricities, but also at complex mean anomalies.

When trying to apply an analytic version of a theorem of stability (KAM or Nekhoroshev), we require two different things. First, the variables must be symplectic variables to be able to apply Hamilton's equation (hence making the previous researches in elliptic coordinates out of use); secondly, we have to consider a complex neighborhood of a real domain for each variable. This problematic has been studied by Niederman [54], while trying to apply a Nekhoroshev's theorem to the planetary three-body problem. To perform this work, we choose to study the plane case, and use the Poincaré coordinates. Considering a complex neighborhood of these coordinates, we are able to determine a polydisc in which Kepler's equation induces a diffeomorphism, as well as a sufficient condition to avoid the singularities of the perturbing function. We are also able to bound the norm of the perturbation on some complex domain, in which it is analytic. Besides, we consider anisotropic polydiscs to obtain better result; we use three different analyticity widths: one for the action Poincaré variables, one for the angles, and one for the Cartesian coordinates. With this result, it is now possible to consider relating the stability time of a system with its initial geometry.

We choose to work in the Jacobi framework ([58]), and use the expansion in semi-major axes of the perturbing function. This choice has two main interests: first, there is no complementary term that appears as opposed as in the heliocentric framework, and secondly, for a suitable choice of masses, the perturbation can be made smaller. After introducing Poincaré variables, we have to consider the analytic continuation of the perturbation, and therefore, a discussion about this step will be necessary. Then, we study deeply Kepler's equation. First, if possible, we determine the singularities of this equation explicitly, and if not, try to approach them precisely. Secondly, we establish a domain on which the change of variables induced by Kepler's equation is a diffeomorphism. With this work, we are finally able to bound the norm of the perturbing function in the plane planetary three-body problem.

1.1 Preliminaries

1.1.1 Initial Hamiltonian and Jacobi coordinates

In the plane planetary 3-body problem, we consider a massive body, called the star, and two lighter bodies, called planet 1 and planet 2, orbiting the star. Call (p_0, q_0) the coordinates (momentum-position) of the star, and (p_i, q_i) the coordinates of the planet i, with $i \in \{1, 2\}$. The phase space we are going to consider is obtained by removing the collisions between the bodies:

$$D = \left\{ (p_i, q_i)_{0 \leq i \leq 2} \in \left(\mathbb{R}^{2\star} \times \mathbb{R}^2 \right)^3 \mid \forall 0 \leq i < j \leq 2,\ q_i \neq q_j \right\}. \tag{1.1}$$

The Hamiltonian of this isolated system is the following:

$$H(p_0, p_1, p_2, q_0, q_1, q_2) = \frac{1}{2} \sum_{0 \leq j \leq 2} \frac{|p_j|^2}{m_j} - \sum_{0 \leq i < j \leq 2} \frac{G_{grav} m_i m_j}{|q_j - q_i|}, \tag{1.2}$$

where m_0 is the mass of the star, m_i the mass of the planet i, G_{grav} the gravitational constant (we will keep track of this constant along the calculations), and $|\cdot|$ is the Euclidian norm.
We wish to use the Jacobi coordinates, so we do not have to deal with the complementary term $p_1 \cdot p_2$ and can focus only on the q_i. We will follow the work of Féjoz [19] to obtain these coordinates. First, one needs to consider the following symplectic change of variables:

$$\begin{cases} P_0 = p_0 + p_1 + p_2 \\ P_1 = p_1 + \sigma_1 p_2 \\ P_2 = p_2 \end{cases} \qquad \begin{cases} Q_0 = q_0 \\ Q_1 = q_1 - q_0 \\ Q_2 = q_2 - \sigma_0 q_0 - \sigma_1 q_1 \end{cases}$$

where the two coefficients σ_0 and σ_1 are given by $\sigma_0 = \dfrac{m_0}{m_0 + m_1}$, $\sigma_1 = \dfrac{m_1}{m_0 + m_1}$. The new coordinates are represented in figure 1.1.

Now, let us consider the following masses, these choices simplifying the expansion of the perturbing function in powers of the semi-major axes:

$$\mu_1 = \frac{m_0 m_1}{m_0 + m_1},\ \mu_2 = \frac{(m_0 + m_1) m_2}{m_0 + m_1 + m_2},$$
$$M_1 = m_0 + m_1,\ M_2 = m_0 + m_1 + m_2.$$

Considering the frame of reference attached to the center of mass of the system, and $Q_2 \neq 0$ (which will always be satisfied in the planetary motion), the Hamiltonian H of the system is:

$$H(P_1, P_2, Q_1, P_2) = H_{Kep}(P_1, P_2, Q_1, P_2) + H_{pert}(P_1, P_2, Q_1, P_2),$$

$$\begin{cases} H_{Kep}(P_1, P_2, Q_1, P_2) = \dfrac{|P_1|^2}{2\mu_1} - \dfrac{G_{grav} \mu_1 M_1}{|Q_1|} + \dfrac{|P_2|^2}{2\mu_2} - \dfrac{G_{grav} \mu_2 M_2}{|Q_2|} \\ \\ H_{pert}(P_1, P_2, Q_1, Q_2) = G_{grav} m_2 \left(\dfrac{m_0 + m_1}{|Q_2|} - \dfrac{m_0}{|Q_2 + \sigma_1 Q_1|} - \dfrac{m_1}{|Q_2 - \sigma_0 Q_1|} \right) \end{cases} \tag{1.3}$$

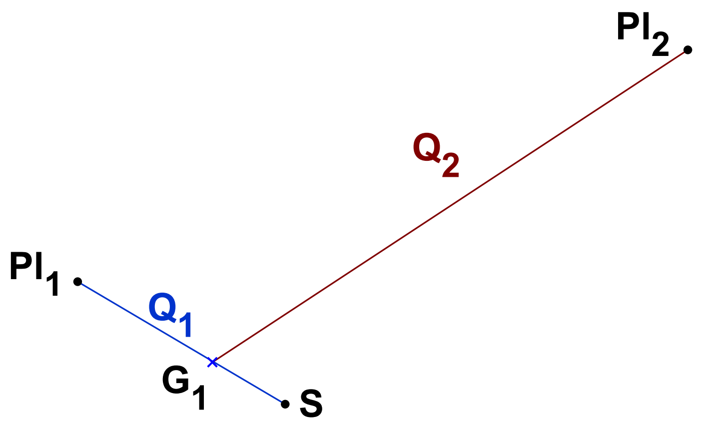

Figure 1.1: In blue, the coordinate Q_1 represents the vector between the star S and the interior planet Pl_1. In brown,, the coordinate Q_2 represents the vector between the center of mass G_1 of the system { Star - Interior Planet } and the exterior planet Pl_2

Observe that the Hamiltonian does not depend on Q_0 (because of the symmetry by translation). The system is then described by the 4 Jacobi coordinates (P_1, P_2, Q_1, Q_2), outside the collisions.
The first part H_{Kep} can be thought of as the sum of two integrable fictitious Kepler problem Hamiltonian. The first one corresponds to the planet 1 orbiting the star (with fictitious masses), and the second one to the planet 2 orbiting the center of mass of the star and the planet 1 (as well with fictitious masses).
One can now develop formally this formula in expansion of the semi-major axes, obtaining the following series:

$$H_{pert}(P_1, P_2, Q_1, Q_2) = \frac{G_{grav}\mu_1 m_2}{|Q_2|} \sum_{n\geq 2} \sigma_n P_n(\cos(S)) \left(\frac{|Q_1|}{|Q_2|}\right)^n, \quad \text{with } \sigma_n = \sigma_0^{n-1} + (-1)^n \sigma_1^{n-1}, \tag{1.4}$$

where $P_n(\cos(S))$ is the n^{th} Legendre polynomial and S is the oriented angle between Q_1 and Q_2.

Observe that in Jacobi coordinates, the series starts at $n \geq 2$, and therefore will decrease quickly with the ratio of the semi-major axes.
The perturbation of the Kepler problem Hamiltonian contains three terms, that are proportional to

$$\frac{1}{|Q_2|}, \qquad \frac{1}{|Q_2 + \sigma_1 Q_1|}, \qquad \frac{1}{|Q_2 - \sigma_0 Q_1|}.$$

The singularities in these coordinates corresponds to the points where the denominators are null. When looking at the analytic continuation of these functions, the two last terms depend on 8 complex variables. The difficulty of solving such an equation makes it relevant to start studying the first term, which depends only on 4 complex variables.

1.1.2 Reminder on the plane Delaunay and Poincaré variables

When studying celestial mechanics, physicists use mainly the elliptic coordinates. These coordinates prove to be very useful when studying the Kepler problem. Indeed, they simplify

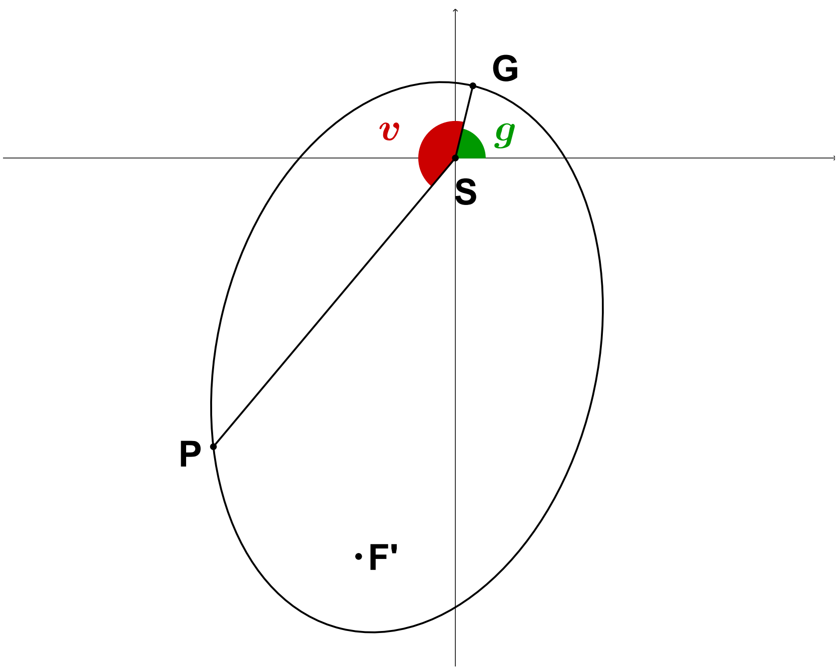

Figure 1.2: The periapsis of the ellipse of foci S (the star) and F' is in G. The green angle is g, the angle of the periapsis. The planet is on the ellipse, in P, and in red is represented the true anomaly v.

greatly the expression of the solution, taking their roots in basic geometry of the ellipse. In the plane case, those are written (a, e, v, g), where a is the semi-major axis, e the eccentricity, v the true anomaly and g the angle of the perihelion. In these variables, one can consider the ellipse given by the values (a, e, g), the position of the planet being determined by v. In a Cartesian framework, we have the relations

$$\begin{cases} x = r\cos(v+g) \\ y = r\sin(v+g) \\ z = x + \imath y = r\exp(\imath v)\exp(\imath g) \end{cases} \tag{1.5}$$

where r is the distance from the planet to the star. The two variables (v, g) are represented on figure 1.2.
Nevertheless, these coordinates are not symplectic, and one cannot work directly on the Hamiltonian equations with those. It is therefore necessary to change these variables, though trying to keep the simplicity of the expression of Kepler's laws of planetary motion.
A first set of symplectic variables that can be used is the Delaunay variables. To see their construction, one can refer to [14] and [40]. Before introducing them, let us define the eccentric anomaly u, another common angle in celestial mechanics or in the study of ellipses. It satisfies the relations:

$$\begin{cases} r\cos(v) = a(\cos(u) - e) \\ r\sin(v) = a\sqrt{1-e^2}\sin(u) \end{cases} . \tag{1.6}$$

The construction of the eccentric anomaly u, as well as the mean anomaly l (that will be defined hereinbelow), is shown in figure 1.3, where we let $g = 0$ to simplify the drawing. The Delaunay coordinates are the coordinates (L, G, l, g): from the elliptic coordinates, they can be

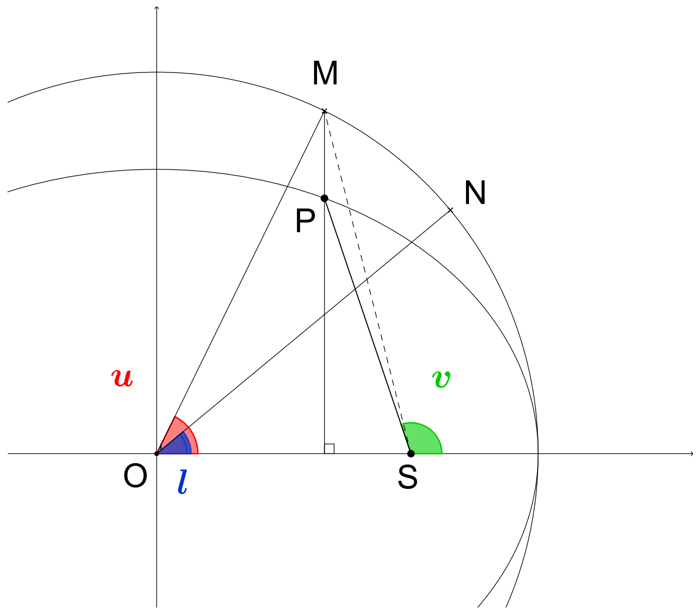

Figure 1.3: S represents the star, P the planet and O the center of the two foci of the ellipse. The circle is the circle of center O and passing through the periapsis of the ellipse (here $g = 0$). In green is represented the true anomaly v, in red the eccentric anomaly u, and in blue the mean anomaly v.

defined as follows,

$$\begin{cases} L = \mu\sqrt{G_{grav}Ma} \\ G = L\sqrt{(1-e^2)} \\ l = u - e\sin u = u - \Im\left(e\exp(\imath u)\right) \end{cases} \tag{1.7}$$

and g being the angle of the perihelion. G is the angular momentum, and l is given by the Kepler's equation, to which we shall return later. Those coordinates are therefore action-angle variables. Observe that for $L = G$, in other words for a zero eccentricity, the angle of the perihelion is not defined. This singularity can lead to issues in computation, since it introduces a singular point due to the definition of these coordinates.

This problem can be avoided using Poincaré coordinates, which do not present a singularity for zero eccentricities. For a complete construction of these variables, peruse [14] and [21]. In the plane case, those are composed of two action-angle coordinates, and two Cartesian coordinates for each body: $(\Lambda, \lambda, \xi, \eta)$. Their formulas can be summarized, using Delaunay variables, as follows:

$$\begin{cases} \Lambda = L \\ \lambda = l + g \\ \Gamma = L - G \\ \xi = \sqrt{2\Gamma}\cos(-g) \\ \eta = \sqrt{2\Gamma}\sin(-g) \end{cases} . \tag{1.8}$$

The angle λ is called the mean longitude (we could have defined as well the eccentric longitude adding g to u). To simplify the use of the Cartesian coordinates, we can define the following variable:

$$F = \xi + \imath\eta = \sqrt{2\Gamma}\exp(-\imath g) = \sqrt{2\Lambda}\sqrt{1 - \sqrt{1-e^2}}\exp(-\imath g).$$

Another variable that can prove to be useful while switching from Delaunay point of view to Poincaré's one is the eccentricity vector $E = e\exp(\imath g)$. We then have the following relation:

$$F = \frac{\sqrt{2L}}{\sqrt{1+\frac{G}{L}}}\bar{E}.$$

Recall now that the Hamiltonian of the two body problem can be written

$$H = \frac{G^2}{2\mu r^2} - \frac{G_{grav}M\mu}{r}.$$

In the planetary case, the distance r of the planet to the star oscillates between the extremal values $a(1+e)$ and $a(1-e)$ on its ellipse. Those are the roots of the polynomial of degree 2 given by the Kepler problem Hamiltonian (the energy being constant along the motion). In Poincaré coordinates, it takes the form

$$H = -\frac{G_{grav}^2 M^2 \mu^3}{2L^2}.$$

Replacing L by the variable Λ gives directly its expression in Poincaré coordinates.

1.1.3 Analytic continuation

The analytic continuation consists in expanding some function in a complex neighborhood of a set. The KAM theorem we use relies on the analyticity of the disturbing function, and more precisely in its analyticity widths. To obtain these values in our case, we consider the analytic continuation of the perturbation, and we study it precisely in order to find the singularities (or complex collisions) of this function. For some insight about the theory of analytic continuation for functions of several complex variables, one can consult [50] or [73].
The motivation of this section is the introduction of some notations that can be useful when studying our particular perturbation of the plane planetary three-body problem. To understand better the need of these notations, consider for instance the following variables for $i = 1, 2$:

$$x_i = a_i(\cos(u_i) - e_i),$$
$$y_i = a_i\sqrt{1-e_i^2}\sin(u_i).$$

Those are the coordinates on an ellipse for which $g_i = 0$. We will be later interested in the solution of an equation of the type:

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 = 0.$$

The mixing of all the variables $a_i,\ e_i,\ u_i$ in this equation makes it difficult to solve. This difficulty comes from the fact that we lost the initial structure of the ellipse. Instead, to keep this structure, we could have defined $z_i = x_i + \imath y_i$. The previous equation then becomes:

$$(z_2 - z_1)(\bar{z}_2 - \bar{z}_1) = 0.$$

Hence, a straightforward necessary condition to be a solution of the last is $|z_1| = |z_2|$. In the elliptic variables, it corresponds to the simple condition:

$$a_1(1 - e_1\cos(u_1)) = a_2(1 - e_2\cos(u_2))$$

Now, if the variables $a_i,\ e_i,\ u_i$ are in a complex neighborhood of their initial values, this necessary condition holds. This simplification makes it important, when looking for singularities, to keep the complex structure of the initial equation to solve, so as to simplify greatly the computation.

1. Analytic continuation of a real analytic function: Let f be a real analytic function defined on some domain $D \subset \mathbb{R}$.

$$\begin{aligned} f : D &\to \mathbb{R} \\ x &\mapsto f(x) = \sum_{n \geq 0} a_n x^n \end{aligned}$$

Let $D_{\mathbb{C}}$ be a complex neighborhood of the set D. The analytic continuation of f over the set $D_{\mathbb{C}}$ is simply given by:

$$\begin{aligned} \tilde{f} : D_{\mathbb{C}} &\to \mathbb{C} \\ \tilde{x} &\mapsto f(\tilde{x}) = \sum_{n \geq 0} a_n \tilde{x}^n \end{aligned}$$

Observe that we used the sign ”˜” to indicate that we are looking at a complex variable, and the analytic continuation of f. Moreover, we will decompose $\tilde{x}$ in the following way:

$$\tilde{x} = x + \imath x',$$

where $(x, x') \in \mathbb{R}^2$. To keep track of our variables, whether they are complex or real, we will use this notation throughout the chapter.

2. Analytic continuation of some simple functions: In the example that motivated this section, we used the following analytic function of two variables:

$$\begin{aligned} g : D \subset \mathbb{R}^2 &\to \mathbb{C} \\ (x, y) &\mapsto x + \imath y = z \end{aligned}$$

Define the analytic continuation of this function:

$$\begin{aligned} \tilde{g} : D_{\mathbb{C}} \subset \mathbb{C}^2 &\to \mathbb{C} \\ (\tilde{x}, \tilde{y}) &\mapsto \tilde{x} + \imath \tilde{y} \end{aligned}$$

First, let us call $\tilde{z} = \tilde{g}(\tilde{x}, \tilde{y})$. Using the previous notations, we have:

$$\tilde{z} = \tilde{x} + \imath \tilde{y} = x - y' + \imath(x' + y).$$

Observe that we can define as well $\tilde{z} = z + \imath z'$, where $(z, z') \in \mathbb{R}^2$ with $z = x - y'$ and $z' = x' + y$. Define now $h(x, y) = x - \imath y = \bar{z}$. Its analytic continuation can be written as:

$$\begin{aligned} \tilde{h} : D_{\mathbb{C}} \subset \mathbb{C}^2 &\to \mathbb{C} \\ (\tilde{x}, \tilde{y}) &\mapsto \tilde{x} - \imath \tilde{y} \end{aligned}$$

Writing $\tilde{\bar{z}} = \tilde{h}(\tilde{x}, \tilde{y})$, we have

$$\tilde{\bar{z}} = x + y' + \imath(x' - y).$$

An important remark is that $\tilde{\bar{z}} \neq \bar{\tilde{z}}$. Indeed:

$$\tilde{\bar{z}} = x + y' + \imath(x' - y) \neq x - y' - \imath(x' + y) = \bar{\tilde{z}}.$$

These definitions of the variables $\tilde{z}$ and $\tilde{\bar{z}}$ are the one that motivated this discussion. We will use them as well throughout the chapter, when it simplifies the computations. In the next paragraphs, we give some formulas using those, which will be useful later on.

3. Real part and imaginary part: For a real x and a real y, we defined $z = x + \imath y$. Hence,

$$x = \Re(z) = \frac{z + \bar{z}}{2}, \qquad y = \Im(z) = \frac{z - \bar{z}}{2\imath}.$$

Considering now a complex x and y, with our previous notations we have:

$$\begin{aligned} \tilde{x} &= x + \imath x' = \widetilde{\Re(z)} = \frac{\tilde{z} + \tilde{\bar{z}}}{2}, \\ \tilde{y} &= y + \imath y' = \widetilde{\Im(z)} = \frac{\tilde{z} - \tilde{\bar{z}}}{2\imath}. \end{aligned}$$

4. Analytic continuation of a function of two variables: Let f be an analytic function depending on two variables:

$$\begin{aligned} f : D \subset \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) \quad &\mapsto \sum_{m,n=0}^{\infty} a_{m,n} x^m y^n \end{aligned}$$

where $a_{m,n} \in \mathbb{R}$. The analytic continuation of f is defined on some domain $D_{\mathbb{C}}$, by:

$$\begin{aligned} \tilde{f} : D_{\mathbb{C}} \subset \mathbb{C}^2 &\to \mathbb{C} \\ (\tilde{x}, \tilde{y}) \quad &\mapsto \sum_{m,n=0}^{\infty} a_{m,n} \tilde{x}^m \tilde{y}^n \end{aligned}$$

Since the variables $\tilde{x}$ and $\tilde{y}$ can be expressed using $\tilde{z}$ and $\tilde{\bar{z}}$, there exists a domain $D'_{\mathbb{C}}$ and a sequence $(b_{m,n})_{m,n \geq 0}$ with $b_{m,n} \in \mathbb{C}$, such that we can write:

$$\begin{aligned} \tilde{g} : D'_{\mathbb{C}} \quad &\to \mathbb{C} \\ (\tilde{z}, \tilde{\bar{z}}) &\mapsto \sum_{m,n=0}^{\infty} b_{m,n} \tilde{z}^m \tilde{\bar{z}}^n \end{aligned}$$

and that $\tilde{f}(\tilde{x}, \tilde{y}) = \tilde{g}(\tilde{x} + \imath \tilde{y}, \tilde{x} - \imath \tilde{y})$ on $D_{\mathbb{C}}$.
This correspondence will be of great importance in our work.

5. Cosine of the oriented angle between two vectors: In the formula of the perturbation of the three-body problem, the Legendre polynomial composed with the cosine of the angle S appears, S being the oriented angle between the vectors Q_1 and Q_2. Let us compute the analytic continuation of this function, using our previous work. Let

$$\mathbf{u}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

be two vectors of $\mathbb{R}^2$, with $\|\mathbf{u}_i\| \neq 0$, and $S = \widehat{(\mathbf{u}_1, \mathbf{u}_2)}$. We have:

$$\cos(S) = \frac{x_1 x_2 + y_1 y_2}{\|\mathbf{u}_1\| \|\mathbf{u}_2\|}.$$

Working with $z_i = x_i + \imath y_i$ for $i = 1, 2$, we get

$$\begin{aligned} \cos(S) &= \frac{1}{2} \frac{z_1 \bar{z}_2 + \bar{z}_1 z_2}{\sqrt{z_1 \bar{z}_1 z_2 \bar{z}_2}} \\ &= \frac{1}{2} \left(\sqrt{\frac{z_1 \bar{z}_2}{\bar{z}_1 z_2}} + \sqrt{\frac{\bar{z}_1 z_2}{z_1 \bar{z}_2}} \right). \end{aligned} \tag{1.9}$$

One needs to take care of the fact that none of the fractions $\tilde{z}_i/\tilde{\bar{z}}_i$ goes to zero. Though, to be in this case, one has to consider a large neighborhood of the real domain we will be looking at. The analytic continuation is hence straightforward, replacing z by $\tilde{z}$ and $\bar{z}$ by $\tilde{\bar{z}}$.

6. Euclidean norm of the difference between two vectors: To complete our study, let us take a look at the Euclidean norm, and more precisely at its analytic continuation on a domain not containing zero.
Following our notations, we have for $(\tilde{x}, \tilde{y}) \neq (0,0)$:

$$\widetilde{\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|} \equiv \sqrt{\widetilde{x^2+y^2}} = \sqrt{\tilde{z}\tilde{\bar{z}}}.$$

Several remarks are necessary. Observe that we are using the analytic continuation of the square root function. We can indeed consider a determination of the root of a complex number under some conditions on the initial set. Let D be the initial set of real values we are considering such that $0 \notin D$. Consider the function $d(x,y) = x^2+y^2$. Assume that $D_{\mathbb{C}}$ is a complex neighborhood of D such that $0 \notin D_{\mathbb{C}}$, and consider the analytic continuation of d on this set. If $D_{\mathbb{C}}$ is close enough to D, then $\tilde{d}(D_{\mathbb{C}})$ will be close to $d(D)$, and hence close to the real axis. Hence, the square root will be well-defined on the set $\tilde{d}(D_{\mathbb{C}})$.
Consider now the Euclidean norm of the difference between two vectors, which appears when looking at the singularities of the perturbing function. Let us consider for instance two points $M_1 = (x_1, y_1)$ and $M_2 = (x_2, y_2)$ in the plane case. The vector between these two points is $\mathbf{u} = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}$. We are interested in the case where the norm of $\mathbf{u}$ goes to zero. This question is closely related to the singularities of the perturbation, since we have terms of the form $1/(Q_2 - cQ_1)$. In the real case, we can let $z = x + \imath y$, and then the distance is the following:

$$d = \sqrt{z\bar{z}}.$$

Letting $z_1 = x_1 + \imath y_1$, $z_2 = x_2 + \imath y_2$, and $z = z_2 - z_1$, we get

$$d = \sqrt{z_2 - z_1}\sqrt{\bar{z}_2 - \bar{z}_1}.$$

Using the analytic continuation, we see that we have a singularity in the case

$$\tilde{z_1} = \tilde{z_2} \quad \text{or} \quad \tilde{\bar{z}}_1 = \tilde{\bar{z}}_2.$$

This remark will make it easier to find the singularities related to the complex collisions between the two planets. Observe that in this case, we have two singularities occurring at the same time. One singularity comes from the value of a denominator of the disturbing function going to zero, and the other coming from the fact that the square root function is not analytic at zero.

1.2 Complex Kepler's equation

This section is devoted to the study of the complex Kepler equation. This equation arises when defining the mean anomaly as a function of the eccentric anomaly. It depends on a parameter $0 \leq e < 1$, that represents the eccentricity in celestial mechanics. In the study of the analyticity width of the disturbing function, not only the angles can be taken in a complex set, but also the eccentricity. We are mainly interested in two problems: first, we wish to determine the singular points of a local diffeomorphism between the two previously mentioned angles; secondly, we want to find a precise set on which the Kepler equation induces a diffeomorphism.

1.2.1 Complex elliptic coordinates

Preliminaries: The real case

Let us start with a short study of the real case. Recall that u corresponds to the eccentric anomaly. We can release the constraint on the eccentricity to be positive when focusing on the Kepler equation, and consider $e \in \mathbb{R}$. We call $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, and define the mean anomaly using the function

$$\begin{aligned} f : \mathbb{T} &\to \mathbb{T} \\ u &\mapsto u - e \sin u \end{aligned}$$

It is easy in the real case to see if f is a diffeomorphism, its derivative being, for $u \in \mathbb{T}$:

$$f'(u) = 1 - e \cos u.$$

In the case $|e| = 1$, f is not a diffeomorphism on the set $\mathbb{T}$, since for $e = 1$ (respectively $e = -1$), at the angle $u = 0$ (resp. $u = \pi$), $f'(u) = 0$. In the case $|e| > 1$, there exists two points where the derivative goes to zero. Hence, a necessary and sufficient condition for f to be a diffeomorphism on the set $\mathbb{T}$ is $0 \leq |e| < 1$.

After the analytic continuation

The first difficulties in the study of the singular points of the Kepler equation arise when we consider the complex case. We will start, for the sake of simplicity, with the study of the elliptic coordinates.
Consider the complex eccentricity $\tilde{e} = e + \imath e'$, with $e, e' \in \mathbb{R}$. To simplify the discussions, we will as well consider that $e \geq 0$. Kepler's equation is completely symmetric in $\tilde{e}$, and therefore the results can be extended easily to the negative case. The complex angles will take their values in the set $\mathbb{T}_{\mathbb{C}} = \mathbb{T} \times \mathbb{R}$ (we identified $\mathbb{C}$ and $\mathbb{R}^2$). Consider the function:

$$\begin{aligned} \tilde{f} :\ \mathbb{T}_{\mathbb{C}} &\to \mathbb{T}_{\mathbb{C}} \\ \tilde{u}\ &\mapsto\ \tilde{l} = \ \tilde{u} - \tilde{e} \sin \tilde{u} \end{aligned}$$

We are interested in determining if $\tilde{f}$ is an analytic diffeomorphism (at least locally), in every points of a set $\mathbb{T} \times (-u'_{max}, u'_{max})$. Hence, we want to determine the singular points of $\tilde{f}$. Consider the function $(\Re \tilde{f}, \Im \tilde{f}) : (\Re \tilde{u}, \Im \tilde{u}) \mapsto (\Re \tilde{l}, \Im \tilde{l})$, it has the same singular points as $\tilde{f}$, that is why we will shift from one point of view to another without self-restraint. This function is defined by the formulas

$$\left\{ \begin{array}{l} \Re \tilde{f} = l \ = u - e \sin(u) \cosh(u') + e' \cos(u) \sinh(u') \\ \Im \tilde{f} = l' = u' - e' \sin(u) \cosh(u') - e \cos(u) \sinh(u') \end{array} \right. \tag{1.10}$$

The derivative of $\tilde{f}$ is $\tilde{f}'(\tilde{u}) = 1 - \tilde{e} \cos \tilde{u}$. Hence, there is a singular point if and only if

$$\left\{ \begin{array}{l} \Re(\tilde{e} \cos \tilde{u}) = 1 \\ \Im(\tilde{e} \cos \tilde{u}) = 0 \end{array} \right. .$$

In the real variables, it gives:

$$\left\{ \begin{array}{l} e \sin(u) \sinh(u') = e' \cos(u) \cosh(u') \\ e \cos(u) \cosh(u') + e' \sin(u) \sinh(u') = 1 \end{array} \right. . \tag{1.11}$$

Consider first the case $\tilde{e} = e + \imath e' = 0$. Then it is straightforward to see that there exists no singular point, since $\tilde{f} = Id$.

The second case we can consider is the case $e' = 0$, in other words for a real eccentricity. There exists three different type of singular points. For $0 < |e| < 1$, for $e > 0$ (*respectively* $e < 0$) there exists two singular points in $(u, u') = (0, \pm \operatorname{arccosh}(1/e))$ (*resp.* $(\pi, \pm \operatorname{arccosh}(1/e))$. When $e = 1$, there is one singular point at $(0, 0)$, and when $e = -1$ at $(\pi, 0)$. Finally, when $|e| > 1$, there exists two singularities for $(u, u') = (\arccos(1/e), 0)$.
Now let us assume $e' \neq 0$. For the sake of simplicity, we consider that $e \geq 0$ (the case $e < 0$ can be obtained by symmetry consideration). To find the singular points, we deduce the angle u in the first equation, and solve the second in u'.
Assume $u \in [0, \pi/2]$, $e' > 0$, $u' > 0$ (again, by symmetry of the equations, the study of this case is enough to compute the other cases), the first equation gives

$$e^2(1 - \cos^2(u)) \sinh^2(u') = e'^{\,2} \cos^2(u) \cosh^2(u').$$

Call $\mu = \sqrt{e'^{\,2} \cosh^2(u') + e^2 \sinh^2(u')}$ (since $e' \neq 0$, we have $\mu > 0$):

$$\begin{cases} \cos(u) = \dfrac{e \sinh(u')}{\mu} \\ \sin(u) = \dfrac{e' \cosh(u')}{\mu} \end{cases} \tag{1.12}$$

Observe that in every case, *i.e.* for $e' > 0, u' > 0$ or $e' < 0, u' > 0$ or $e' > 0, u' < 0$ or $e' < 0, u' < 0$, there exists two solutions (obtained by adding π to the first solution) of the first equation in $\mathbb{T}$.
The second equation can now be written

$$e^2 \cosh(u') \sinh(u') + e'^2 \cosh(u') \sinh(u') = \mu.$$

Squaring this equation, and using hyperbolic trigonometry identities, we obtain:

$$(e^2 + e'^2)^2 \cosh^4(u') - ((e^2 + e'^2)^2 + (e^2 + e'^2)) \cosh^2(u') + e^2 = 0$$

Let $m = e^2 + e'^2 > 0$, there is a unique possible value for u':

$$u'_{max} = \operatorname{arccosh}\left(\sqrt{\frac{1}{2}\left(1 + \frac{1}{m}\left(1 + \sqrt{(m+1)^2 - 4e^2}\right)\right)}\right). \tag{1.13}$$

By symmetry arguments, we obtain two curves of singular points. We represented those in figure 1.4 and 1.5, for small eccentricities, and in figure 1.6 and 1.7 for large eccentricities. Observe that if we release the constraint e positive, and authorize it to be negative, we obtain the same figure as figure 1.4 and 1.6, though shifted by a factor π, which represents the real singular point for the case $e = -1$. The case $|e| = 1$ corresponds to a change of mode, where there is the existence of a singular point on the real axis when $e' = 0$.
To summarize, we have the following singular point:

- $m = 0$: There exists no singular points, for all $\tilde{u} \in \mathbb{T}_{\mathbb{C}}$
- $\tilde{e} \in \mathbb{R} \setminus \{0\}$: for real eccentricities $|e| \notin \{0, 1\}$, there are 2 singular points depending on the value of e. When $|e| = 1$, only one singular points exists.
- $e' \neq 0$: For each possible sign of the variables u', e, e' there exists two singular points.

When fixing the value of the eccentricity, the complex Kepler equation defines an analytic local diffeomorphism at every point that is not singular, where the latter are given by the formulas above.

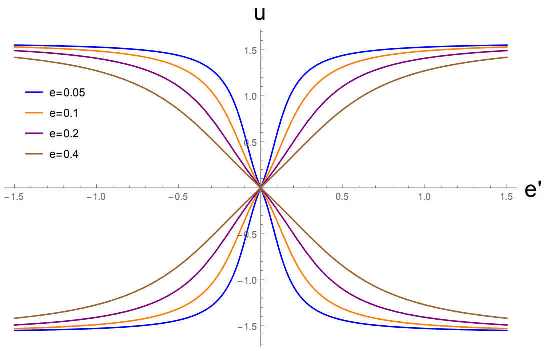

Figure 1.4: Real part u of the eccentric anomaly of singular points as a function of the imaginary part e' of the eccentricity for a fixed small real part e (respectively 0.05, 0.1, 0.2, 0.4). Increasing values of e are represented respectively on the blue, orange, purple, and brown curve.

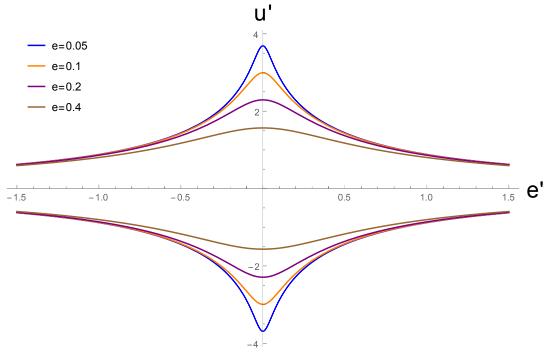

Figure 1.5: Imaginary part u' of the eccentric anomaly of singularities as a function of the imaginary part e' of the eccentricity for a fixed small real part e (respectively 0.05, 0.1, 0.2, 0.4). Increasing values of e are represented respectively on the blue, orange, purple, and brown curve.

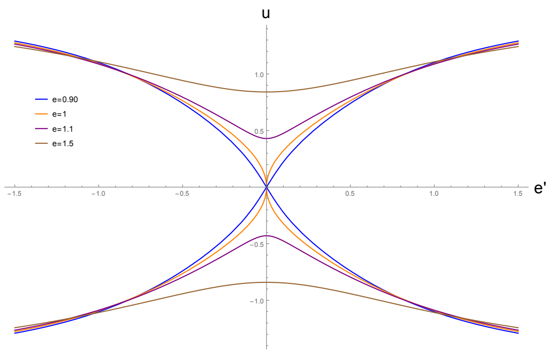

Figure 1.6: Real part u of the eccentric anomaly of singular points as a function of the imaginary part e' of the eccentricity for a fixed large eccentricity e (respectively 0.9, 1, 1.1, 1.5). Increasing values of e are represented respectively on the blue, orange, purple, and brown curve. The orange line divides the two different modes.

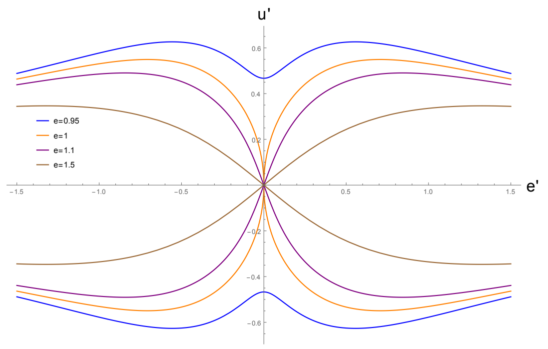

Figure 1.7: Imaginary part u' of the eccentric anomaly of singularities as a function of the imaginary part e' of the eccentricity for a large eccentricity e (respectively 0.9, 1, 1.1, 1.5). Increasing values of e are represented respectively on the blue, orange, purple, and brown curve.

1.2.2 Singular points in Poincaré coordinates

When switching to Poincaré coordinates, Kepler's equation loses its symmetry. Besides, the relation between the Cartesian coordinates $F = (\xi, \mu)$ and the eccentricity is quite complicated, and makes the search of singular points more difficult. The elliptic case involved only two parameters, whereas in Poincaré case there are six.
In the continuity of our work, we fix the value of $\tilde{\Lambda}$ and $\tilde{F}$, which fixes the eccentricity vector $\tilde{E}$. Using this vector, we give the expression of the singular points in the eccentric longitude $\tilde{w} = (w, w')$.
This section is divided in three parts. First, we start with a study of the real case to get familiar with the transformation, then we compute the singular points using the variable $\tilde{E}$, and finally, we define a domain in which Kepler's equation induces an analytic local diffeomorphism at each points in Poincaré coordinates.

Real Case

In terms of the eccentric and mean longitude $w = u + g$ and $\lambda = l + g$, Kepler's equation can be rewritten as:

$$\lambda = w - e\sin(w - g). \tag{1.14}$$

As mentioned before, we use the temporary variable $E = e\exp(\imath g)$.

$$\begin{aligned}\lambda &= w - \frac{e}{2\imath}\left(\exp(\imath(w-g)) - \exp(-\imath(w-g))\right)\\ &= w - \frac{\bar{E}}{2\imath}\exp(\imath w) + \frac{E}{2\imath}\exp(-\imath w)\\ &= h(w).\end{aligned}$$

In order for h to define a local diffeomorphism, we want its derivative with respect to w to be non-null. The singular points verify

$$1 - \frac{\bar{E}}{2}\exp(\imath w) - \frac{E}{2}\exp(-\imath w) = 0. \tag{1.15}$$

We preferred the exponential notation to simplify our upcoming work. Decomposing the variable E into $E = E_1 + \imath E_2$, the previous equation becomes

$$E_1 \cos w + E_2 \sin w = 1.$$

The first singular points arising in coordinates (E_1, E_2), in other words the minimal modulus of E such that we have a singular point, corresponds to the case $|E| = 1$, the singular points is then in $w = \arg(E)$. Considering the variable F, recall that its relation with E is given by the formula

$$F = \sqrt{\frac{2\Lambda}{1 + (1 - E\bar{E})^{\frac{1}{2}}}}\bar{E}.$$

In our case, we have $E\bar{E} = 1$. Hence, the first singular points occurring in term of the modulus of F are on a circle of radius $|F| = \sqrt{2\Lambda}$. The angle of the singular point is then $w = -\arg(F)$. In contrast with the elliptic case, the singulars point are not located in 0, it can take every values in $\mathbb{T}$.Indeed, the angle of the periapsis is not originally related to Kepler's equation, hence, when studying the mean longitude, the singular points describe the whole circle with the value of g. Besides, we can reach continuously a negative value $e = -1$ when changing g from 0 to π.

Singular points for a fixed $\tilde{E}$

Considering the analytic continuation of the previous function defined in the real case, its singular points satisfy

$$\frac{\tilde{\tilde{E}}}{2}\exp(\imath w)\exp(-w') + \frac{\tilde{E}}{2}\exp(-\imath w)\exp(w') = 1. \quad (1.16)$$

Fixing $\tilde{E}$ and $\tilde{\tilde{E}}$, we are looking at singular points in the plane (w, w'). We can distinguish different cases.

- Case $\tilde{E} = \tilde{\tilde{E}} = 0$: it is clear that equation (1.16) cannot be verified. There is no singular points as soon as $\tilde{E}$ and $\tilde{\tilde{E}}$ are null, this corresponds to the case of a zero-eccentricity.
- Case $\tilde{E} = 0$, $\tilde{\tilde{E}} \neq 0$: call $\tilde{\tilde{E}}/2 = s\exp(\imath\sigma)$ with $s > 0$ and $\sigma \in \mathbb{T}$, equation (1.16) becomes:

  $$s\exp(-w')\cos(w+\sigma) + \imath s\exp(-w')\sin(w+\sigma) = 1.$$

  We obtain $w = -\sigma + k\pi$ for $k \in \mathbb{Z}$, and $w' = \log(s) = \log\frac{|\tilde{\tilde{E}}|}{2}$. There exists two singularities $w \in [0, 2\pi]$, at the same distance $w' = \log(s)$ from the set of real angles $\mathbb{T} \times \{0\}$.
- Case $\tilde{\tilde{E}} = 0$, $\tilde{E} \neq 0$: as well, there exists singular points for $w = -\arg\tilde{E} + k\pi$, $k \in \mathbb{Z}$, and $w' = \log\frac{|\tilde{E}|}{2}$.
- Case $\tilde{E} \neq 0$, $\tilde{\tilde{E}} \neq 0$: in the general case, let $\tilde{E}/2 = r\exp(\imath\theta)$ and $\tilde{\tilde{E}}/2 = s\exp(\imath\sigma)$.

  Lemma 1.1. *The singular points of the complex Kepler equation* (1.16) *are given by the formulas:*

  $$\begin{cases} w_1' = \log(2s) - \frac{1}{2}\log\left((1+\sqrt{\Delta}) + \sqrt{(\sqrt{\Delta}+1)^2 - 16(rs)^2}\right) \\ w_2' = -\log(2r) + \frac{1}{2}\log\left((1+\sqrt{\Delta}) + \sqrt{(\sqrt{\Delta}+1)^2 - 16(rs)^2}\right) \end{cases}, \quad (1.17)$$

  where $\Delta = (1-4rs)^2 + 8rs(1-\cos\gamma)$ *and* $\gamma = \theta + \sigma$.

  Proof. Equation (1.16) can be rewritten:

  $$s\exp(-w')\exp(\imath(\sigma+w)) + r\exp(w')\exp(\imath(\theta-w)) = 1.$$

  Call $x = \sigma + w$ and $\gamma = \theta + \sigma$, $a = r\exp(w')$ and $b = s\exp(-w')$. The previous equation becomes:

  $$b\exp(\imath x) + a\exp(\imath(\gamma - x)) = 1. \quad (1.18)$$

  The equation verified by the singular points of the function induced by the complex Kepler equation are:

  $$\begin{cases} (b + a\cos\gamma)\cos x + a\sin\gamma\sin x = 1 \\ (b - a\cos\gamma)\sin x + a\sin\gamma\cos x = 0 \end{cases} \quad (1.19)$$

  Case $\gamma = \pi$***:*** since $a, b > 0$, the second equation implies $\sin x = 0$, and hence the first one becomes: $(b-a)\cos(v) = 1$. This implies that there exists two singular points, one for $x = 0$, the other one for $x = \pi$. In $x = 0$, we obtain $w' = \log(-1+\sqrt{1+4rs}) - \log(2r)$ and in $x = \pi$, $w' = \log(1+\sqrt{1+4rs}) - \log(2r)$.
  Case $\gamma = 0$***:*** the second equation of (1.19) implies $(b-a)\sin x = 0$, and the first one $(b+a)\cos x = 1$. Notice that since $a + b > 0$, there exist no singular point as soon as $\cos x \leq 0$. We can again divide the study in two cases. In $x = 0$, the equation to solve

is $b + a = 1$. There exists at least a singular point if and only if $rs \leq 1/4$, and their coordinates are given by $w' = \log(1 \pm \sqrt{1 - 4rs}) - \log(2r)$.
General case **($\gamma \neq 0, \gamma \neq \pi$):** in order to study the general case, we do as in the elliptic case. The second equation of the system (1.19) gives:

$$\begin{cases} \cos x = \pm \dfrac{a \cos\gamma - b}{\mu} \\ \\ \sin x = \pm \dfrac{a \sin\gamma}{\mu} \end{cases}$$

where $\mu = \sqrt{a^2 + b^2 - 2ab\cos\gamma} \neq 0$, and $\cos x$ and $\sin x$ have same signs. Injecting in the first equation, we get

$$(a^2 - b^2)^2 = a^2 + b^2 - 2ab\cos\gamma.$$

Changing to our coordinate w', we have:

$$r^4\exp(4w') - r^2\exp(2w') - 2r^2s^2 + 2rs\cos\gamma - s^2\exp(-2w') + s^4\exp(-4w') = 0 \quad (1.20)$$

This equation is of order 4 in $\exp(2w')$. It turns out that it has either a repeated root, either 2 real roots. Indeed, letting $x' = 2(w' - \log(\sqrt{s/r})) = 2w' - \log(s/r)$, the equation in x' is:

$$\begin{aligned} &(rs)^2(\exp(2x') + \exp(-2x') - 2) - rs(\exp(x') + \exp(-x') - 2\cos\gamma) = 0 \\ \Leftrightarrow\ &2(rs)^2(\cosh(2x') - 1) - 2rs(\cosh(x') - \cos\gamma) = 0. \end{aligned}$$

Putting this equation under the form of a polynomial of order 2 in $\cosh x'$:

$$2rs\cosh^2(x') - \cosh(x') - (2rs - \cos\gamma) = 0. \quad (1.21)$$

This polynomial is even, hence, if $x' > 0$ is a solution, then $-x'$ is also one. By making this equation symmetric using the variable x', we can obtain explicit solutions. Let $V = \cosh(x')$. the discriminant of this polynomial is $\Delta = (1 - 4rs)^2 + 8rs(1 - \cos\gamma)$, and is always positive. Observe how the other cases $\cos\gamma = 1$ and $4rs = 1$ arise naturally. We have two solutions in the case $\cos\gamma \neq 1$:

$$\begin{cases} V_1 = \dfrac{1}{4rs} + \dfrac{1}{4rs}\sqrt{(1 - 4rs)^2 + 8rs(1 - \cos\gamma)} = \dfrac{1}{4rs}\left(1 + \sqrt{\Delta}\right) \\ \\ V_2 = \dfrac{1}{4rs} - \dfrac{1}{4rs}\sqrt{(1 - 4rs)^2 + 8rs(1 - \cos\gamma)} = \dfrac{1}{4rs}\left(1 - \sqrt{\Delta}\right) \end{cases}$$

Yet $V = \cosh x'$, therefore only the solutions $V_i \geq 1$ must be considered, which is V_1. The solutions of equation (1.21) are

$$x' = \pm \operatorname{arccosh} V_1 = \pm \log\left(V_1 + \sqrt{V_1^2 - 1}\right).$$

Observe as well that $V_1^2 - 1 = V_1/(2rs) - \cos\gamma/(2rs)$, hence the two symmetrical solutions are

$$x' = \pm\left(\log(4rs) - \log\left((1 + \sqrt{\Delta}) + \sqrt{(\sqrt{\Delta} + 1)^2 - 16(rs)^2}\right)\right). \quad (1.22)$$

Using the relation between x' and w', we recover the expressions of the lemma. □

With the formulas of the lemmas, we get interested in two different matters, that are interesting when trying to find a domain on which the analyticity widths of the diffeomorphism induced by Kepler's equation is non-null. First, for a fixed r, s and γ, we want to determine the closest singular point to the real axis $\mathbb{T} \times \{0\}$. Secondly, we want to determine if it is possible that one of the singular points is on the real axis.
In the case $\gamma = \pi$, when $r < s$, the closest singular point is in $w' = \log(-1 + \sqrt{1 + 4rs}) - \log(2r)$, and in the opposite case $r > s$, it is in $w' = \log(1 + \sqrt{1 + 4rs}) - \log(2r)$. There exists two lines for which $w' = 0$, the line of equation $s = r + 1$ and the one of equation $s = r - 1$.
In the case $\gamma = 0$, if $s > r$, then the the first singular point occurs for $w' = \log(1 + \sqrt{1 - 4rs}) - \log(2r)$, in the opposite case for $w' = \log(1 - \sqrt{1 - 4rs}) - \log(2r)$, and they are at the same distance when $r = s$. Moreover, if $s + r = 1$, then the width w' is zero. Besides, after passing through the line $s + r = 1$ (hence for $s + r > 1$ but $rs \leq 1/4$), both solutions have same sign, they are positive if $s > r$ and negative if $r < s$. When $x \neq 0$, we get the equation $b - a = 0$, and therefore $2a \cos v = 1$. There exists solutions only if $\sqrt{rs} \geq 1/2$, and in this case, $w' = 1/2 \log(s/r)$ and $x = \pm \arccos(1/2\sqrt{rs})$. The width w' is null for $r = s$, and we can see that it is continuous at $\sqrt{rs} = 1/2$.
In the general case, by symmetry, the closest solution to the origin depends on the ratio s/r. Indeed, the two solutions are symmetrical about the value $\log(\sqrt{s/r})$. They are both of the same sign if $\log(\sqrt{s/r})$ is greater than the norm of the solution x' of the equation (1.21), which corresponds to passing through a point where the width w' is zero. Hence, as in the case $\gamma = 0$, the "line" of width zero indicates a shift from two solutions of different signs to a case of identical sign. Moreover, the only case of repeated root occurs in the cases $rs = 1/4$ or $\cos\gamma = 1$, and after passing through this line, *i.e.* for $r = s$ and $rs > 1/4$, the width w' stays null.

Null-width line ($w' = 0$): To finish our study, we are interested in the singular points on the real axis for $\tilde{w}$, *i.e.* for which $w' = 0$ (the other variables being complex). We have already seen that in the case $\gamma = 0$, there exists singular points such that $w' = 0$ on the line $r + s = 1$, and the half-line $r = s$ for $rs \geq 1/4$. In $\gamma = \pi$, this corresponds to the two lines $s = 1 + r$ et $s = -1 + r$. Finally, in $(r, s) = (0, 1)$ and $(r, s) = (1, 0)$, we also have $w' = 0$.
In the other cases, we ask that:

$$(r^2 - s^2)^2 = r^2 + s^2 - 2rs\cos\gamma. \tag{1.23}$$

For each γ, this equation defines implicitly s as a function of r. Notice that the couples $(0, 1)$ and $(1, 0)$ verify this equation for every γ. Rewrite (1.23) as follows:

$$(r - s)^2 \left((r + s)^2 - 1)\right) = 2rs(1 - \cos\gamma).$$

Since the second term is strictly positive, and that $(r-s)^2$ is positive (if $\cos\gamma \neq 1$ there are no solution such that $r = s$), a necessary condition is $r+s > 1$. Hence, when $\cos\gamma = 1$, the singular points on the real axis are the closest in norm l_1 for (r, s). In norm l_2 (the Euclidean one), the closest singular point occurs at the coordinates $(r, s, \gamma, w') = (1/2, 1/2, 0, 0)$.
To be able to study more precisely the solutions of equation (1.23) when $\cos\gamma \neq 1$, let $z = r + s$, and $y = r - s$. The previous equation can then be written:

$$z^2 y^2 = \frac{z^2}{2}(1 + \cos\gamma) + \frac{y^2}{2}(1 - \cos\gamma).$$

We can describe y as a function of z, there exists two branches:

$$y_\pm(z) = \pm\sqrt{\frac{z^2(1 + \cos(\gamma))}{2z^2 + \cos\gamma - 1}} \tag{1.24}$$

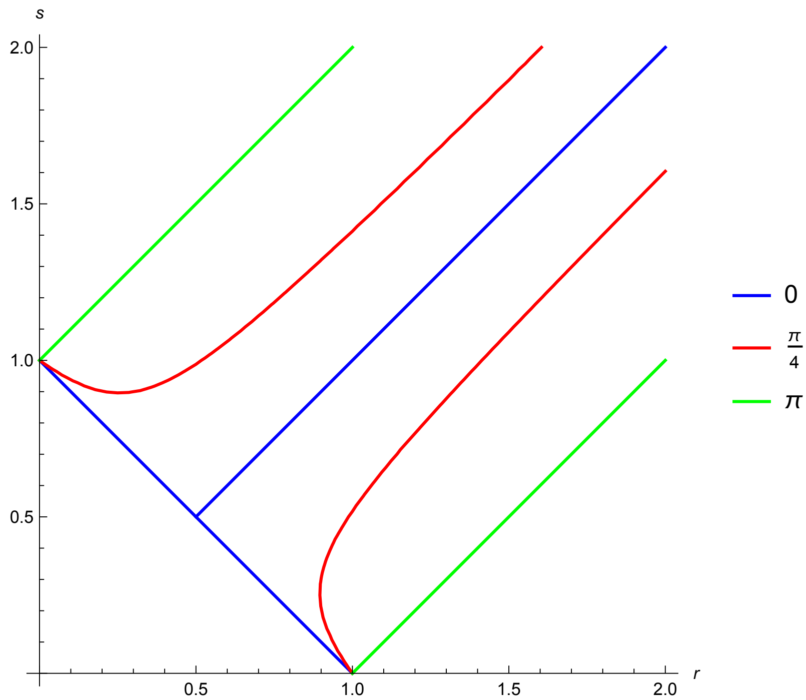

Figure 1.8: Singular points on the real axis ($w' = 0$) on the plane (r, s) for different values of γ: in blue, $\gamma = 0$, in red $\gamma = \frac{\pi}{4}$, in green $\gamma = \pi$.

We can determine the domain of definition of z. We mentioned before the fact that we restrict the discussion to the case $r + s > 1$, *i.e.* $z > 1$. Indeed, we could consider values smaller than 1, but they would give negative values for either r or s, which does not interest us. The function y is then well-defined, since the denominator in the square root is strictly positive under the condition $z > 1$. Let us study the limit of y when z goes to infinity:

$$\lim_{z \to \infty} y_{pm}(z) = \pm \sqrt{\frac{1 + \cos \gamma}{2}}$$

Hence, in the plane (r, s), the solution $s(r)$ defined implicitly by equation (1.23) have two branches, which start in $(r, s) = (1, 0)$ (*respectively* $(0, 1)$) and which admits, when r goes to infinity, an asymptote $\mathcal{D}_- : s = r - \sqrt{\frac{1+\cos\gamma}{2}}$ (*respectively* $\mathcal{D}_+ : s = r + \sqrt{\frac{1+\cos\gamma}{2}}$). The figure 1.8 represents the cases $\gamma = 0$, $\gamma = \pi$, and a transitional case (here $\gamma = \pi/4$), it helps identifying the difference between those cases, and understanding the evolution from the curve of singular points for a null-width with the value γ. When γ increases from 0 to π, we can see on the figure the detachment of the curves from the case $\gamma = 0$ to finally reach the two curves of the case $\gamma = \pi$. We can therefore identify three main zones (or connected components): the first one is the one containing the origin $r = s = 0$, which corresponds to two singular points in w' of different signs, the case $r = s$ corresponding to the line where the two singularities are at same distance of the origin; the second one is the upper zone and corresponds to the existence of singular points with w'_1 and w'_2 positive; finally the lower zone for which the singular points have negative values for w'_i.

Singular points using the Cartesian coordinates $\tilde{F}$

A general formula of the singular points of the complex Kepler equation in Poincaré coordinates depends solely on the complex domains of our variables. Our choice of considering polydiscs will prove to be convenient later, though these domains do not transfer in an intuitive way when

changing the variable $\tilde{E}$ into $\tilde{F}$. Given a point $\tilde{F}$, it is possible to compute the value of $\tilde{E}$ and hence to deduce the width w' at this exact point. Though, when considering a polydisc, it is not straightforward to understand well the domain in terms of the variable $\tilde{E}$, and to deduce an exact maximum width. We will therefore only look for a sufficient condition to avoid the singular points of the change of variables, trying to be as optimal as possible.
First, consider the coordinate $\tilde{\Lambda} \in B(\Lambda_0, r)$ with $r < \Lambda_0$. The set $\mathbb{T}$ where w takes its value is extended to a set for the $\tilde{w}$ that is $\mathbb{T}\times] - w'_{max}, w'_{max}[$. Considering the variable $F = \mu + \imath\xi$, the real domain we choose is a disc around the origin (the point of null eccentricity): define $0 < \rho \leq \sqrt{2(\Lambda_0 - r)}$, then $F \in B(0, \rho)$, or in other words $F = r_0 \exp(\imath\theta)$ with $0 \leq r_0 < \rho$ and $\theta \in \mathbb{T}$. Considering the complex coordinates $\tilde{\xi}$ and $\tilde{\eta}$: define an identical width $\rho' < \sqrt{\Lambda_0 - r} - \rho/\sqrt{2}$, these variables are defined with the help of $0 \leq r_0 < \rho$, $\theta \in \mathbb{T}$, $0 \leq r_1, r_2 < \rho'$, $\theta_1, \theta_2 \in \mathbb{T}$, and we have $(\tilde{\mu}, \tilde{\xi}) = (r_0 \cos\theta + r_1 \exp(\imath\theta_1), r_0 \sin\theta + r_2 \exp(\imath\theta_2)$. Let us call this set:

$$\begin{aligned}\mathcal{D}^w_{\Lambda_0,r,\rho,\rho',w'_{max}} = \{(\tilde{\Lambda}, \tilde{w}, \tilde{\xi}, \tilde{\eta}) \in \mathbb{C} \times \mathbb{T}_{\mathbb{C}} \times \mathbb{C}^2 \quad &: \tilde{\Lambda} \in B(\Lambda_0, r),\ |\Im\tilde{w}| < w'_{max}, \\ &\exists(\xi_0, \eta_0) \in B(0, \rho) \text{ s.t. } \tilde{\xi} \in B(\xi_0, \rho'),\ \tilde{\eta} \in B(\eta_0, \rho')\}. \end{aligned} \tag{1.25}$$

As shown before, for small values of $|\tilde{E}|$ and $|\tilde{\tilde{E}}|$, the width on which we have a local diffeomorphism is strictly positive and hence we can apply a local inversion theorem. We will determine values of ρ and ρ' small enough to be sure that it is still the case.

Lemma 1.2. *Under the assumption*

$$\frac{1}{\sqrt{\Lambda_0 - r}} \sqrt{1 + \frac{3\rho'^2}{2(\Lambda_0 - r)}} (\rho + 2\rho') < 1, \tag{1.26}$$

the value of w'_{max} on which the complex Kepler equation induces a local diffeomorphism at every point of the set $\mathcal{D}^w_{\Lambda_0,r,\rho,\rho',w'_{max}}$ is strictly positive. Besides, the following inequality holds:

$$w'_{max} > \operatorname{arccosh}\left(\sqrt{\frac{\Lambda_0 - r}{1 + \frac{3\rho'^2}{2(\Lambda_0 - r)}}} \frac{1}{(\rho + 2\rho')}\right). \tag{1.27}$$

Proof. Consider the complex Kepler equation (1.16) and the definition of $\tilde{E}$ and $\tilde{\tilde{E}}$ as a function of $\tilde{F}$ and $\tilde{\tilde{F}}$, it gives

$$\frac{1}{2\sqrt{\tilde{\Lambda}}} \sqrt{1 - \frac{\tilde{F}\tilde{\tilde{F}}}{4\tilde{\Lambda}}} \left(\tilde{F} \exp(\imath w) \exp(-w') + \tilde{\tilde{F}} \exp(-\imath w) \exp(w')\right) = 1. \tag{1.28}$$

The sufficient condition will rely on determining an upper bound for each of these terms independently (instead of maximizing them at the same time), and checking that this bound is less than 1. Indeed, it is not immediate to maximize together the term under the square root sign and the terms where w' appears.
Let us start by bounding the first term $\tilde{\Lambda}^{-\frac{1}{2}}$: the norm of $\tilde{\Lambda}$ is simply bounded from below by $\Lambda_0 - r$, hence $|\tilde{\Lambda}^{-\frac{1}{2}}| \leq (\Lambda_0 - r)^{-\frac{1}{2}}$ on the wanted set.
By definition of the set $\mathcal{D}^w_{\Lambda_0,r,\rho,\rho',w'_{max}}$, there exists $r_0, r_1, r_2 \in \mathbb{R}^+ = [0, +\infty)$, $\sigma_0, \sigma_1, \sigma_2 \in \mathbb{T}$ such that $F = \xi + \imath\eta = r_0 \exp(\imath\theta_0)$, $\tilde{\xi} - \xi = r_1 \exp(\imath\theta_1)$ and $\tilde{\eta} = r_2 \exp(\imath\theta_2)$. In the previous equation, considering the second term of the product under the square root sign, since in real coordinates

we have $\bar{F}F = \xi^2 + \eta^2$, we get:

$$\begin{aligned}\left|1 - \frac{\tilde{F}\tilde{\bar{F}}}{4\tilde{\Lambda}}\right| &= \left|1 - \frac{\tilde{\xi}^2 + \tilde{\eta}^2}{4\tilde{\Lambda}}\right| \\ &< 1 + \frac{1}{4(\Lambda_0 - r)} \sup_{r_0,r_1,r_2} \left|-r_0^2 + 2r_0(r_1 + r_2) + r_1^2 + r_2^2\right| \\ &< 1 + \frac{1}{4(\Lambda_0 - r)} \sup_{r_0,r_1,r_2} \left(-(r_0 - r_1 - r_2)^2 + (r_1 + r_2)^2 + r_1^2 + r_2^2\right) \\ &< 1 + \frac{3\rho'^2}{2(\Lambda_0 - r)}.\end{aligned}$$

This bound is independent of ρ, using the fact that $|F\bar{F}| = r_0^2$, and that it has a negative sign in front of it. It therefore depends only on ρ' and r'.
Regarding the last term, observe that the maximal value taken by $|\tilde{\bar{F}}|$ is the same as the one taken by $|\tilde{F}|$ because of the symmetry of the domains. We have

$$\begin{aligned}\left|\left(\tilde{F}\exp(\imath w)\exp(-w') + \tilde{\bar{F}}\exp(-\imath w')\exp(w')\right)\right| &< \sup_{\mathcal{D}^w_{\Lambda_0,r,\rho,\rho',w'_{max}}} |\tilde{F}| \times \sup_{|w'|<w'_{max}} 2\cosh(w') \\ &< 2(\rho + 2\rho')\cosh(w'_{max}).\end{aligned}$$

Hence, a sufficient condition for the initial set to avoid the singular points in our domain is:

$$\frac{1}{\sqrt{\Lambda_0 - r}}\sqrt{1 + \frac{3\rho'^2}{2(\Lambda_0 - r)}}(\rho + 2\rho')\cosh(w'_{max}) < 1.$$

□

1.2.3 Induced diffeomorphism

Injectivity and semi-global inversion

The complex Kepler equation, seen as a complex change of variables is associated to the function:

$$\begin{aligned} f : \mathbb{T}_{\mathbb{C}} &\to \mathbb{C} \\ \tilde{w} &\mapsto \tilde{w} - \frac{\tilde{\bar{E}}}{2\imath}\exp(\imath\tilde{w}) + \frac{\tilde{E}}{2\imath}\exp(-\imath\tilde{w})\end{aligned}$$

This function is clearly holomorphic, and is a sum of the identity function plus a perturbation. Therefore, (Id, u) is a holo-decomposition of f (see appendix B), where:

$$\begin{aligned} u : \mathbb{T}_{\mathbb{C}} &\to \mathbb{C} \\ \tilde{w} &\mapsto -\frac{\tilde{\bar{E}}}{2\imath}\exp(\imath\tilde{w}) + \frac{\tilde{E}}{2\imath}\exp(-\imath\tilde{w})\end{aligned}$$

To apply the semi-global inversion theorem B.11, it remains to find an open set A such that for every closed subset $B \subset A$, we have $f'(z) \neq 0$ and $\|u'\|_A < 1$. In the set $\mathcal{D}^w_{\Lambda_0,r,\rho,\rho',w'_{max}}$ determined before, these two conditions are verified.

Proposition 1.3. *Let* $0 < r < \Lambda_0$, $0 < \rho \leq \sqrt{2(\Lambda_0 - r)}$, $0 < \rho' \leq \sqrt{\Lambda_0 - r} - \rho/\sqrt{2}$ *and*

$$w'_{max} = \operatorname{arccosh}\left(\sqrt{\frac{\Lambda_0 - r}{1 + \frac{3\rho'^2}{2(\Lambda_0 - r)}}}\frac{1}{(\rho + 2\rho')}\right).$$

In the set $\mathcal{D}^w_{\Lambda_0,r,\rho,\rho',w'_{max}}$*, defined in* (1.25)*, the complex Kepler equation induces a diffeomorphism* $(\Re f, \Im f)$ *from the set*
$A = \mathbb{T}\times]-w'_{max}, w'_{max}[$ *in its image.*

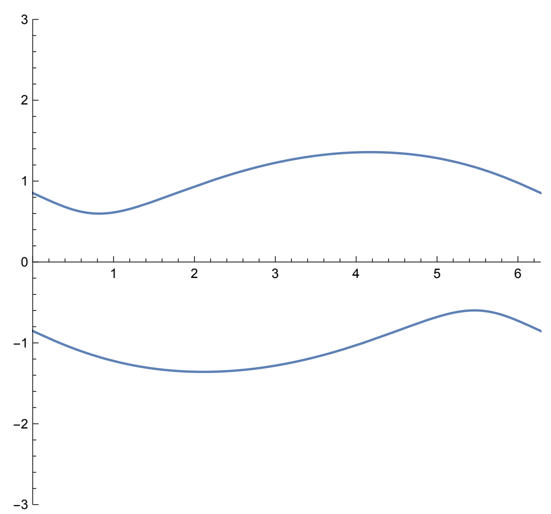

Figure 1.9: Image of the sets $\mathbb{T} \times \{-t\}$ and $\mathbb{T} \times \{t\}$ by $\tilde{f}$, for $t = u'_{max}/2$, $e = 0.2$, $e' = 0.2$.

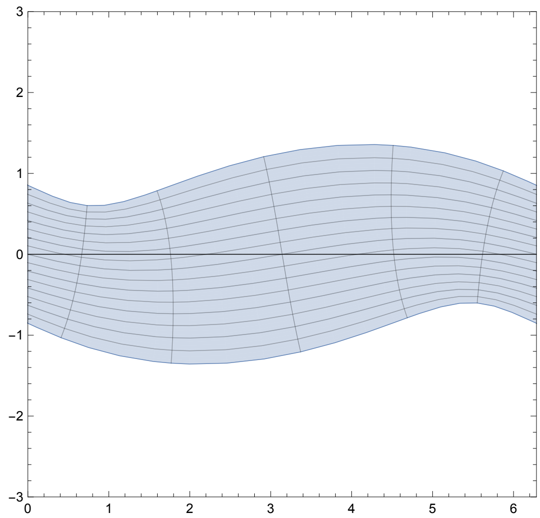

Figure 1.10: Image of the set $U_t\mathbb{T}\times]-t, t[$ by $\tilde{f}$, for $t = u'_{max}/2$, $e = 0.2$, $e' = 0.2$.

Proof. The proof is straightforward using our previous work and theorem B.11 of appendix B. □

Surjectivity and domain for the mean longitude

After finding a domain on the eccentric longitudes on which Kepler's equation induces a diffeomorphism, it remains to find a domain in the mean longitude for which this property holds. The image of a domain $U = \mathbb{T}\times]-w'_{max}, w'_{max}[$ by this diffeomorphism is for now unknown. In Poincaré coordinates, we want to find a domain $V = \mathbb{T}\times]-\lambda'_{max}, \lambda'_{max}[$ whose image by the inverse map is contained in U.

Elliptic coordinates:
Consider, for $(e, e') \in \mathbb{R}^+ \times \mathbb{R}$ and $t \leq u'_{max}$:

$$\begin{aligned} \tilde{f}\ U_t = \mathbb{T} \times (-t, t) &\to g(U) \\ (u, u') &\mapsto (u - e\sin(u)\cosh(u') + e'\cos(u)\sinh(u'), \\ &\qquad u' - e'\sin(u)\cosh(u') - e\cos(u)\sinh(u')) \end{aligned}$$

To understand better the image of U_t by $\tilde{f}$, we can consider the image of the sets $\mathbb{T} \times \{t\}$ and $\mathbb{T} \times \{-t\}$. An example for $e = 0.2$, $e' = 0.2$ and $t = u'_{max}/2$ is given in figure 1.9. The image of the set U_t by $\tilde{f}$ is contained between those curves (figure 1.10).

To define a rectangular set $\mathbb{T}\times]-l'_{max}, l'_{max}[$ in the image of the set U_t by $\tilde{f}$, it is therefore enough to concentrate on the minimum of l' when u varies. Besides, we show that this value is maximal for $t = u'_{max}$, which means that the maximum value l'_{max} on which there exists a diffeomorphism is as well a singular point of Kepler's equation. Indeed, we have

$$l' = u' - e'\sin(u)\cosh(u') - e\cos(u)\sinh(u'), \tag{1.29}$$

and for $u' = u'_{max}$, the location of a minimum of this function is a point for which u and u' goes to zero at the same time. Therefore, the width l'_{max} corresponds to the minimal value of l' such that the inverse map has a singular point.
Recall the definition of $\mu = \sqrt{e'^{\,2}\cosh^2(u') + e^2\sinh^2(u')}$. If $\mu = 0$ then $l' = u'$. Now if $\mu > 0$,

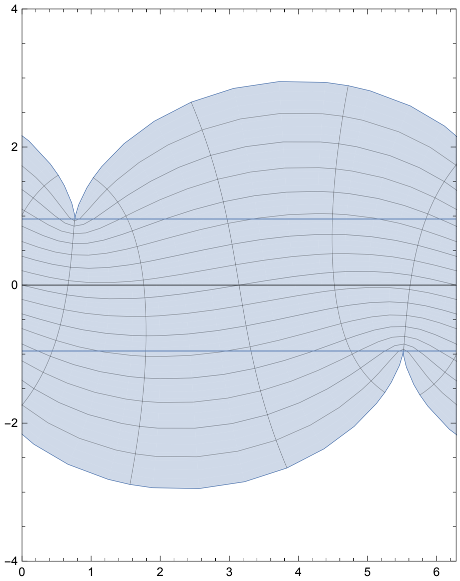

Figure 1.11: Image of the set $\mathbb{T}\times]-u'_{max}, u'_{max}[$ by the function g and of the lines of equation $l' = \pm(u'_{max} - \mu)$, with $e = 0.2$, $e' = 0.2$.

equation (1.12) was giving:

$$\begin{cases} \cos(u) = \pm\dfrac{e\sinh(u')}{\mu} \\ \sin(u) = \pm\dfrac{e'\cosh(u')}{\mu} \end{cases}$$

Injecting in the equation of l', it gives

$$l' = u' \mp \mu.$$

For $\mu \geq 0$, the minimum of the right term is $l' = u' - \mu$. As said before, this value is maximal in the set $u' \in [u'_{max}, u'_{max}]$ when $u' = u'_{max}$. Observe that for e and e' sufficiently small, l' is positive. The width l'_{max} is hence given by the formula:

$$l'_{max} = u'_{max} - \sqrt{e'^{\,2}\cosh^2(u'_{max}) + e^2\sinh^2(u'_{max})}.$$

In our work, we are interested in the case $l'_{max} > 0$. Figure 1.11 shows the limit case when $t = u'_{max}$, on which we drew the limit value l'_{max} and $-l'_{max}$. We can observe different limit cases for which l'_{max} goes to zero:

- In the case the eccentricity is large enough, then l'_{max} defined as before might be negative. Hence, there would exist angles l for which the fiber of $(l, 0)$ is not in the set $U_{u_{max}}$. The width l' will then be considered to be null. This case is represented in figure 1.12.

- In the case of a non-null eccentricity, when choosing a value $t < u'_{max}$, if t is too small then we would have $l'_{max} < 0$. Indeed, a small enough t implies that we did not "fill" any set of the form $\mathbb{T}\times]-l'_{max}, l'_{max}[$. The width l' will again be considered to be 0. This case is represented in figure 1.13.

Since solving the equation $l'_{max} > 0$ requires to solve an equation of the form $\exp x + cx = 0$, we will express the solution implicitly.

Lemma 1.4. *Let $e > 0$, $e' \in \mathbb{R}$, and $l'_{max} = u'_{max} - \sqrt{e'^{\,2}\cosh^2(u'_{max}) + e^2\sinh^2(u'_{max})}$. If $l'_{max} > 0$, then g^{-1} is a (analytic) diffeomorphism on the set $\mathbb{T}\times]-l'_{max}, l'_{max}[$, and its image is contained in the set $U_{u'_{max}}$.*

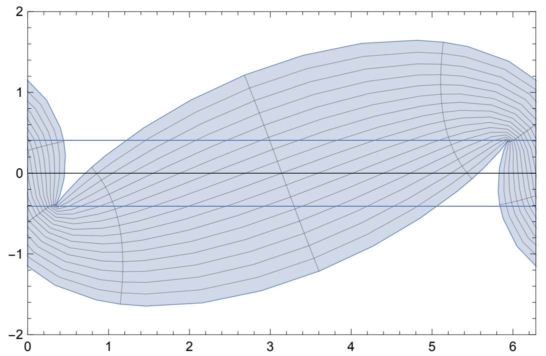

Figure 1.12: Image of the set $\mathbb{T}\times]-u'_{max}, u'_{max}[$ by g and of the lines of equation $l' = \pm(u'_{max} - \mu)$, with $e = 0.9$, $e' = 0.7$.

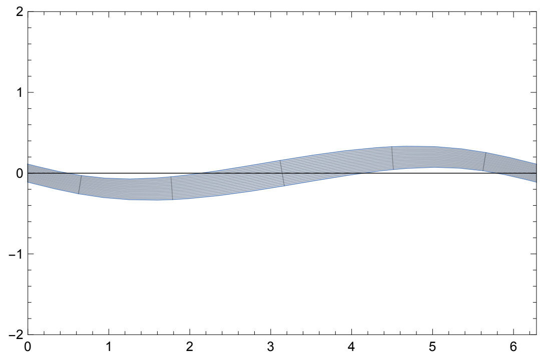

Figure 1.13: Image of the set $\mathbb{T}\times]-t, t[$ by g, with $t = u'_{max}/15$, $e = 0.2$, $e' = 0.2$.

Poincaré coordinates:
Let $\Lambda_0 > 0$, $0 < r < \Lambda_0$, $0 < \rho \leq \sqrt{2(\Lambda_0 - r)}$, $0 < \rho' \leq \sqrt{\Lambda_0 - r} - \rho/\sqrt{2}$. The function we want to consider this time is the following:

$$\begin{aligned} \tilde{h}\ \mathbb{T}\times]-t, t[&\to \mathbb{C} \\ \tilde{w} \quad &\mapsto \tilde{w} - \frac{1}{2\imath\sqrt{\tilde{\Lambda}}}\sqrt{1 - \frac{\tilde{F}\tilde{\bar{F}}}{4\tilde{\Lambda}}}\left(\tilde{F}\exp(\imath\tilde{w}) - \tilde{\bar{F}}\exp(-\imath\tilde{w})\right) \end{aligned}$$

Now let $w' > 0$, and consider the following definitions:

$$\begin{cases} a_1 = \dfrac{1}{\sqrt{\Lambda_0 - r}} & b_1 = \dfrac{r}{\sqrt{\Lambda_0}(\Lambda_0 - r)} \\ a_2 = 1 + \dfrac{3\rho'^2}{2(\Lambda_0 - r)} & b_2 = \dfrac{2\rho'(\Lambda_0 + r)(\rho' + \sqrt{2}\rho) + r(\rho + \sqrt{2}\rho')^2}{4(\Lambda_0 - r)^2} \\ a_3(w') = (\rho + 2\rho')\cosh w' & b_3(w') = \rho\sinh w' + 2\rho'\cosh w' \end{cases} \tag{1.30}$$

We shall prove the following statement:

Proposition 1.5. *Let $\tilde{\Lambda} \in B(\Lambda_0, r)$ and $(\tilde{\xi}, \tilde{\eta}) \in \mathbb{C}^2$ such that there exists $\xi_0 + \imath\eta_0 \in B(0, \rho)$ with $\tilde{\xi} \in B(\xi_0, \rho')$ and $\tilde{\eta} \in B(\eta_0, \rho')$. Let t be such that*

$$0 < t \leq w'_{max} = \operatorname{arccosh}\left(\sqrt{\frac{\Lambda_0 - r}{1 + \frac{3\rho'^2}{2(\Lambda_0 - r)}}}\frac{1}{(\rho + 2\rho')}\right).$$

Define:

$$\lambda'_{max} = t - (a_1 a_2 b_3(t) + a_2 a_3(t) b_1 + a_1 a_3(t) b_2 + b_1 b_2 b_3(t)). \tag{1.31}$$

If $\lambda'_{max} > 0$ and $4r < 3\Lambda_0$, then the function $(\tilde{h})^{-1}$ is a diffeomorphism on the set $\mathbb{T}\times]-\lambda'_{max}, \lambda'_{max}[$ onto its image, which is a subset of the set $\mathbb{T}\times]-t, t[$.

Observe that the variable λ'_{max} for a fixed t is positive if r and ρ' are small enough. Hence, we can always find a value of r and ρ such that we have a diffeomorphism.

To prove this proposition, we will need a mini-lemma.

Lemma 1.6. *Let z be a complex number of the form $z = a + \imath b$. Consider a complex number $c + \imath d$ such that $(c + \imath d)^2 = 1 + z$. If $|a| \leq \frac{3}{4}$, then the inequality $|d| \leq |b|$ holds.*

Proof. Let a, b, c, d defined as in the statement. We have the relations:

$$c^2 = 1 + a + d^2, \qquad 2cd = b$$

Hence:

$$|a| \leq \frac{3}{4} \Rightarrow |1 + a| \geq \frac{1}{4} \Rightarrow |c| \geq \frac{1}{2}$$

Besides:

$$|c| \geq \frac{1}{2} \Rightarrow |b| = |2cd| \geq |d| \tag{1.32}$$

□

We can now prove the proposition.

Proof. First, define $\lambda' = \Im(\tilde{h}(w, w'))$. We want to know, as before, if for a value of $w' > 0$ this function have a minimum that is strictly positive. Using the same arguments as in the previous paragraph, we will then have a width λ' strictly positive such that the inverse map is a diffeomorphism.
Let us take a look at the value

$$\lambda'_{max} = w' - \sup_{\mathcal{D}^{w,1}_{\Lambda_0, r, \rho, \rho', w'_{max}}} (|\tilde{E}|) \cosh w'.$$

We showed earlier that:

$$\sup_{\mathcal{D}^{w,1}_{\Lambda_0, r, \rho, \rho', w'_{max}}} |\tilde{E}| < \frac{1}{\sqrt{\Lambda_0 - r}} \sqrt{1 + \frac{3\rho'^2}{2(\Lambda_0 - r)|}} (\rho + 2\rho') \,.$$

Yet, this estimate is not precise enough to work with. Indeed, looking at the case $w' = w'_{max}$, and injecting the results in the value of λ'_{max}, we obtain $\min \lambda' = u'_{max} - 1$. The 1 comes from the maximal value of the derivative we found before, when we bounded at the same time the factor $\sin x$ aand $\cos x$. We need here to be more precise, trying not to "break" the real structure of Kepler's equation when considering its analytic continuation. We therefore have to take a closer look to the imaginary part of $\tilde{h}$, which can be seen as a perturbation of the real case for small values of the r, ρ', w'.
Let us study the expression of $\left(\tilde{F} \exp(\imath \tilde{w}) + \bar{\tilde{F}} \exp(-\imath \tilde{w})\right)$, in order to separate its real and imaginary part of it. First, there exists $0 \leq r_0 < \rho$, $0 \leq r_1, r_2 < \rho'$, and $\theta, \theta_1, \theta_2 \in \mathbb{T}$ such that:

$$\begin{cases} \tilde{F} = r_0 \exp(\imath\theta) + r_1 \exp(\imath\theta_1) + \imath r_2 \exp(\imath\theta_2) \\ \bar{\tilde{F}} = r_0 \exp(-\imath\theta) + r_1 \exp(\imath\theta_1) - \imath r_2 \exp(\imath\theta_2) \end{cases}$$

Hence,

$$\frac{1}{2\imath} \left(\tilde{F} \exp(\imath \tilde{w}) + \bar{\tilde{F}} \exp(-\imath \tilde{w})\right) = r_0 \sin(\theta + \tilde{w}) + r_1 \exp(\imath\theta_1) \sin(\tilde{w}) + r_2 \exp(\imath\theta_2) \cos(\tilde{w}).$$

We deduce the 2 following inequalities:

$$\begin{cases} \left| \Re \left(\frac{1}{2\imath} \left(\tilde{F} \exp(\imath \tilde{w}) + \bar{\tilde{F}} \exp(-\imath \tilde{w})\right) \right) \right| < (\rho + 2\rho') \cosh(w') = a_3(w') \\ \\ \left| \Im \left(\frac{1}{2\imath} \left(\tilde{F} \exp(\imath \tilde{w}) + \bar{\tilde{F}} \exp(-\imath \tilde{w})\right) \right) \right| < \rho \sinh(w') + 2\rho' \cosh(w') = b_3(w') \end{cases} \tag{1.33}$$

In the imaginary part, we have the small terms ρ' or $\sinh w'$ in factor. We will now try to bound from above the imaginary and real parts of the two other terms appearing in the imaginary part of $\tilde{h}$. observe that we are not interested in a precise estimate for the real terms, since they are not small. Yet, it is necessary to highlight the fact that the imaginary parts remains small. Recall that we showed

$$\left|\Re\left(\sqrt{1-\frac{\tilde{F}\bar{\tilde{F}}}{4\tilde{\Lambda}}}\right)\right| \leq \left|\sqrt{1-\frac{F\bar{F}}{4\Lambda}}\right| \leq 1+\frac{3\rho'^2}{2(\Lambda_0-r)} = a_2.$$

As for the imaginary part, we will use lemma 1.6. Before, we need to make a further assumption to verify the hypothesis of the lemma. Observe that $|a| \leq |a+\imath b| = |z|$. Hence, we can ensure that the norm of z is small enough to verify our hypothesis. We know that

$$\left|\frac{\tilde{F}\bar{\tilde{F}}}{4\tilde{\Lambda}}\right| < \frac{(\rho+\sqrt{2}\rho')^2}{4(\Lambda_0-r)},$$

hence, to apply the lemma, we require

$$\frac{(\rho+\sqrt{2}\rho')^2}{(\Lambda_0-r)} < 3 \;\Leftrightarrow\; \rho' < \sqrt{\frac{3}{2}(\Lambda_0-r)}-\frac{\rho}{\sqrt{2}}, \tag{1.34}$$

which is a hypothesis of the proposition.
We can now determine the norm of the imaginary part of the term we are studying, so as to deduce the desired bound. We have:

$$\frac{\tilde{F}\bar{\tilde{F}}}{4\tilde{\Lambda}} = \frac{\tilde{\xi}^2+\tilde{\eta}^2}{4\tilde{\Lambda}} \;\Rightarrow\; \Im\left(\frac{\tilde{F}\bar{\tilde{F}}}{4\tilde{\Lambda}}\right) = \Im\left(\frac{\tilde{\xi}^2+\tilde{\eta}^2}{4\tilde{\Lambda}}\right) = \frac{1}{4|\tilde{\Lambda}|^2}\Im\left(\bar{\tilde{\Lambda}}(\tilde{\xi}^2+\tilde{\eta}^2)\right).$$

At the frontier of our domain, we have $\tilde{\xi} = \rho\cos\theta + \rho'\exp(\imath\theta_1)$, and $\tilde{\eta} = \rho\sin\theta + \rho'\exp(\imath\theta_2)$. We get:

$$\begin{aligned}\tilde{\xi}^2+\tilde{\eta}^2 &= \rho^2+\rho'^2(\cos(2\theta_1)+\cos(2\theta_2)) + 2\rho\rho'(\cos\theta\cos\theta_1+\sin\theta\cos\theta_2)\\ &\quad + \imath\left(\rho'^2(\sin(2\theta_1)+\sin(2\theta_2)+2\rho\rho'(\cos\theta\sin\theta_1+\sin\theta\sin\theta_2)\right).\end{aligned}$$

Whence,

$$|\Im(\tilde{\xi}^2+\tilde{\eta}^2)| \leq 2\rho'(\sqrt{2}\rho+\rho'),$$

as for the real and imaginary parts of $\bar{\tilde{\Lambda}}$, the upper bounds $|\Re\bar{\tilde{\Lambda}}| < \Lambda_0 + r$, and $|\Im\bar{\tilde{\Lambda}}| < r$ are straightforward. Multiplying these two terms, we obtain the following upper bound

$$\left|\Im\left(\sqrt{1-\frac{\tilde{F}\bar{\tilde{F}}}{4\tilde{\Lambda}}}\right)\right| < \frac{1}{4(\Lambda_0-r)^2}\left(2\rho'(\Lambda_0+r)(\rho'+\sqrt{2}\rho)+r(\rho+\sqrt{2}\rho')^2\right) = b_2.$$

Gathering the computations, we have the following implication:

$$\frac{(\rho+\sqrt{2}\rho')^2}{(\Lambda_0-r)} < 3 \Rightarrow \left|\Im\left(\sqrt{1-\frac{\tilde{F}\bar{\tilde{F}}}{4\tilde{\Lambda}}}\right)\right| \leq b_2.$$

It remains to consider the term $\sqrt{\tilde{\Lambda}}^{-1}$. It can be put under the form:

$$\frac{1}{\sqrt{\tilde{\Lambda}}} = \frac{\sqrt{\bar{\tilde{\Lambda}}}}{|\tilde{\Lambda}|}.$$

Yet $\tilde{\Lambda} = \Lambda_0 + s\exp(\imath\theta)$ with $0 < s < r$. Using again lemma 1.6, we get the upper bounds:

$$\begin{cases} \left|\Re\left(\frac{1}{\tilde{\Lambda}}\right)\right| \leq \frac{1}{\sqrt{|\tilde{\Lambda}|}} < \frac{1}{\sqrt{\Lambda_0 - r}} = a_1 \\ \\ \left|\Im\left(\frac{1}{\tilde{\Lambda}}\right)\right| < \frac{r}{\sqrt{\Lambda_0}(\Lambda_0 - r)} = a_2 \quad \text{if} \quad \frac{r}{\Lambda_0} \leq \frac{3}{4} \end{cases}$$

With all these estimates, one can bound the imaginary part of the product of the three terms, using either their real parts a_i or their imaginary parts b_i for $i = 1, 2, 3$. Then, using the considerations we made while studying the same problem in elliptic coordinates, the result follows. □

1.3 Estimates on the norm of the disturbing function

In this part, we determine a bound on the norm of the perturbation in the plane planetary three-body problem. Recall the formula (1.4) of the perturbation:

$$H_{pert}(P_1, P_2, Q_1, Q_2) = \frac{G_{grav}\mu_1 m_2}{|Q_2|} \sum_{n\geq 2} \sigma_n P_n(\cos(S)) \left(\frac{|Q_1|}{|Q_2|}\right)^n,$$

with $\sigma_n = \sigma_0^{n-1} + (-1)^n \sigma_1^{n-1}$. There are different terms to study before deriving an explicit bound. First, we need to study the terms $|Q_1|$ and $|Q_2|$. Secondly, we need to study the norm of $P_n(\cos S)$ for a complex angle S. We will need to do some initial work on Legendre polynomials, and on the analytic continuation of their expression, as well as finding a bound on the complex cosine.

1.3.1 Discussion on the singularities

Before studying the norm of the perturbation, let us take a closer look at the singularities. Before expanding the perturbation with respect to the semi-major axis, we had three terms in the expression of the perturbation. These terms had the following denominators: $|Q_2|$, $|Q_2 - \sigma_0 Q_1|$ and $|Q_2 + \sigma_1 Q_1|$. For real variables, singularities can arise in different cases. The first term corresponds to the distance between the outer planet and the center of mass of the star and the inner planet: this singularity is artificial and depends on our choice of masses of the Kepler Hamiltonian. The second corresponds to the distance between the two planets, and the third one to the distance between the star and the outer planet. In the planetary problem we consider, the first singularity that can arise corresponds to the crossing of the orbits of the two planets, and hence to the term: $|Q_2 - \sigma_0 Q_1|$.

When studying the analytic continuation, we saw that there were two possibilities for this denominator to go to zero. First, if $\tilde{Q}_1 = \tilde{Q}_2$, and secondly if $\tilde{\tilde{Q}}_1 = \tilde{\tilde{Q}}_2$. Since we are looking at a symmetrical complex neighborhood (more exactly polydiscs) around the real set of orbits, these two conditions are equivalent. Now, looking at the elliptic coordinates, we have $\tilde{Q}_i = \tilde{a}_i(1 - \tilde{e}_i \cos(\tilde{u}_i + \tilde{g}_i))$, for $i = 1, 2$. Therefore, bounding from above $\tilde{a}_1$ and from below $\tilde{a}_2$, we are left with the derivative of Kepler's equation. Since $|Q_1| > 0$, we are not looking at the singular points of Kepler's equation. Indeed, those corresponds to the singularity between the star and the outer planet, though the singularity of the distance between the two planets occurs before this one. When $|\tilde{a}_2(1 - \tilde{e}_2 \cos(\tilde{u}_2 + \tilde{g}_2)| \sim |\tilde{a}_1|$, we reach this singularity. Hence, the formula of a singularity depends on the variables of the two bodies, and it occurs before the eventual singular points of the diffeomorphism between the eccentric and mean longitudes of the second planet. Besides, if we let the eccentricities of the two planets belong to the same set, this singularity happens before the singularity between the inner planet and the star, as well as the singular points of the diffeomorphism between the eccentric and mean longitudes for the first planet. We need to derive a sufficient condition in Poincaré coordinates to avoid this singularity. Yet a necessary condition would necessitate to overcome the difficulties of finding exactly the singular points of the complex Kepler equation in Poincaré coordinates.

1.3.2 Bound on the norm of the complex distances star-planet

We are looking at the terms $|Q_i|$ for $i = 1, 2$. Before the step of analytic continuation, their formulas are given by

$$q = |Q| = a(1 - e\cos u).$$

After this step, we therefore have

$$\tilde{q} = \tilde{a}(1 - \tilde{e}\cos\tilde{u}).$$

Notice that the estimates on $1 - \tilde{e}\cos\tilde{u}$ have already been obtained while studying the singular points of the complex Kepler equation. It remains to adapt these here. We can start by defining a domain of analyticity on which we want our estimates.
Let $\Lambda_{0,i} > 0$ for $i = 1, 2$, $r < \Lambda_{0,1}, \Lambda_{0,2}$, $0 < \rho < \sqrt{2(\Lambda_{0,i} - r)}$, $\rho' < \sqrt{\Lambda_{0,i} - r} - \rho/\sqrt{2}$ and $\lambda'_{max} > 0$. Define the following set:

$$\begin{aligned}\mathcal{D}_{\Lambda_0,r,\rho,\rho',\lambda'_{max}} = \{&(\Lambda_1, \Lambda_2, \lambda_1, \lambda_2, \xi_1, \xi_2, \eta_1, \eta_2) \in \mathbb{C}^2 \times \mathbb{T}^2_{\mathbb{C}} \times \mathbb{C}^4 | \\ &\text{for i=1,2:} \Lambda_i \in B(\Lambda_{0,i}, r),\ |\Im\lambda_i| < \lambda'_{max}, \\ &\exists(\xi_{0,i}, \eta_{0,i}) \in B(0, \rho) \text{ s.t. } \xi_i \in B(\xi_{0,i}, \rho'),\ \eta_i \in B(\eta_{0,i}, \rho')\}.\end{aligned}$$

Forgetting the indices, we write:

$$\tilde{h} : (\tilde{w}, \tilde{\xi}, \tilde{\eta}) \mapsto \tilde{w} - \frac{\bar{\tilde{E}}}{2\imath}\exp(\imath\tilde{w}) + \frac{\tilde{E}}{2\imath}\exp(-\imath\tilde{w}),$$

where the variable E depends on the coordinates ξ and η. We have

$$\tilde{q} = \frac{\tilde{\Lambda}^2}{G_{grav}M\mu^2}\,\frac{\partial\tilde{h}}{\partial\tilde{w}}(\tilde{w}, \tilde{\xi}, \tilde{\eta}).$$

If $\tilde{\Lambda} \in B(\Lambda_0, r)$ with $r < \Lambda_0$, then the term $\tilde{\Lambda}^2$ verifies $(\Lambda_0 - r)^2 < |\tilde{\Lambda}^2| < (\Lambda_0 + r)^2$. Let us define the variable t (specific to each body) implicitly with the help of λ'_{max} and of the other analyticity width, such as done in proposition 1.5, in the following way:

$$\lambda'_{max} = t - (a_1a_2b_3(t) + a_2a_3(t)b_1 + a_1a_3(t)b_2 + b_1b_2b_3(t)).$$

If there exists indeed $t > 0$ solution of this implicit equation, then the study of singular points in the Poincaré coordinates gave:

$$\left|1 - \frac{\partial\tilde{h}}{\partial\tilde{w}}(\tilde{w}, \tilde{\xi}, \tilde{\eta})\right| < \frac{1}{\sqrt{\Lambda_0 - r}}\sqrt{1 + \frac{3\rho'^2}{2(\Lambda_0 - r)}}(\rho + 2\rho')\cosh t = l(\Lambda_0, r, \rho, \rho', t).$$

Moreover, if we had $l(\Lambda_0, r, \rho, \rho', t) < 1$, then Kepler's equation was inducing a diffeomorphism between the eccentric longitudes and the mean longitudes for $|\Im\lambda| < \lambda'_{max}$.
Hence, we deduce that, under the assumption $l(\Lambda_0, r, \rho, \rho', t) < 1$ and $\lambda'_{max} > 0$:

$$\begin{aligned}&\frac{(\Lambda_0 - r)^2}{G_{grav}M\mu^2}(1 - l(\Lambda_0, r, \rho, \rho', t)) < |\tilde{q}| < \frac{(\Lambda_0 + r)^2}{G_{grav}M\mu^2}(1 + l(\Lambda_0, r, \rho, \rho', t)) \\ &\frac{(\Lambda_0 - r)^2}{G_{grav}M\mu^2}(1 - l(\Lambda_0, r, \rho, \rho', t)) < |\tilde{q}| < 2\frac{(\Lambda_0 + r)^2}{G_{grav}M\mu^2}. \end{aligned} \tag{1.35}$$

The second inequation is not optimal, though it is succinct and simple to use. Indeed, we replaced the term $1 + l$ by 2, which is a choice of convenience. It corresponds to consider the second body on a circle of radius $2\tilde{a}_2$ instead of its ellipse. Nevertheless, it implies that when injecting this value in the series of the perturbation, it increases its norm artificially.

1.3.3 Estimates on Legendre polynomials

The complex Legendre polynomials

In this section, we first recall some results on the Legendre polynomials, and then we give an upper bound on the Legendre polynomials evaluated on some complex set. We focus here on relations that are interesting in our study, although a lot of work have been done on these polynomials (see [1, 3] for instance). In our problem, these polynomials arise in the plane planetary three-body problem because of the following relation:

$$\frac{1}{\sqrt{1-2xz+z^2}} = \sum_{n=0}^{\infty} P_n(x) z^n.$$

One can work on those using the polynomials $U_n(x) = (x^2-1)^n$ for $n \geq 0$, we have:

$$P_n(x) = U_n^{(n)}(x) = \frac{d^n U_n}{dx^n}(x).$$

Observe that the polynomial U_n is of degree $2n$, and admits two repeated roots of order n: -1 and 1. Hence, considering the successive derivatives of these polynomials, the polynomial P_n is of degree n, and all of its roots are simple roots belonging to the set $]-1,1[$. One more definition will be important for us, it is the explicit form of P_n for $n \geq 0$:

$$P_n(x) = \frac{1}{2^n} \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n}{k} \binom{2n-2k}{n} x^{n-2k}.$$

From this relation, we can deduce several results. First the Legendre polynomial are either even or odd, depending on the parity of n. Next, the leading coefficient of P_n is

$$\frac{(2n)!}{2^n (n!)^2} = \prod_{k=1}^{n} \frac{2k-1}{k}. \tag{1.36}$$

The polynomials are also defined by a recurrence relation, the formula for $n > 0$ being

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x),$$

with $P_0(x) = 1$, and $P_1(x) = x$. We therefore deduce that $P_n(1) = 1$ and $P_{2n+1}(0) = 0$ for all n. We want to find a suitable way of expressing the Legendre polynomials, and derive estimates on an upper bound. We have, for all $n \geq 0$:

$$\begin{cases} P_{2n}(x) = \dfrac{(4n)!}{2^{2n}(2n)!} \displaystyle\prod_{i=1}^{n} (x^2 - \lambda_{2n,i}^2) \\ \\ P_{2n+1}(x) = \dfrac{(4n+2)!}{2^{2n+1}(2n+1)!} x \displaystyle\prod_{i=1}^{n} (x^2 - \lambda_{2n+1,i}^2) \end{cases},$$

where the roots $\lambda_{n,i} \in]-1,1[$. Define for $\rho > 0$:

$$B_0(\rho) = \{z \in \mathbb{C}, |z| < \rho\}. \tag{1.37}$$

We have:

Proposition 1.7. *$\forall x \in [-1,1]$, $\forall n \geq 0$:*

$$|P_n(x)| \leq 1. \tag{1.38}$$

$\forall z \in B_0(\rho)$ with $\rho \geq 0$, $\forall n \geq 0$

$$|P_n(z)| \leq \frac{(2n)!}{2^n n!} \left(\sqrt{1+\rho^2}\right)^n. \tag{1.39}$$

Proof. For $n \geq 1$, and $x \in [-1, 1]$:

$$Q_n(x) = (P_n(x))^2 + \frac{1-x^2}{n(n+1)}(P_n'(x))^2.$$

The polynomial Q_n is even, let us show that it is increasing for $x \in [0, 1]$. The fact that $Q_n(1) = 1$ and that $(P_n(x))^2 \leq Q_n(x)$ then finishes the proof.

$$\begin{aligned} Q_n'(x) &= 2P_n'(x)P_n(x) - \frac{2xP_n'(x)}{n(n+1)} + \frac{2(1-x^2)P_n'(x)P_n''(x)}{n(n+1)} \\ &= \frac{2P_n'(x)}{n(n+1)}\Big(n(n+1)P_n(x) - xP_n'(x) + (1-x^2)P_n''(x)\Big). \end{aligned} \tag{1.40}$$

Besides, $U_n'(x) = 2nx(x^2-1)^{n-1}$, whence

$$(x^2-1)U_n'(x) = 2nxU_n(x).$$

Computing the (n+1)-th derivative, we obtain:

$$(x^2-1)P_n''(x) + 2xP_n'(x) = n(n+1)P_n(x).$$

Injecting this relation in equation (1.40), we obtain:

$$Q_n'(x) = \frac{2x(P_n'(x))^2}{n(n+1)}$$

Thus, for $x \geq 0$, the derivative is increasing, and $|P_n(x)| \leq Q_n(x) \leq 1$ for all $x \in [-1, 1]$.

To show the second result, notice that the Legendre polynomials can be decomposed in two types of polynomials: $z^2 - \lambda^2$ with $0 < \lambda < 1$ or z. Let us begin by a study on these polynomials to deduce the final result.

- $R(z) = z$: it is straightforward that the norm of R on $B_0(\rho)$ is bounded by ρ.
- $R(z) = z^2 - \lambda^2$, $0 < \lambda < 1$: let $z = r\exp(\imath\theta)$, with $r \leq \rho$. We can start by bounding the norm of this polynomial using the argument of z. We have $R(z) = r^2\exp(2\imath\theta) - \lambda^2$and hence,

  $$|R(z)|^2 = r^4 + \lambda^4 - 2\lambda^2 r^2\cos(2\theta).$$

  The norm of this polynomial is therefore maximal for $\theta \in \{\pi/2, 3\pi/2\}$, in other words, if z the real part of z is null. We then have $|R(z)| = r^2 + \lambda^2$. Given the domain of definition of these value, we have the following bound:

  $$|R(z)| < \rho^2 + 1 = \lim_{r\to\rho,\ \lambda\to 1} r^2 + \lambda^2$$

The study of these two types of polynomials is in fact enough to conclude. Indeed the maximum of those terms can be taken on the imaginary axis, or in other words letting z go to $\pm\imath\rho$. If we do not know where the roots of the nth Legendre polynomial exactly are, we know that they are simple and in the set $]-1, 1[$. We will therefore consider the worst case, and take $|\lambda| \leq 1$. As well, even if 0 is the root of the polynomial $R(z) = z$, we can bound it by $\sqrt{1+\rho^2}$ on our domain in order to simplify the final result. In the end, we have:

$$\prod_{i=1}^{\lfloor n/2 \rfloor}(x^2 - \lambda_{n,i}^2) \leq \left(\sqrt{1+\rho^2}\right)^n.$$

Multiplying this expression by the leading coefficient we determined before, and finally, for $z \in B_0(\rho)$:

$$\left|\frac{1}{\sqrt{1-2zy+y^2}}\right| \leq \sum_{n=0}^{\infty} \frac{(2n)!}{2^n(n!)^2}\left(|y|\sqrt{1+\rho^2}\right)^n$$

This series is convergent for $y \in B_0\left(\frac{1}{\sqrt{1+\rho^2}}\right)$, in other words for $|y|\sqrt{1+\rho^2} \leq 1$. □

Application to the complex oriented angle between the planets

Recall the formula (1.9) of the complex oriented angle between the two planets:

$$\cos S = \frac{1}{2}\left(\sqrt{\frac{z_1\bar{z}_2}{\bar{z}_1 z_2}} + \sqrt{\frac{\bar{z}_1 z_2}{z_1\bar{z}_2}}\right)$$

Call $d = \sqrt{\frac{z_1\bar{z}_2}{\bar{z}_1 z_2}}$, we have $\cos S = \frac{1}{2}(d + 1/d)$. Considering the analytic continuation of this formulas, we call $d_m = \sup_D |\tilde{d}|$, where D is the complex neighborhood of our real domain. We then have $|\cos\tilde{S}| \leq 1/2(d_m + 1/d_m)$.

Notice that the difference between z and $\bar{z}$ consists only in changing the angle u in $-u$. Hence, when considering the analytic continuation of our function, by symmetry of the domains (which are polydiscs), the maxima and minima of z and $\bar{z}$ are the same. It is therefore enough to study only one of those terms (for instance z), and to determine the extrema of this value over the domain to find a bound on d_m.

First, let us work on d_m to simplify our work. We can observe that both $\tilde{z}$ and $\tilde{\bar{z}}$ are proportional to $\tilde{a}$, hence we can already remove this factor from our computation. The formula for z we are considering is

$$\tilde{z} = \cos\tilde{u} - \tilde{e} + \imath\sqrt{1-\tilde{e}^2}\sin\tilde{u}.$$

Call as well $d_i = \sqrt{z_i/\bar{z}_i}$, and hence $d = d_1/d_2$. We have $d_1 = \sqrt{z_1^2/(z_1\bar{z}_1)}$, where we let appear the value $z_1\bar{z}_1 = 1 - e\cos u$, which will help us using the previous work. In order to bound the term $|\tilde{z_1}^2|$ more subtly, we can see it as a perturbation of the real case. Indeed, we will compare the two terms $\tilde{z}\tilde{\bar{z}}$ and $|\tilde{z}\tilde{\bar{z}}|$. Recall:

$$\begin{aligned}\tilde{z}\tilde{\bar{z}} &= \tilde{x}^2 + \tilde{y}^2 \\ &= x^2 - x'^2 + y^2 - y'^2 + 2\imath(xx' + yy').\end{aligned}$$

As for the term $\tilde{z}\bar{\tilde{z}}$:

$$\begin{aligned}\tilde{z}\bar{\tilde{z}} &= (x-y')^2 + (x'+y)^2 \\ &= x^2 + y^2 + x'^2 + y'^2 - 2xy' + 2x'y \\ &= \tilde{z}\tilde{\bar{z}} + 2(x'^2 + y'^2 - xy' + x'y + \imath(xx' + yy')) \\ &= \tilde{z}\tilde{\bar{z}} + 2\imath(x - y' - \imath(x'+y))(x' + \imath y') \\ &= \tilde{z}\tilde{\bar{z}} + 2\imath\bar{\tilde{z}}z',\end{aligned}$$

where $z' = x' + \imath y'$. Hence, we have the bound:

$$\begin{aligned}|\tilde{z}|^2 &\leq |\tilde{z}\tilde{\bar{z}}| + 2|\bar{\tilde{z}}||z'| \\ &\leq |\tilde{z}\tilde{\bar{z}}| + 2|\tilde{z}||z'|.\end{aligned}$$

In this inequality, there is still a term $\tilde{z}$ appearing on the right side, whereas we want to bound it. Though, it is multiplied by a factor $2|z'|$, and since the latter goes to zero when we are considering the real set, it can be made as small as wanted by choosing a small polydisc. We will therefore consider it much smaller than $|\tilde{z}\tilde{\bar{z}}|$. We can rewrite this equation as follows:

$$|\tilde{z}|^2 - 2|\tilde{z}||z'| - |\tilde{z}\tilde{\bar{z}}| \leq 0.$$

We now have an inequation involving a polynomial of order 2, with a positive leading coefficient. Hence, the maximum value that $|\tilde{z}|$ can take corresponds to the equality case.

$$\Delta = 4(|z'|^2 + |\tilde{z}\tilde{\bar{z}}|).$$

The positive root of this equation is therefore:

$$|\tilde{z}|_{max} = |z'| + \sqrt{|\tilde{z}\tilde{\bar{z}}| + |z'|^2}.$$

Now, using the fact that $\sqrt{1+x} \leq 1 + \sqrt{x}$ for all $x > 0$, we have the following bound:

$$|\tilde{z}|_{max} \leq \sqrt{|\tilde{z}\tilde{\bar{z}}|}\left(1 + 2\frac{|z'|}{\sqrt{|\tilde{z}\tilde{\bar{z}}|}}\right).$$

Regarding the value of $\tilde{d}_1$:

$$|\tilde{d}_1| = \left|\sqrt{\frac{\tilde{z}_1}{\tilde{\bar{z}}_1}}\right| \leq \frac{|\tilde{z}_1|}{\sqrt{|\tilde{z}_1\tilde{\bar{z}}_1|}} \leq 1 + 2\frac{|z_1'|}{\sqrt{|\tilde{z}_1\tilde{\bar{z}}_1|}}$$

We can also estimate the maximum of $1/d_2$, it suffices to change the angle u_2 in $-u_2$: consider the ratio $1/d_2 = \sqrt{\tilde{\bar{z}}_2^2/(\tilde{z}_2\tilde{\bar{z}}_2)}$ and change the angle. After computation, we obtain

$$|\tilde{d}| \leq \left(1 + 2\frac{|z_1'|}{\sqrt{|\tilde{z}_1\tilde{\bar{z}}_1|}}\right)\left(1 + 2\frac{|z_2'|}{\sqrt{|\tilde{z}_2\tilde{\bar{z}}_2|}}\right).$$

We now have to bound $|z_i'|$. Consider

$$z' = x' + \imath y'$$
$$\begin{cases} x + \imath x' = \cos\tilde{w} - \tilde{e} = \cos w \cosh w' - \imath \sin w \sinh w' + e + \imath e' \\ y + \imath y' = \sqrt{1-\tilde{e}^2}\sin\tilde{w} = \sqrt{1-\tilde{e}^2}(\sin w \cosh w' + \imath \cos w \sinh w') \end{cases},$$

from which we deduce

$$\begin{cases} x' = -\sin w \sinh w' + e' \\ y' = \Re(\sqrt{1-\tilde{e}^2})\cos w \sinh w' + \Im(\sqrt{1-\tilde{e}^2})\sin w \cosh w' \end{cases}.$$

When trying to bound our terms, we will ensure that the imaginary terms remain small, by keeping in factor the analyticity width. Let us study more precisely the different terms.

1. $\Im(\sqrt{1-\tilde{e}^2})$: recall that by definition, $\sqrt{1-\tilde{e}^2} = 1 - \frac{\tilde{\xi}^2+\tilde{\eta}^2}{2\tilde{\Lambda}}$. We already studied this term, which was arising while looking at the complex Kepler equation.

$$\Im(\sqrt{1-\tilde{e}^2}) = \Im\left(\frac{\tilde{\xi}^2+\tilde{\eta}^2}{2\tilde{\Lambda}}\right) = \frac{1}{2|\tilde{\Lambda}|^2}\Im\left(\tilde{\bar{\Lambda}}(\tilde{\xi}^2+\tilde{\eta}^2)\right).$$

   And a bound on this term was:

$$|\Im(\sqrt{1-\tilde{e}^2})| < \frac{1}{2(\Lambda_0 - r)^2}\left(r\rho^2 + 2\rho'(\Lambda_0 + 2r)(\rho' + \sqrt{2}\rho)\right).$$

2. $\Re(\sqrt{1-\tilde{e}^2})$: this term is in factor of $\sinh u'$, it is not supposed to be small (it is close to 1 in the real domain), since the hyperbolic sine will play this role. A rough bound is:

$$|\Re(\sqrt{1-\tilde{e}^2})| = \left|\Re\left(1-\frac{\tilde{\xi}^2+\tilde{\eta}^2}{2\tilde{\Lambda}}\right)\right| \leq 1 + \left|\frac{\tilde{\xi}^2+\tilde{\eta}^2}{2\tilde{\Lambda}}\right|$$
$$\leq 1 + \frac{(\rho+\sqrt{2}\rho')^2}{2(\Lambda_0 - r)}.$$

3. e': for this term, we want to have one of the analyticity width in factor. By definition:

$$e = \sqrt{1-\frac{F\bar{F}}{4\Lambda}}\frac{\sqrt{F\bar{F}}}{\sqrt{\Lambda}}.$$

The first factor has already been studied in lemma 1.6. We had:

$$\left|\Re\left(\sqrt{1-\frac{F\bar{F}}{4\Lambda}}\right)\right| \leq 1 + \frac{3\rho'^2}{2(\Lambda_0 - r)},$$

and:

$$\frac{(\rho+\sqrt{2}\rho')^2}{(\Lambda_0 - r)} < 3 \Rightarrow \left|\Im\left(\sqrt{1-\frac{F\bar{F}}{4\Lambda}}\right)\right| \leq \frac{1}{4(\Lambda_0 - r)^2}\left(r\rho^2 + 2\rho'(\Lambda_0+2r)(\rho'+\sqrt{2}\rho)\right).$$

Let us now bound the real part of $\sqrt{\frac{\tilde{F}\tilde{\bar{F}}}{\tilde{\Lambda}}}$.:

$$\left|\Re\left(\sqrt{\frac{\tilde{F}\tilde{\bar{F}}}{\tilde{\Lambda}}}\right)\right| \leq \left|\sqrt{\frac{\tilde{F}\tilde{\bar{F}}}{\tilde{\Lambda}}}\right| = \sqrt{\left|\frac{\tilde{F}\tilde{\bar{F}}}{\tilde{\Lambda}}\right|} \leq \frac{(\rho+\sqrt{2}\rho')}{\sqrt{\Lambda_0 - r}}.$$

The computation of a bound on the imaginary part must let appear a small factor. The main difficulty will be again the presence of the square root, lemma 1.6 being not applicable to this case. Recall that $\tilde{\xi} = r_0\cos\theta + r_1\exp(\imath\theta_1)$ and $\tilde{\eta} = r_0\sin\theta + r_2\exp(\imath\theta_2)$, hence:

$$\begin{aligned}\tilde{F}\tilde{\bar{F}} &= \tilde{\xi}^2 + \tilde{\eta}^2 \\ &= r_0^2 + r_1^2\cos(2\theta_1) + r_2^2\cos(2\theta_2) + 2r_0r_1\cos\theta\cos\theta_1 + 2r_0r_2\sin\theta\cos\theta_2 \\ &\quad + \imath\left(r_1r_2(\sin(2\theta_1)+\sin(2\theta_2) + 2r_0r_1\cos\theta\sin\theta_1 + 2r_0r_2\sin\theta\sin\theta_2\right).\end{aligned}$$

We also obtained the inequality:

$$\frac{r}{\Lambda_0} \leq \frac{3}{4} \Rightarrow \left|\Im\left(\sqrt{\tilde{\bar{\Lambda}}}\right)\right| < \frac{r}{\sqrt{\Lambda_0}}$$

It remains to compute the imaginary part of $\sqrt{\tilde{\xi}^2+\tilde{\eta}^2}$. The bound will be this time more delicate to compute given the shape of the domain we are considering. Consider a complex number $z = a + \imath b$, and a root $z = (c+\imath d)^2$. We have the equality $c^2 - d^2 = a$ and $b = 2cd$. Hence, we deduce that

$$d^2 = -\frac{a}{2} + \frac{1}{2}\sqrt{a^2+b^2} = -\frac{a}{2} + \frac{|z|}{2}.$$

In our case, by symmetry of the domains in which r_1 and r_2 take their value, we will bound $|z| = |\xi^2+\eta^2|$ by $(r_0+\sqrt{2}\rho')^2$, and it only remains to calculate a lower bound on the real part of z. We have

$$|a| = |r_0^2 + r_1^2\cos(2\theta_1) + r_2^2\cos(2\theta_2) + 2r_0r_1\cos\theta\cos\theta_1 + 2r_0r_2\sin\theta\cos\theta_2|,$$

which also satisfies

$$|a| \geq r_0^2 - 2\sqrt{2}r_0\rho' - 2r_0'^2.$$

Whence,

$$\begin{aligned}\max|d|^2 &\leq \frac{1}{2}\left((r_0+\sqrt{2}\rho')^2 - r_0^2 - 2\sqrt{2}r_0\rho' - 2r_0'^2\right)\\ &\leq \sqrt{2}\rho'(\sqrt{2}\rho' + 2r_0)\\ &\leq \sqrt{2}\rho'(\sqrt{2}\rho' + 2\rho).\end{aligned}$$

We can now derive an upper bound on the imaginary part of the product of these terms, and it indeed goes to zero when considering null analyticity widths (*i.e.* in the real domain):

$$\begin{aligned}\left|\Im\left(\sqrt{\frac{\tilde{F}\tilde{\bar{F}}}{\tilde{\Lambda}}}\right)\right| &\leq \frac{1}{|\tilde{\Lambda}|}\left|\Re\left(\sqrt{\tilde{\bar{\Lambda}}}\right)\right|\left|\Im(\sqrt{\tilde{F}\tilde{\bar{F}}})\right| + \frac{1}{|\tilde{\Lambda}|}\left|\Im\left(\sqrt{\tilde{\bar{\Lambda}}}\right)\right|\left|\Re(\sqrt{\tilde{F}\tilde{\bar{F}}})\right|\\ &\leq \frac{\sqrt{\Lambda_0+r}}{\Lambda_0-r}\sqrt{\sqrt{2}\rho'(\sqrt{2}\rho'+2\rho)} + \frac{(\rho+\sqrt{2}\rho')}{\Lambda_0-r}\frac{r}{\sqrt{\Lambda_0}}\end{aligned}$$

Regarding the bound on the variable e', under the following hypotheses, we have

$$\begin{aligned}&(1)\quad \frac{(\rho+\sqrt{2}\rho')^2}{(\Lambda_0-r)} < 3,\\ &(2)\quad \frac{r}{\Lambda_0} \leq \frac{3}{4}.\end{aligned}$$

Hence $|e'| \leq e'_{max}(\Lambda_0, r, \rho, \rho')$ with

$$\begin{aligned}e'_{max}(\Lambda_0, r, \rho, \rho') &= \frac{(\rho+\sqrt{2}\rho')}{4(\Lambda_0-r)^{\frac{5}{2}}}\left(r\rho^2 + 2\rho'(\Lambda_0+2r)(\rho'+\sqrt{2}\rho)\right)\\ &+\left(1+\frac{3\rho'^2}{2(\Lambda_0-r)}\right)\left(\frac{\sqrt{\Lambda_0+r}}{\Lambda_0-r}\sqrt{\sqrt{2}\rho'(\sqrt{2}\rho'+2\rho)} + \frac{(\rho+\sqrt{2}\rho')}{\Lambda_0-r}\frac{r}{\sqrt{\Lambda_0}}\right)\end{aligned} \quad (1.41)$$

Observe that for the first time we have the presence of a term that does not decrease linearly with the analyticity width, but with the square root of ρ'; this is due to the expression of e.

4. $\sinh w'$, $\cosh w'$: regarding these two terms, we will not worry about their bound in Poincaré variables yet, but will simply bound $|w'|$ using w'_{max}. It is obvious that for a small value of w'_{max}, $\sinh w'$ is close to 0, and $\cosh w'$ is close to 1. To simplify again the calculation, we will use a rough bound on $\cosh w'$

$$\begin{aligned}|\cosh w'| &\leq 1 + |\sinh w'|\\ &\leq 1 + \sinh(w'_{max}),\end{aligned}$$

and will use the inequation $|\sinh w'| \leq \sinh w'_{max}$.

We can now bound z', and more particularly x' and y'. We have $|x'| \leq x'_{max}(\Lambda_0, r, \rho, \rho', w'_{max})$ and $|y'| \leq y'_{max}(\Lambda_0, r, \rho, \rho', w'_{max})$ with:

$$\begin{cases} x'_{max} &= \sinh w'_{max} + e'_{max}\\ y'_{max} &= \left(1+\frac{3\rho'^2}{2(\Lambda_0-r)}\right)\sinh w'_{max} + \frac{1+\sinh w'_{max}}{2(\Lambda_0-r)^2}\left(r\rho^2 + 2\rho'(\Lambda_0+2r)(\rho'+\sqrt{2}\rho)\right)\end{cases} \quad (1.42)$$

With the help of equation (1.35), we now bound d_m in a suitable way, using all the variables.

Proposition 1.8. *On the set $\mathcal{D}_{\Lambda_0,r,\rho,\rho',\lambda'_{max}}$, assume that the following hypotheses are satisfied for $i = 1, 2$:*

(1) *There exists $t_i > 0$ s.t. t_i verifies*
$\lambda'_{max} = t_i - (a_{1,i}a_{2,i}b_{3,i}(t_i) + a_{2,i}a_{3,i}(t_i)b_{1,i} + a_{1,i}a_{3,i}(t_i)b_{2,i} + b_{1,i}b_{2,i}b_{3,i}(t_i))$

(2) $l_i(\Lambda_{0,i}, r, \rho, \rho', t_i) = \dfrac{1}{\sqrt{\Lambda_{0,i} - r}}\sqrt{1 + \dfrac{3\rho'^2}{2(\Lambda_{0,i} - r)}}(\rho + 2\rho')\cosh t_i < 1$

(3) $\dfrac{r}{\Lambda_{0,i}} \leq \dfrac{3}{4}$.

where the variables a_i and b_i are defined in (1.30)*; Then the following inequalities hold:*

$$(i)\ |\tilde{d_i}| \leq d_{i,max} = 1 + 2\frac{x'_{i,max}(\Lambda_{0,i}, r, \rho, \rho', t_i) + y'_{i,max}(\Lambda_{0,i}, r, \rho, \rho', t_i)}{1 - l_i(\Lambda_{0,i}, r, \rho, \rho', t_i)}$$

$$(ii)\ d_m \leq d_{max} = d_{1,max} \times d_{2,max}$$

$$(iii)\ \cos(S) \leq \eta = \frac{1}{2}\left(d_{max} + \frac{1}{d_{max}}\right)$$

$$(iv)\ |P_n(\cos S)| \leq a_n(\Lambda_0, r, \rho, \rho', t_1, t_2) = \frac{(2n)!}{2^n(n!)^2}(\sqrt{1+\eta^2})^n$$

1.3.4 Final majoration

Using the work of the two last sections, we are able to derive a bound on the norm of the Hamiltonian. We will give two versions of it. A first version will be a raw one, with a lot of variables depending on other variables, though cumbersome to compute by hand: it is intended to a computer calculation. Yet, this formula will be closer to a realistic result, since we will not lose more optimality artificially. It is still dependent on some choices, and one is encouraged to make one's own choices when trying to compute a suitable expression. A second version will be a compact one. With more assumptions, one can simplify greatly the raw version, and make a statement simpler to use, though less optimal. With all the elements we computed, it is only a matter of choice to decide when to stop simplifying the equations.

Notations for the computation of the bound

We recall here every equations we need to state the result.
Let the variables $\Lambda_{0,1}, \Lambda_{0,2}, r, \rho, \rho', \lambda'_{max} \in \mathbb{R}^{+\star}$ be such that

$$r < \Lambda_0, \qquad \rho < \sqrt{2(\Lambda_0 - r)}, \qquad \rho' < \sqrt{\Lambda_0 - r} - \frac{\rho}{\sqrt{2}},$$

where $\Lambda_0 = \min_{i=1,2}\Lambda_{0,i}$. Call as well $\Lambda = (\Lambda_{0,1}, \Lambda_{0,2})$. Now define

$$\begin{cases} a_1 = \dfrac{1}{\sqrt{\Lambda_0 - r}} & b_1 = \dfrac{r}{\sqrt{\Lambda_0}(\Lambda_0 - r)} \\[2ex] a_2 = 1 + \dfrac{3\rho'^2}{2(\Lambda_0 - r)} & b_2 = \dfrac{2\rho'(\Lambda_0 + r)(\rho' + \sqrt{2}\rho) + r(\rho + \sqrt{2}\rho')^2}{4(\Lambda_0 - r)^2} \\[2ex] a_3(t) = (\rho + 2\rho')\cosh t & b_3(t) = \rho\sinh t + 2\rho'\cosh t \end{cases}$$

and

$$e'_{max}(\Lambda,r,\rho,\rho') = \frac{(\rho+\sqrt{2}\rho')}{4(\Lambda-r)^{\frac{5}{2}}}\left(r\rho^2+2\rho'(\Lambda_0+2r)(\rho'+\sqrt{2}\rho)\right) + \left(1+\frac{3\rho'^2}{2(\Lambda-r)}\right)\left(\frac{\sqrt{\Lambda+r}}{\Lambda-r}\sqrt{\sqrt{2}\rho'(\sqrt{2}\rho'+2\rho)}+\frac{(\rho+\sqrt{2}\rho')}{\Lambda_0-r}\frac{r}{\sqrt{\Lambda}}\right),$$

$$\begin{cases} x'_{max} = \sinh t + e'_{max} \\ y'_{max} = \left(1+\frac{3\rho'^2}{2(\Lambda-r)}\right)\sinh t + \frac{1+\sinh t}{2(\Lambda-r)^2}\left(r\rho^2+2\rho'(\Lambda+2r)(\rho'+\sqrt{2}\rho)\right) \end{cases}$$

where the e'_{max} is in fact a function of (Λ,r,ρ,ρ') and x'_{max}, y'_{max} are functions of (Λ,r,ρ,ρ',t). Call, for $t_1,t_2>0$:

$$\begin{cases} l_i = \frac{1}{\sqrt{\Lambda_{0,i}-r}}\left(1+\frac{3\rho'^2}{2(\Lambda_{0,i}-r)}\right)^{\frac{1}{2}}(\rho+2\rho')\cosh t_i, \\ d_{i,max} = 1+2\frac{x'_{i,max}(\Lambda_{0,i},r,\rho,\rho',t_i)+y'_{i,max}(\Lambda_{0,i},r,\rho,\rho',t_i)}{1-l_i}, \\ d_{max} = d_{1,max}\times d_{2,max}, \\ \eta = \frac{1}{2}\left(d_{max}+\frac{1}{d_{max}}\right), \\ \mathfrak{A} = 2\frac{(m_0+m_1)^2m_2^2}{(m_0+m_1+m_2)m_0m_1^2}\left(\frac{\Lambda_{0,1}+r}{\Lambda_{0,2}-r}\right)^2\frac{1+l_1}{1-l_2}\sqrt{1+\eta^2}, \\ \mathfrak{B} = \frac{m_1}{m_0}\mathfrak{A}, \\ \mathfrak{M} = G_{grav}^2\frac{(m_0+m_1)^2m_1m_2^3}{m_0+m_1+m_2}. \end{cases}$$

A raw theorem for a computer use

Theorem 1.9. *On the set $\mathcal{D}_{\Lambda,r,\rho,\rho',\lambda'_{max}}$, if the following hypotheses are true:*

(1) *There exists $t_i>0$ s.t. t_i verifies*
$\lambda'_{max} = t_i - (a_{1,i}a_{2,i}b_{3,i}(t_i)+a_{2,i}a_{3,i}(t_i)b_{1,i}+a_{1,i}a_{3,i}(t_i)b_{2,i}+b_{1,i}b_{2,i}b_{3,i}(t_i))$

(2) $l_1,l_2<1$

(3) $\frac{r}{\Lambda_0}\le\frac{3}{4}$,

(4) $\mathfrak{A}(\Lambda_{0,1},\Lambda_{0,2},r,\rho,\rho',t_1,t_2)<1$,

then the following inequality holds

$$\left\|\tilde{H}_{pert}\right\|_{\mathcal{D}_{\Lambda,r,\rho,\rho',\lambda'_{max}}} < \frac{3}{8}\frac{\mathfrak{M}}{(\Lambda_{0,2}-r)^2}\frac{1}{1-l_2}\left(\frac{\mathfrak{A}^2}{1-\mathfrak{A}}+\frac{m_1}{m_0}\frac{\mathfrak{A}^2}{1+\frac{m_1}{m_0}\mathfrak{A}}\right). \quad (1.43)$$

Proof. The proof is pretty much straightforward. We divide the series of the perturbation (1.4) into two terms, because of the expression $\sigma_n = \sigma_0^{n-1} + (-1)^n \sigma_1^{n-1}$. As for the leading coefficient of the Legendre polynomials that appears in the series, observe that

$$\frac{(2n)!}{2^n (n!)^2} \leq \frac{3}{8} 2^n.$$

We then gather all the terms into $\mathfrak{A}$ and $\mathfrak{B}$. The sum of these series are straightforward to compute. The hypotheses (1), (2) and (3) ensure that we can apply proposition 1.8, the fourth hypothesis ensuring the convergence of the series. □

Remark: Observe that we used the bound on the Legendre polynomial, starting from $n = 2$. This estimate is the only estimate where we did not make appear the analyticity widths as factors, and therefore, it is the less optimal bound in this sense. Though, to improve the computation, since the first Legendre polynomials are easy to compute, one can work on their expression directly, instead of the estimates. By doing this, one should be able to reduce the leading coefficient in front of the first terms, and hence the factor 3/8 in the formula of the theorem.

A simplified bound

Making further assumptions on the smallness of the analyticity widths allows us to simplify importantly the conditions and expressions of the previous terms.
Again, let $\Lambda_{0,1}, \Lambda_{0,2}, r, \rho, \rho' \in \mathbb{R}^{+\star}$ and $\Lambda_0 = \min_{i=1,2} \Lambda_{0,i}$.

Corollary 1.10. *Let* $0 < t \leq 0.1$*, under the assumptions*

$$100r < \Lambda_0, \qquad 100\sqrt{2}\rho' < \rho, \qquad \rho < \frac{25}{52}\sqrt{\Lambda_0}, \qquad 57\frac{(m_0+m_1)^2 m_2^2}{m_0^2 m_1^2}\left(\frac{\Lambda_{0,1}}{\Lambda_{0,2}}\right)^2 < 1,$$

define:

$$\lambda'_{max} = t - 1.02\frac{\rho}{\sqrt{\Lambda_0}}\sinh t - 3.7\frac{\rho'}{\sqrt{\Lambda_0}} - 2.4\frac{r}{\Lambda_0}.$$

If $\lambda'_{max} > 0$*, then the perturbation is analytic on the set* $\mathcal{D}_{\Lambda,r,\rho,\rho',\lambda'_{max}}$*, and the following inequation holds:*

$$\left\|\tilde{H}_{pert}\right\|_{\mathcal{D}_{\Lambda,r,\rho,\rho',\lambda'_{max}}} < 280 G_{grav}^2 \frac{(m_0+m_1)^5 m_2^7}{m_0^4 m_1^3}\left(\frac{\Lambda_{0,1}^4}{\Lambda_{0,2}^6}\right)\left(3 + 11\sinh t + 10\frac{r}{\Lambda_0} + 11\sqrt{\frac{\rho'}{\sqrt{\Lambda_0}}}\right).$$

With this corollary, we obtain the rate of decrease of the bound on the norm of the perturbation with respect to the analyticity widths. Observe that for a choice of width t for the angles, the bound decreases as $\sinh t$ (therefore as t) when it is small. Considering the variable λ'_{max}, when t is small, we have $\lambda'_{max} \sim (1 - 1.02\rho/\sqrt{\Lambda_0})t$, and $\sinh t \sim t$, therefore, the rate of decrease of the bound with the value λ'_{max} is

$$\frac{1}{1 - 1.02\frac{\rho}{\sqrt{\Lambda_0}}}\lambda'_{max},$$

which means that it is linear in λ'_{max}. Observe that the bound is as well linear in the width r, but not in ρ'. Instead, the square root of ρ' appears, which comes from the bound on the imaginary part of $\sqrt{\tilde{\xi}^2 + \tilde{\eta}^2}$. The mixing induced by the presence of the square root, and the non-trivial shape of the domain where $\tilde{\xi}^2 + \tilde{\eta}^2$ are defined, imply that the bound we found was in $\sqrt{\rho'}$.

Proof. The proof of this corollary requires a lot of simple computations (that we will not make entirely explicit), and a change of point of view considering the angle we consider. Indeed, instead of starting with an angle λ'_{max}, we start with a width in the eccentric longitude, which simplifies the computation.
First, we consider the expressions (1.30) of the a_i, b_i for $i = 1, 2, 3$ (here a_i stand for the temporary variables expressing the real part of some terms, not the semi-major axis of the planet). Regarding the a_i, the idea is to leave only a dependence in the variable Λ_0. For a_3, we use as well the fact that $\cosh t < 1.005$. Regarding the b_i, we want to leave a dependence in the small terms, either r or ρ' or $\sinh t$ (notice that we keep ρ as a possible high value term). We obtain:

$$\begin{cases} a_1 < 1.01/\sqrt{\Lambda_0} & b_1 < 1.02r/\Lambda_0^{3/2} \\ a_2 < 1.001 & b_2 < 1.04\rho'/\sqrt{\Lambda_0} + 0.53r/\Lambda_0 \\ a_3 < 1.5\sqrt{\Lambda_0} & b_3 < \rho \sinh t + 2.01\rho' \end{cases}$$

We exhibited the results to give an idea of the computations we were doing, though we will not display the next ones since it would be cumbersome. One can multiply the latter expressions to obtain a lower bound for λ'_{max}. When multiplying the terms b_1, b_2, b_3, we obtain a cubic dependence in the small variables, which does not interest us. Hence, we bound b_1 and b_2 using the assumptions, and we keep the expression of b_3; this choice is not unique.
Observe that the expression of λ'_{max} is of the form $y = x - a\sinh x - b$, with $a < 0.5$. This function is increasing from zero to some value, and then decreasing. Since $a < 0.5$, the point where it starts to decrease is further than 0.1, our limit for t. Therefore, the value of λ'_{max} corresponds to the first singular point of the complex Kepler equation. Hence, it defines an analytic diffeomorphism on this domain.
All the assumptions we made relied on the variable $\Lambda_0 = \min(\Lambda_{0,1}, \Lambda_{0,2})$, and λ'_{max} is defined using the value t and this specific $\Lambda_{0,i}$. Let us assume first that $\Lambda_{0,1} = \Lambda_0$ to simplify the discussion, which implies $t_1 = t$. Looking at the right side of the relation between t and λ'_{max}, if we let Λ_0 increase, then the right side increases as well. The solution t of this equation needs to decrease to compensate (λ'_{max} is fixed), and it gives $t_2 \leq t_1$. The bounds can therefore be computed using t_1 everywhere. The value of t in the theorem is hence associated to the body with the minimum value $\Lambda_{0,i}$.
Now one can bound every other term to obtain the result, keeping each time the smaller exponent in the analyticity width. Starting with e'_{max}, we see $\sqrt{\rho'}$ appears, and each time a factor ρ' appears, we decompose it into $\sqrt{\rho'} \times \sqrt{\rho'}$, bounding the first root and keeping the second one. Then, for x'_{max} and y'_{max}, we use the fact that $t_1, t_2 \leq t$ and $\Lambda_0 \leq \Lambda_{0,1}, \Lambda_{0,2}$ so as to find a bound independent on the bodies. The assumption on the variable ρ implies in fact that l_1 and l_2 are smaller than $1/2$, which simplifies again the computation.
Using these facts, we deduce a bound on $d_{i,max}$ that is again independent of i, and therefore, $d_{max} \leq d_{i,max}^2$. As for η, since $d_{max} > 1$, we bound it by

$$\eta \leq \frac{d_{max}}{2} + \frac{1}{2}.$$

Again, instead of considering $\mathfrak{A}$ and $\mathfrak{B}$, we simplify the expression using the fact that for $n \geq 2$, we have

$$\sigma_n \leq 1.$$

As well, we remove some dependence on the masse using $m_0 + m_1 < m_0 + m_1 + m_2$. The last two inequalities we will need are:

$$\frac{1 + l_1}{1 - l_2} < 3,$$
$$\sqrt{1 + x^2} < \sqrt{2} - 1 + x.$$

The series we obtain is the following:

$$\sum_{n=2}^{\infty}\left(6\frac{(m_0+m_1)^2m_2^2}{m_0^2m_1^2}\left(\frac{\Lambda_{0,1}}{\Lambda_{0,2}}\right)^2\left(3+11\sinh t+10\frac{r}{\Lambda_0}+11\sqrt{\frac{\rho'}{\sqrt{\Lambda_0}}}\right)^2\right)^n.$$

We can bound $\left(3+11\sinh t+10\frac{r}{\Lambda_0}+11\sqrt{\frac{\rho'}{\sqrt{\Lambda_0}}}\right)$ by 4.55. It converges if the main term is less than 1, hence if

$$28.5\frac{(m_0+m_1)^2m_2^2}{m_0^2m_1^2}\left(\frac{\Lambda_{0,1}}{\Lambda_{0,2}}\right)^2<1.$$

Requiring this term to be less than 1/2 (hence our last assumption) allows us a last simplification: $1/(1-x)<2$ for $x<1/2$. Finally, using the bound 4.55 on the term dependent on the analyticity width, one can remove the square. □

Analysis and comparison of the bound on the real domain

Let us do a short comparison between the real perturbation, that can be computed directly from its expression, and the two bounds given in the theorem and the corollary in the real case (*i.e.* for null analyticity width).

Consider the first theorem. We are looking at some point $(\Lambda_{0,1},\Lambda_{0,2})\in\mathbb{R}^2$, some eccentricity defined by $\rho>0$ and null analyticity width $r,\rho',t=0$. In this case, we obtain

$$e'_{max}=x'_{max}=y'_{max}=0.$$

Hence:

$$l_i=\frac{\rho}{\sqrt{\Lambda_{0,i}}},\qquad d_{i,max}=1,\qquad d_{max}=1,\qquad \eta=1,$$
$$\mathfrak{A}=2\sqrt{2}\frac{(m_0+m_1)^2m_2^2}{(m_0+m_1+m_2)m_0m_1^2}\frac{1+\frac{\rho}{\sqrt{\Lambda_{0,1}}}}{1-\frac{\rho}{\sqrt{\Lambda_{0,2}}}}\left(\frac{\Lambda_{0,1}}{\Lambda_{0,2}}\right)^2.$$

Finally,

$$\left\|\tilde{H}_{pert}\right\|_{\mathcal{D}_{\Lambda,0,\rho,0,0}}<\frac{3}{8}\frac{\mathfrak{M}}{\Lambda_{0,2}^2}\frac{1}{1-\rho/\sqrt{\Lambda_{0,2}}}\left(\frac{\mathfrak{A}^2}{1-\mathfrak{A}}+\frac{m_1}{m_0}\frac{\mathfrak{A}^2}{1+\frac{m_1}{m_0}\mathfrak{A}}\right).$$

An equivalent of this expression when the ratio $\Lambda_{0,1}/\Lambda_{0,2}$ goes to infinity is the following:

$$3G_{grav}^2\frac{\Lambda_{0,1}^4}{\Lambda_{0,2}^6}\frac{(m_0+m_1)^7m_2^7}{(m_0+m_1+m_2)^3m_0^3m_1^3}\frac{\left(1+\rho/\sqrt{\Lambda_{0,1}}\right)^2}{\left(1-\rho/\sqrt{\Lambda_{0,2}}\right)^3}\tag{1.44}$$

Regarding the corollary 1.10, under the two assumptions

$$\rho<25/52\sqrt{\Lambda_0},\qquad\text{and}\qquad 57\frac{(m_0+m_1)^2m_2^2}{m_0^2m_1^2}\left(\frac{\Lambda_{0,1}}{\Lambda_{0,2}}\right)^2<1,$$

we obtain directly:

$$\left\|\tilde{H}_{pert}\right\|_{\mathcal{D}_{\Lambda,0,\rho,0,0}}<820G_{grav}^2\frac{(m_0+m_1)^5m_2^7}{m_0^4m_1^3}\left(\frac{\Lambda_{0,1}^4}{\Lambda_{0,2}^6}\right).\tag{1.45}$$

To understand the limits in celestial mechanics given by these assumptions, we need to fix the masses of the bodies: we choose $m_1 = m_2 = 10^{-3} m_0$. To compare the different results obtained, we need the ratio $\Lambda_{0,1}/\Lambda_{0,2}$ to verify the assumption of the corollary, it has to verify:

$$\frac{\Lambda_{0,1}}{\Lambda_{0,2}} < \frac{1}{\sqrt{57}} \frac{1000}{1001}.$$

In terms of semi-major axes a_1 and a_2 (in what follows, we will not need the variables a_i and b_i, therefore no confusion can be made), it corresponds to the condition:

$$\frac{a_1}{a_2} < \frac{1}{57} \frac{1001}{1002}.$$

It is therefore very restrictive on the ratio of semi-major axes we can consider. Considering the eccentricities, notice first that we have $\Lambda_0 = \Lambda_{0,1}$. By definition:

$$e_1 < \sqrt{\frac{\rho^2}{\Lambda_{0,1}} - \frac{\rho^4}{4\Lambda_{0,1}^2}}.$$

Hence, a rough estimate on the maximal value of the eccentricity we can consider is $e_1 < 0.467$. As for e_2:

$$e_2 = \sqrt{\frac{\rho^2}{\Lambda_{0,1}} \frac{\Lambda_{0,1}}{\Lambda_{0,2}} - \frac{\rho^4}{4\Lambda_{0,1}^2} \left(\frac{\Lambda_{0,1}}{\Lambda_{0,2}}\right)^2}.$$

Given the assumptions, the value of e_2 is always increasing with the ratio of the variables $\Lambda_{0,i}$. Therefore, the maximal value for e_2 corresponds to the highest ratio of the $\Lambda_{0,i}$, and we have $e_2 < 0.174$. This time, compared to the examples in celestial mechanics, these values of eccentricities are not very restrictive.
Recall that this discussion of the maximal values we can consider comes from the corollary 1.10, theorem 1.9 requiring not such restrictive bounds.
It remains to determine the norm of the real perturbation. In this aim, consider equation (1.3). We can write

$$\begin{aligned} |H_{pert}| &\leq \max \left(\frac{G_{grav}\mu_1 m_2}{|Q_2|} \left(\frac{|Q_1|}{|Q_2|}\right)^2 \left(\frac{\sigma_0}{1 - \sigma_0 \frac{|Q_1|}{|Q_2|}} + \frac{\sigma_1}{1 + \sigma_1 \frac{|Q_1|}{|Q_2|}} \right) \right) \\ &\leq \max \left(\frac{G_{grav}\mu_1 m_2}{|Q_2|} \left(\frac{|Q_1|}{|Q_2|}\right)^2 \left(1 + \sigma_0\sigma_1 \left(\frac{|Q_1|}{|Q_2|}\right)^2 \frac{1}{1 - (\sigma_0 - \sigma_1)\frac{|Q_1|}{|Q_2|} - \sigma_0\sigma_1 \left(\frac{|Q_1|}{|Q_2|}\right)^2} \right) \right) \end{aligned}$$

With the last equation, we can see that the maximum is reached when $|Q_2|$ is minimal, and $|Q_1|$ maximal. Hence:

$$|H_{pert}| \leq G_{grav}\mu_1 m_2 \frac{(a_1(1+e_1))^2}{(a_2(1-e_2))^3} \left(\frac{\sigma_0}{1 - \sigma_0 \frac{a_1(1+e_1)}{a_2(1-e_2)}} + \frac{\sigma_1}{1 + \sigma_1 \frac{a_1(1+e_1)}{a_2(1-e_2)}} \right).$$

An equivalent when the ratio $\Lambda_{0,1}/\Lambda_{0,2}$ goes to infinity is then:

$$G_{grav}^2 \frac{\Lambda_{0,1}^4}{\Lambda_{0,2}^6} \frac{m_2^7}{m_0^3 m_1^3} \frac{(m_0+m_1)^7}{(m_0+m_1+m_2)^3} \frac{(1+e_{max})^2}{(1-e_{max})^3}. \tag{1.46}$$

We can now compare the results in the real case. First, let us talk about the equivalent, and to simplify the discussion, we can consider the eccentricities to be null. Comparing the

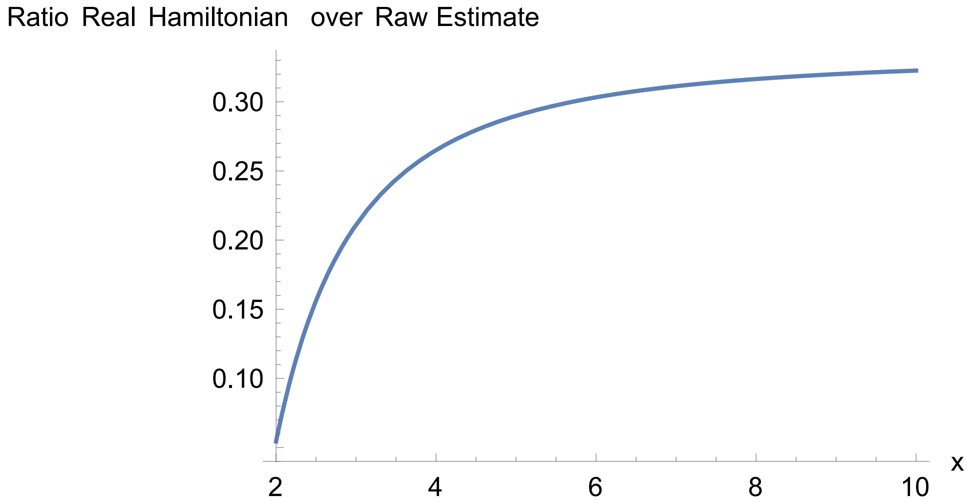

Figure 1.14: Ratio $|H_{pert}|/|H_{pert}|_{\mathcal{D}_{\Lambda_0,0,\rho,0,0}}$ as a function of $x = \Lambda_{0,2}/\Lambda_{0,1}$, for $\rho = \sqrt{\Lambda_1}/10$ ($e_1 \sim 0.1$), $m_1 = m_2 = 10^{-3}m_0$.

equivalent (1.44) given by the theorem and the real equivalent (1.46), we observe that only a factor 3 appears. This factor can be attributed to the difference between the estimates on the real Legendre polynomials and the complex one, which was far from being optimal. Indeed, the highest value of the Legendre polynomials in the real case is 1, whereas in our computation, we bounded the second Legendre polynomial by 3, the other ones not appearing since we consider the case $\Lambda_{0,1}/\Lambda_{0,2} \to +\infty$.

Comparing the equivalent (1.45) to the real case (1.46), we can see appearing a factor that is at least 820 (we did not compute the difference coming from the masses). Therefore, the corollary is far from being optimal. Its use is hence mostly to obtain an equivalent in term of the analyticity width.

Going back to the formula given by the theorem 1.9 in the case of null analyticity width, we plotted a graph 1.14 depending on the ratio of the semi-major axis to better understand the impact of the non-optimal bound of the Legendre polynomials.

Chapter 2

Some transformations of the Hamiltonian, and their constants

The Kepler part of the Hamiltonian of the plane planetary three-body problem suffers a degeneracy. Indeed, it depends on two coordinates, whereas there are four degrees of freedom. The Hamiltonian equations associated to it imply that only two angles, the mean anomalies, change over the time, the motion therefore takes place on a 2-torus. Physically speaking, the planets move on their ellipses, but the ellipses do not move over the time.

To apply the KAM theorem, we need an unperturbed Hamiltonian that depends only in the action of our action-angle variables and that is non-degenerate: hence we consider the secular Hamiltonian. This idea goes back to the work of Laplace and Lagrange, who developed the perturbation in the three-body problem to try to prove the stability of the solar system ([16], [17]). The secular Hamiltonian corresponds to the sum of the Kepler problem plus a part of the perturbation that depends only in the four action variables. Though, the frequencies associated to the angles are not of the same order: the two mean longitudes have frequencies associated to the Kepler problem part of the Hamiltonian, whereas the angles of the perihelion have their frequencies of the order of the perturbation. The first two are called the "fast angles", and the other two the "slow angles". When averaging the perturbation over the fast angles, the new term still depends on the two slow angles. It is possible to remove the dependency in one of these two angles using conservation of the angular momentum. It then remains a dependency of this term in one angle, which needs to be eliminated.

For this work, we consider the two first actions as parameters. Then, our object of study depends on two actions and one angle. It has been shown that it was possible to expand this object in the eccentricities, and that only terms of even positive order in these variables exist. Therefore, there exists an elliptic fixed point at zero for the two actions we are considering (related to the eccentricities). Removing the dependency in the last angle relies on two main operations: first we need to remove the dependency on the angle by doing a rotation on the variables, secondly, one has to use a Birkhoff normal form theorem to some order. The secular Hamiltonian will then consist in this development truncated at the order at which we removed the dependency in the angle, the highest order being considered as part of the perturbation. Putting the unperturbed Hamiltonian up to the first order requiring only a simple rotation, we will deal with it while computing explicitly the secular Hamiltonian. In this chapter, we will give a theorem of Birkhoff normal form with explicit constants.

Another requirement to the application of the KAM theorem is that the perturbation has small norm compared to the total Hamiltonian, or more precisely, that the perturbation in the frequencies induced by the interaction between the two planets is small enough compared to the frequencies of the unperturbed Hamiltonian. In the division we made before into secular and non-secular Hamiltonian, it is a priori not the case. Indeed, the frequencies associated to the slow angles are of the order of the perturbation, and we cannot apply the KAM theorem right

away. Before, it is necessary to make this perturbation smaller. This will be done in the second part of this chapter, and it relies on solving the classic cohomological equation in a restricted space (considering this time the third and fourth action as parameters), and on evaluating the remainders.
Most of the methods that we are using can be found in the two fundamental books [5] and [6].

2.1 Birkhoff Normal Form

This section will deal with the Birkhoff Normal Form (BNF), which corresponds to performing symplectic changes of variables around an elliptic point of equilibrium of a Hamiltonian, in order to put it into a normal form (yet to be defined) up to a certain order. As usual, in the statements we derive, we determine precisely the different constants involved in the proofs, such as the norm of the change of variables, the loss of analyticity, and the norm of the remainders. We will derive an explicit theorem, and several corollaries that could be useful when studying the secular Hamiltonian.
The general case is an old subject of study, and lots of references about this operation exist, for instance see [65], [32], and [30]. We will show a general theorem on BNF, and then apply it to obtain the constants we are looking for. In this second approach, we will largely follow the work of Bambusi [7], and try to be precise on the conditions of application.
Before getting into the details, let us take a look at the result we wish to show. Let $n \geq 1$ and a Hamiltonian $H : \mathbb{R}^{2n} \to \mathbb{R}$ (the coordinates are Cartesian here), such that the point $(p, q) = (0, 0)$ is an equilibrium point of the Hamiltonian equations associated to H. The Hamiltonian can be written:

$$
\begin{aligned}
H : &\mathcal{D} \subset \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \\
&(p, q) \longmapsto H_2(p, q) + H_3(p, q) + \dots \quad ,
\end{aligned} \tag{2.1}
$$

where H_m represents all the terms of degree m in p and q (in the case of the three-body problem, the Hamiltonian in Poincaré coordinates is already under this shape, and has only even terms). The Hamiltonian H_2 being only composed of terms of order two in p and q, the Hamiltonian equations associated to it are linear, and thus the motion can be easily deduced. In the elliptic case, one can write it as follows:

$$
H_2(p, q) = \sum_{i=1}^{n} \omega_i \left(\frac{p_i^2 + q_i^2}{2} \right) \tag{2.2}
$$

The Hamiltonian is said to be under normal form, if for $m \geq 3$ we have $\{H_2, H_m\} = 0$, where $\{\cdot, \cdot\}$ are the Poisson brackets. The BNF hence consists in doing several canonical changes of variables to put the Hamiltonian H under normal form if it is not already the case. In the calculations, it corresponds to removing some terms of order 3, then some of order 4, etc, though a finite number of time to ensure the convergence. The final expression of the Hamiltonian will be a specific normal form, called Birkhoff Normal Form up to some order.
The aim of this section is therefore to derive a theorem as follows:

Theorem 2.1. *Let H be a Hamiltonian of the form* (2.1) *with H_2 written as in* (2.2)*, and analytic on a domain $\mathcal{D}$. Then, under a non-resonance condition on the vector ω, and a condition on the norm of H, there exists a symplectic change of variable τ such that:*

1. $\tau : \ \mathcal{D}' \subset \mathcal{D} \to \mathcal{D}$

2. *$H \circ \tau$ is in BNF up to the order $N \geq 3$.*

Moreover, the set $\mathcal{D}'$ can be explicitly expressed as a function of the initial conditions.

From this theorem, we will deduce several corollaries, in different cases, that could be useful while studying the secular Hamiltonian.
Notice as well that we are considering an analytic Hamiltonian defined on subsets of $\mathbb{R}^n$. Nevertheless the theorem can be applied for any analytic Hamiltonian defined on some subset of $\mathbb{C}^n$, it requires only to take care of the domains of definition of the different variables.

2.1.1 Set of removable terms and first elimination

To introduce the general transformation and the theorem we want to prove, we will perform a transformation to remove one term. This will allow us to understand precisely the nature of the transformation as well as the removable terms, and therefore the precise form of the BNF. Before that, let us describe the Hamiltonian in a more suitable way for the incoming computation.
The first step to put the Hamiltonian under BNF is to perform a classic change of variables and to express the Hamiltonian in complex variables. Let:

$$x_i = \frac{1}{\sqrt{2}}(p_i + \imath q_i)$$
$$x_{i+n} = \frac{1}{\sqrt{2}}(q_i + \imath p_i)$$

Notice that we do not give information about the initial set, this will be done later. In these variables, we have the relation $x_i x_{i+n} = \frac{\imath}{2}(p_i^2 + q_i^2)$, and one can write $H_2(x) = \frac{1}{\imath}\sum_{i=1}^n \omega_i(x_i x_{i+n})$. For the terms of higher order, define the set

$$A_{m,n} = \left\{(i_1, ...i_{2n}) \in \mathbb{N}^{2n} \mid i_1 + ... + i_{2n} = m\right\},$$

and for $i \in A_{m,n}$, $x^i = x_1^{i_1}...x_{2n}^{i_{2n}}$. Therefore, for $m \geq 3$, the Hamiltonian H_m in these variables is written

$$H_m(x) = \sum_{i\in A_{m,n}} \omega_{m,i} x^i$$

Now define another set of exponents that will be proven useful later:

$$B_{m,n} = \left\{i = (i_1, ..., i_{2n}) \in \mathbb{N}^{2n} \mid i \in A_{m,n},\ (i_1, ..., i_n) \neq (i_{n+1}, ...i_{2n})\right\}$$

Let us divide the part of order m of the Hamiltonian in two different terms in the following way:

$$H_m^{\perp}(x) = \sum_{i\in A_{m,n}\setminus B_{m,n}} \omega_{m,i} x^i, \tag{2.3}$$
$$H_m^{\|}(x) = \sum_{i\in B_{m,n}} \omega_{m,i} x^i. \tag{2.4}$$

We have $H_m(x) = H_m^{\perp}(x) + H_m^{\|}(x)$, and

$$H(x) = H_2(x) + \sum_{i=3}^{\infty} H_i^{\perp}(x) + \sum_{i=3}^{\infty} H_i^{\|}(x)$$

The reason of this division will be clear after trying to remove one term of order more than 2 of the Hamiltonian. We can already say that the terms contained in $H_m^{\|}$ will be removable whereas the terms of $H_m^{\perp}$ will not be.
Another requirement for the theorem will be the non-resonance condition.

Definition 2.2. A vector $\omega \in \mathbb{R}^n$ is called non-resonant of order $k \in \mathbb{N}\setminus\{0\}$ if for every $j \in \mathbb{Z}^n$ with $0 < |j|_1 = \sum_{i=1}^n |j_i| \leq k$, we have $\langle\omega, j\rangle \neq 0$. ω is called non-resonant if it is non-resonant of order k for every $k \in \mathbb{N}\setminus\{0\}$.

With these definitions, we can state a first lemma.

Lemma 2.3. *Let $m > 2$ and H be a Hamiltonian of the form $H(x) = H_2(x) + \omega_{m,i}x^i$ for some $i \in A_{m,n}$, and $H_2(x) = \sum_{j=1}^n \omega_j x_j x_{j+n}$ with $\omega = (\omega_1, ..., \omega_n)$ non resonant of order m. If $i \in B_{m,n}$, there exists a symplectic transformation τ such that $H \circ \tau = H_2(x) + o(x^m)$.*

Proof. Let $P(x) = \beta x^i$ be a monomial of order m and let $\tau = \exp X_P$, where X_P is the vector field $J\nabla P$. Then $H \circ \tau = \exp L_P \cdot H$, with $L_P = \{P, \cdot\}$. One can then write:

$$H \circ \tau = (Id + \{P, \cdot\} + ...) \cdot (H_2 + \omega_{m,i}x^i) \tag{2.5}$$
$$= H_2 + \omega_{m,i}x_i + \{P, H_2\} + ... \tag{2.6}$$

It is easy to see that the terms not written in the last equality are terms of higher order than m in x, while the term $\{P, H_2\}$ is of order $(m-1) + (2-1) = m$. Hence, to prove the lemma, we need to choose β such that $\{P, H_2\} = -\omega_{m,i}x^i$. In this aim, define $x^{i^{(l)}} = x_1^{i_1}...x_l^{i_l-1}...x_{2n}^{i_{2n}}$. We can now compute:

$$\begin{aligned}
\{P, H_2\} &= \sum_{j=1}^n \left(\frac{\partial P}{x_j}\frac{\partial H_2}{x_{j+n}} - \frac{\partial P}{x_{j+n}}\frac{\partial H_2}{x_j} \right) \\
&= \sum_{j=1}^n \left(\beta x_{j+n}\omega_j i_{j+n} x^{i^{(j+n)}} - \beta x_j \omega_j i_j x^{i^{(j)}} \right) \\
&= \beta \left(\sum_{j=1}^n \omega_j (i_{j+n} - i_j) \right) x^i
\end{aligned}$$

Since ω is non resonant of order m, and $i_j \leq m$ for all j, and since $i \in B_{m,n}$, we have $\sum_{j=1}^n \omega_j(i_{j+n} - i_j) \neq 0$. For the next computation, let us write this relation with a simpler notation:

$$\langle\langle \omega, i \rangle\rangle = \sum_{j=1}^n \omega_j (i_{j+n} - i_j).$$

It is then sufficient to choose $\beta = -\frac{\omega_{m,i}}{\langle\langle \omega, i \rangle\rangle}$, and we have:

$$\{P, H_2\} = -\omega_{m,i}x^i \tag{2.7}$$

With this operation, we removed the term of order m as wanted, but as well we created an infinity of new terms of higher order. □

Let us talk about the two hypotheses of the lemma. First, the non-resonant condition is needed to avoid the case $\langle\langle \omega, i \rangle\rangle = 0$ for some $i \in A_{m,n}$. Indeed, in this case, one cannot remove a monomial $\omega_{m,i}x^i$. It is a classical hypothesis of non-resonance of the unperturbed Hamiltonian. The second hypothesis ($i \in B_{m,n}$) ensures that there exists some $j \in [\![1, n]\!]$ such that $i_{j+n} \neq i_j$. Without this condition, even with a non-resonant ω, it would be impossible to remove the wanted term. Indeed, one can remark that $\left\{H^\perp, H_2\right\} = 0$. The terms x^i for $i \in A_{m,n} \setminus B_{m,n}$ are therefore non-removable, they will be part of the Birkhoff normal form. One cannot expect to remove every term of some order greater than 2 for a generic Hamiltonian, the final form will contain the terms of $H_m^\perp$. Observe that for an odd m, the set $B_{m,n} = A_{m,n}$.
With this discussion, we can now consider the general case of the transformation, *i.e.* for a non-monomial perturbation of H_2.

Corollary 2.4. *Let $H(x) = H_2(x) + H_m(x) + H_r(x)$, with $m > 2$, H_m containing only terms of order m, $H_r(x) = o(x^m)$, and H_2 defined as before. Assume ω is non resonant of order m, then, formally, there exists a symplectic transformation τ such that $H \circ \tau = H_2(x) + H_m^\perp(x) + o(x^m)$.*

Proof. Write:

$$H_m = H_m(x) = H_m^{\perp}(x) + H_m^{\|}(x) = H_m^{\perp}(x) + \sum_{i\in B_{m,n}} \omega_{m,i}x^i.$$

Since ω is non-resonant of order m, one can write $P(x) = \sum_{i\in B_{m,n}} \beta_{m,i}x^i$, with $\beta_{m,i} = -\frac{\omega_{m,i}}{\langle\langle\omega,i\rangle\rangle}$. Each term of $H_m^{\|}$ is then removed by a term of P,thus the corollary. □

2.1.2 Proving the classical BNF Theorem

In the previous lemma and corollary, we defined a formal transformation, without taking care of the initial set of definition of the Hamiltonian. We will now try to quantify the size of the transformation necessary to put the Hamiltonian under normal form, or in other words, the loss of analyticity related to this change of variables.
The symplectic application τ is defined by the time-1 flow of some Hamiltonian vector field X_P (associated to the canonical symplectic form).

From the norm of the Hamiltonian vector field X_P, one can derive the distance of τ from the identity. We will therefore need to define a specific norm on the vector fields associated to some polynomial P.
Let P be an homogeneous polynomial of order m. Let $||\cdot||$ be a norm on $\mathbb{R}^n$ (since the norms are equivalent in $\mathbb{R}^n$, we do not choose a specific norm yet). There exists some constant C such that for all $x \in \mathbb{R}^{2n}$: $|P(x)| \leq C||x||^m$, the constant C depending obviously on P and on the choice of the norm. The vector field associated to P is the following: $X_P = J\nabla P$. Thus, it has $2n$ components, which are all homogeneous polynomials of order $m-1$. One can therefore define $C_1, ..., C_{2n}$ such that for $i \in [\![1, 2n]\!]$, $|X_{P,i}(x)| \leq C_i||x||^{m-1}$. Let us then define the following norm on Hamiltonian vector fields associated to a polynomial P:

$$||X_P|| = \inf\{C > 0 : \text{ for all } i \in (1, 2n), \forall x \in \mathbb{R}^{2n},\ |X_{P,i}|(x) \leq C||x||^{m-1}\} \tag{2.8}$$

This norm is well-defined for $m \in \mathbb{N} \setminus \{0\}$. We will need as well the following definitions:

$$B_R = \{x \in \mathbb{R}^n,\ ||x|| < R\};$$
$$\bar{t} = \bar{t}(R, \delta) = \inf_{x\in B_R} \left(sup\left\{t > 0 : \phi^s(x) \in B_{R+\delta}, \forall |s| < \bar{t}\ \right\}\right). \tag{2.9}$$

The latter definition will be called the minimum escape time $\bar{t}$ of ϕ^t from $B_{R+\delta}$ relatively to B_R. With this definition, we can now state the following lemma:

Lemma 2.5. *Let P be an homogeneous polynomial of order $k \geq 1$, to which we associate the vector field X_P. Let $\phi_X^t = \phi^t$ be the flow associated to this vector field, i.e. $\frac{d\phi^t}{dt}(x) = X_P(\phi^t(x))$ and $\phi^0 = Id$.*

The following inequality is verified:

$$\bar{t} \geq \frac{\delta}{||X_P||(R+\delta)^{m-1}}. \tag{2.10}$$

Equivalently, for all $|t| \leq \bar{t}$, we have

$$||\phi^t(x) - x|| \leq |t| \times ||X_P||(R+\delta)^{m-1} \tag{2.11}$$

This lemma shows that the flow associated to a polynomial vector field stays close to the identity for some time, related to the norm of the vector field P. Since we will want to consider the time-one map ϕ^1, we will require $\bar{t}$ to be greater than 1. Considering δ large enough, it will allow us to consider the time-one map, though the analyticity loss related to the change of variables will be great as well.

Proof. By definition of $\bar{t}$, there exists $\bar{x}$ such that we have $||\phi^{\bar{t}}(\bar{x})|| = R + \delta$. Assuming $\bar{t} < \frac{\delta}{||X_P||(R+\delta)^{k-1}}$, one can compute:

$$\begin{aligned} ||\phi^{\bar{t}}(\bar{x})|| = R + \delta \leq ||\bar{x}|| + ||\phi^{\bar{t}}(\bar{x}) - \bar{x}|| &\leq ||\bar{x}|| + \left|\left|\int_0^{\bar{t}} \frac{d\phi^s}{ds}(\bar{x})ds\right|\right| \\ &\leq R + \int_0^{\bar{t}} ||X_P(\phi^s(\bar{x}))||ds \leq R + \int_0^{\bar{t}} ||X_P||(||\phi^s(\bar{x})||)^{m-1}ds \\ &\leq R + \bar{t}||X_P||(R+\delta)^{m-1} \\ &< R + \delta \end{aligned}$$

There is a contradiction, hence $\bar{t} \geq \frac{\delta}{||X_P||(R+\delta)^{m-1}}$. From the previous computation, for all $|t| \leq \bar{t}$, for $x \in B_R$, the inequality $||\phi^t - Id|| \leq |t|||X_P||(R+\delta)^{m-1}$ holds. □

With the previous lemmas, we can state the Birkhoff Normal Form theorem, yet without precise estimates.

Theorem 2.6. *Birkhoff Normal Form:*
Let $H : \mathbb{R}^{2n} \to \mathbb{R}$ be a Hamiltonian, analytic on some ball, of the form $H(x) = H_2(x) + \sum_{m=3}^{\infty} H_m(x)$, with $H_2(x) = \sum_{i=1}^{n} \omega_i(x_i x_{i+n})$, and $H_m(x) = \sum_{i\in A_{m,n}} \omega_{m,i}x^i$.
Let $k \in \mathbb{N} \setminus \{0, 1, 2\}$ and assume $\omega = (\omega_1, ..., \omega_n)$ is non-resonant of order k. Then there exists a symplectic and analytic change of variables, close to the identity defined in a set close to the origin, such that the Hamiltonian H is under Birkhoff Normal Form up to the order k, that is to say:

1. *$H \circ \tau = H_2 + H^{\perp} + H_r$ with $\left\{H^{\perp}, H_2\right\} = 0$ and $H_r(x) = O(x^{k+1})$*

2. *$\tau(x) = x + O(x^2)$*

Proof. We proceed by induction on the terms H_m for $m \geq 3$ and we use the corollary 2.4. All the maps we will be dealing with are associated to a polynomial, and therefore are symplectic and analytic. For all the maps we want to apply to H, we need to verify that we can consider the time-1 map associated to the vector field.
Let us look at the step m of the induction. Assume that the Hamiltonian is under BNF up to the order $m-1$, and that it is analytic on some ball of size R_{m-1}. Let P_m be the polynomial associated to the terms of H_m, *i.e.* as defined in corollary 2.4, and $X_{P_m} = J\nabla P_m$. We consider the flow ϕ^t associated to the vector field X_{P_m}. Let $\delta < R_{m-1}$ and define $\bar{t}$ is the minimum escape time of ϕ^t from $B_{R_{m-1}}$ relatively to $B_{R_{m-1}-\delta}$. We know that we can consider the flow ϕ^t for all t such that $|t| < \bar{t}$. As done previously, we wish to consider the flow at time one, and therefore, we would like to have $\bar{t} \geq 1$ to be sure the time-1 map is well defined. We require:

$$1 \geq \frac{\delta}{||X_{P_m}||R_{m-1}^{m-1}}$$

Since $m - 1 \geq 2$, by re-scaling R_{m-1} (taking a smaller value), one can always find a suitable value for δ and R_{m-1} such that this inequality is verified. We will require slightly more here to finish the proof, by asking that R_{m-1} is so small that the value of δ is less than half of R_{m-1}. Then define $R_m = R_{m-1}/2 \leq R_{m-1} - \delta$, we then have $\phi^1(B_{R_m}) \subset B_{R_{m-1}}$.
Call $\phi_m = \phi^1$, $\tau_m = \tau_{m-1} \circ \phi_m$ and $\tau_2 = Id$. Then $H \circ \tau_m$ is under BNF up to the order m, and is analytic on B_{R_m}.
Now let us show inductively that $||\tau_m - Id|| = O(R_m^2)$, for $m \geq 2$, and $x \in B_{R_m}$. This is obvious

for $m = 2$. Let $m \geq 2$ and assume $||\tau_m - Id|| = O(R^2)$. Then, on $B_{R_{m+1}}$,

$$\begin{aligned}
||\tau_{m+1} - Id|| &= ||\tau_m \circ \phi_{m+1} - Id|| = ||\tau_m \circ \phi_{m+1} - \tau_m + \tau_m - Id|| \\
&\leq ||\tau_m \circ \phi_{m+1} - \tau_m|| + ||\tau_m - Id|| \\
&\leq \sup_{B_{R_m}} ||D\tau_m|| \times \sup_{B_{R_{m+1}}} ||\phi_{m+1} - Id|| + ||\tau_m - Id|| \\
&\leq C \times \delta + C'' R_m^2 \\
&\leq C \times C'(2R_{m+1})^m + 4C'' R_{m+1}^2 \\
||\tau_{m+1} - Id|| &\leq C_0 R_{m+1}^2
\end{aligned}$$

□

We did not express yet the explicit constants that appear in this change of variables, we showed that in a small ball close to the identity, we could choose a transformation such that we were losing half of the analyticity width at each step. Now we wish to be more precise and to know what initial analyticity width we can consider, at the cost of some additional hypotheses on the Hamiltonian.

2.1.3 Different explicit corollaries for different cases

In this section, we will give explicit corollaries of the BNF theorem. Though, there exists different interesting cases, and we will only consider two of those, that are interesting in our general problem. First, there is a difference between the case of a known Hamiltonian and the general case. Indeed, if the Hamiltonian is known, then we know exactly the polynomial P that generates our vector field, and we can compute explicitly its norm, whereas in the other case, we have to bound an "unknown" vector field. Secondly, we asked the frequency vector ω to verify a non-resonant condition of some order. This condition is pretty weak, and in the three-body problem, we will want this vector to be Diophantine, a much stronger condition. Therefore, we will use the Diophantine condition in the second corollary, even if the non-resonant condition is enough to make the theorem true.

Preliminaries: Choice of a norm and estimates on the vector field

Given the form of the generator of the transformation in the BNF theorem (a polynomial), we can deduce quite easily a bound on the norm of the vector field from the norm of the polynomial. We will then bound the transformation τ with the formula of the minimum escape time.
Consider a homogeneous polynomial P of order $m \geq 3$. It can be written

$$P(x) = \sum_{i \in A_{m,n}} \beta_{m,i} x^i.$$

We know that there exists C such that $|P(x)| \leq C||x||^m$, this constant depending on the norm we choose on $\mathbb{R}^{2n}$. To continue, we will fix this norm: assume we are now working with the sup norm $||x|| = ||x||_\infty = \sup_{i=1,\ldots,2n} |x_i|$, then $C = \sum_{i \in A_{m,n}} |\beta_{m,i}|$ verifies $|P(x)| \leq C||x||^m$. Indeed, we have:

$$\begin{aligned}
|P(x)| &\leq \sum_{i \in A_{m,n}} |\beta_{m,i}||x^i| \leq \sum_{i \in A_{m,n}} |\beta_{m,i}| \prod_{l=1}^{2n} |x_l|^{i_l} \\
&\leq \sum_{i \in A_{m,n}} |\beta_{m,i}| \prod_{i=1}^{2n} ||x||_\infty^{i_l} \leq \sum_{i \in A_{m,n}} |\beta_{m,i}| ||x||_\infty^m \\
&\leq C||x||_\infty^m
\end{aligned}$$

The vector field associated to the polynomial P is by definition $X_P = J\nabla P$. Consider now its different coordinates $X_{P,i}$ for $i \in [\![1, 2n]\!]$, it is a homogeneous polynomial of degree $deg(P) - 1$. Again, by definition, $\pm\frac{\partial P}{\partial x_{i\pm n}}(x)$. Hence, the norm of its different coordinates can be bounded with the help of the constant C defined previously: $|X_{P,i}(x)| \leq mC||x||^{m-1}$. Therefore, by definition of the norm on the vector fields associated to some polynomial:

$$||X_P|| \leq C = \sum_{i\in A_{m,n}} |\beta_{m,i}|. \tag{2.12}$$

Let $\beta_m = \sup_{i\in A_{m,n}} |\beta_{m,i}|$, then $C \leq Card(A_{m,n})\beta_m$. Yet the set $A_{m,n}$ is made of the n-tuples of natural integers such that their sum is equal to m. It is therefore a combination without repetition of some numbers verifying a condition on their sum, and the cardinal of this set is $\frac{(m+2n-1)!}{m!(2n-1)!}$. In practice, we know that the coefficients of P are equal to zero if their index are in the set $A_{m,n} \setminus B_{m,n}$, it is nevertheless easier to leave the computation of the cardinal of $B_{m,n}$ on the side, and to work with this number. Hence:

$$||X_P|| \leq \frac{(m+2n-1)!}{m!(2n-1)!} \sup_{i\in B_{m,n}} |\beta_{m,i}| \tag{2.13}$$

From this estimate on the vector field, we can now work on the size of the transformation ϕ. We know that at a step m, the formula of the transformation is given by $\exp X_P$, in other word it is associated with the time-1 map associated to the vector field X_P.
Consider a Hamiltonian H of the following form: $H(x) = H_2(x) + H_R(x)$, with $H_2(x) = \sum_{i=1}^{n} \omega_i(x_i x_{i+n})$, and with $K \geq 3$, $H_R(x) = \sum_{m=K}^{\infty} H_m(x)$, and
$H_m(x) = \sum_{i\in A_{m,n}} \omega_{m,i} x^i$. Assume that the Hamiltonian is analytic on a ball centered in the origin and of radius r. We wish to find a set $\mathcal{D}' = B_{r'}$ such that the application ϕ verifies $\phi(\mathcal{D}') \subset B_r$ (and that the Hamiltonian $H \circ \phi$ is under a BNF up to the order K).
With the construction we have done before, it is sufficient to verify that for $x \in B(0, r')$, the minimum escape time to leave the ball $B(0, r)$ is greater than 1. Recall its definition, $\bar{t} \geq \frac{r-r'}{||X_P||r^{K-1}}$, we are going to look for a value of $r - r'$ as a fraction of r, in other words $r - r' = a_K \times r$. We want:

$$\frac{a_K \times r}{||X_P||r^{K-1}} = \frac{a_K}{||X_P||r^{K-2}} > 1.$$

Hence, we let

$$a_K = ||X_P||r^{K-2}, \tag{2.14}$$

a_K is the ratio of loss of analyticity width at step K to put the Hamiltonian under a BNF up to the order K. Obviously, we will require that $0 < a_K < 1$, and therefore that either the norm of X_P is small, or that we are looking at a ball close to the origin.
With the definition of a_K, we have $\phi_K : B(0, r(1-a_K)) \to B(0, r)$, and therefore $H \circ \phi_K$ is well defined and analytic.

BNF up to a chosen order

Let a Hamiltonian H be of the following form:

$$H(x) = H_2(x) + H^{\perp}(x) + H_P(x) \tag{2.15}$$

with $H_2(x) = \sum_{i=1}^{n} \omega_i(x_i x_{i+n})$ and $H^{\perp}(x) = \sum_{m=3}^{\infty} \sum_{i\in A_{m,n}\setminus B_{m,n}} \omega_{m,i} x^i$. We will call the perturbation the following Hamiltonian: $K \geq 3$, $H_P(x) = \sum_{m=K}^{\infty} H_m(x)$, and $H_m(x) = \sum_{i\in B_{m,n}} \omega_{m,i} x^i$. We assume moreover that the frequency vector ω is non-resonant of order M, where $M > K$.
Assume H is analytic on $B_0 = B(0, r_0)$, and verifies $||H||_{r_0} \leq \varepsilon$.

The goal of this section is to determine r_f and a symplectic mapping $\tau : B(0, r_f) \to B(0, r_0)$ such that $H \circ \tau$ is under BNF up to the order M, and to finally have

$$\|H \circ \tau\|_{r_f} \leq \varepsilon.$$

Using the lemmas and corollary we derived previously, we know the specific form of the generating polynomials P_m necessary to put the Hamiltonian under BNF up to the order M.

First case: the Hamiltonian is known

We consider here that we know explicitly the Hamiltonian up to the order M, in other words we know the coefficients $\omega_{m,i}$ for $K \leq m \leq M$. At every steps of our scheme, we use a symplectic transformation removing the wanted terms, but we create terms of higher order, that are not straightforward to compute. Therefore, we will consider here that we want to put the Hamiltonian in BNF up to an order such that we did not create any new terms with the previous transformation. The minimal order of the terms appearing after the first step of the scheme comes from the term of order 2 in t of $\exp(tX_P)$: $\{\{H, X_P\}, X_P\}$, it is therefore of order $((\deg(P) - 1) + 2 - 1) + (\deg(P) - 1)) - 1 = 2K - 2$ in the variable x.

In this section, we will consider a scheme to put the Hamiltonian under BNF up to an order $M < 2K - 2$. Hence, the impact of the mapping we used at the order m for $K \leq m \leq M$ will be nonexistent while performing the scheme up to the order M. It could be possible to study higher orders more explicitly, but we choose not to do this, these scheme being enough to work with the plane planetary three-body problem.

Assume now that $M < 2K - 2$. For instance, this study could be applied to a Hamiltonian of the form $H(x) = H_2(x) + H_4(x) + H_5(x) + H_P(x)$, that we wish to put under BNF up to the order 5.

Let m $K \leq m \leq M$, define the following polynomial:

$$P_m(x) = \sum_{i \in B_{m,n}} \beta_{m,i} x^i, \qquad \beta_{m,i} = \frac{\omega_{m,i}}{\langle\langle \omega, i \rangle\rangle}. \tag{2.16}$$

The vector field associated to this polynomial is

$$\|X_{P_m}\| \leq \frac{(m + 2n - 1)!}{m!(2n - 1)!} \sup_{i \in B_{m,n}} \left| \frac{\omega_{m,i}}{\langle\langle \omega, i \rangle\rangle} \right|. \tag{2.17}$$

The ratio of analyticity width we are losing is

$$a_m = ||X_P|| r^{m-2} \leq \frac{(m + 2n - 1)!}{m!(2n - 1)!} \sup_{i \in B_{m,n}} \left| \frac{\omega_{m,i}}{\langle\langle \omega, i \rangle\rangle} \right| r^{m-2} \tag{2.18}$$

At step m, we define the time-1 map $\phi_m : B(0, r_m(1 - a_m)) \to B(0, r_m)$. Let us gather these results in order to obtain r_f as a function of r_0.

Define:

$$\gamma = \sup_{K \leq m \leq M} \left(\sup_{i \in B_{m,n}} \left| \frac{\omega_{m,i}}{\langle\langle \omega, i \rangle\rangle} \right| \right). \tag{2.19}$$

We can also bound (if we want) C_m^{m+2n-1} by C_M^{M+2n-1}. Hence,

$$\|X_{P_m}\| \leq \xi = C_M^{M+2n-1} \gamma. \tag{2.20}$$

Now define $r_{K-1} = r_0$. The mapping ϕ_m is such that if we let $r_m = r_{m-1} - \xi r_{m-1}^{m-1}$,we can define the final symplectic map τ by

$$\tau = \phi_K \circ \ldots \circ \phi_M : B(0, r_M) \to B(0, r_{K-1}).$$

This map puts the Hamiltonian under BNF up to order M.
Recall that we still have to give a condition on r_0 such that $0 < a_m < 1$. To ensure this condition, we will simplify the definition of r_f. First, for $m \geq K$, we have $r_m \leq r_0$. We can therefore define $r'_m = r'_{m-1} - \xi r_0^{m-1}$. We then have:

$$\begin{aligned} r'_m = r_0 - \xi \sum_{i=K-1}^{m-1} r_0^i &= r_0 - \xi r_0^{K-1} \frac{1 - r_0^{M-K}}{1 - r_0} \\ &\geq r_0 - \xi \frac{r_0^{K-1}}{1 - r_0}. \end{aligned}$$

Here we did not consider the case $r_0 = 1$, but the inequality in this case would be straightforward to obtain.
In the case r_0 is small enough, then $r_f = r_0 - \xi \frac{r_0^{K-1}}{1-r_0} > 0$, we have

$$\tau = \phi_K \circ ... \circ \phi_M : B(0, r_f) \to B(0, r_0).$$

This inequality ensures that for every step, the ratio of analyticity is smaller than one. Therefore, under the assumptions we made, we put the Hamiltonian under BNF up to the order $M < 2K-2$. The theorem we proved can be stated as follows:

Corollary 2.7. *Let a Hamiltonian H be of the form* (2.1.3)*, such that $\|H\|_{r_0} < \epsilon$, and let $K \leq M < 2K - 2$ with ω a non-resonant vector of order M. Let*

$$r_f = r_0 - \frac{r_0^{K-1}}{1 - r_0} \times C_M^{M+2n-1} \times \sup_{K \leq m \leq M} \left(\sup_{i \in B_{m,n}} \left| \frac{\omega_{m,i}}{\langle \langle \omega, i \rangle \rangle} \right| \right)$$

If $r_f > 0$ (therefore for r_0 sufficiently small), then there exists an analytic symplectic map τ such that:

1. $\tau : B(0, r_f) \to B(0, r_0)$,

2. $H \circ \tau$ *is under BNF up to the order M.*

Moreover, we have

$$\|H \circ \tau\|_{r_f} \leq \epsilon. \qquad (2.21)$$

Second Case: the Hamiltonian is only bounded
In this paragraph, we will be interested in a specific case that arises in the 3-body problem. Consider a Hamiltonian H that we do not know explicitly. We will consider that we only know that the frequency vector ω verifies a Diophantine condition, and that we can bound H over a set on which it is analytic. This problem is different than the case of a small perturbation of H_2 since the bound is not on the term $H - H_2$, but on the whole Hamiltonian H. In the plane planetary three-body problem, we are bounding the perturbation as a whole, but the Hamiltonian H_2 is part of the perturbation, and we do not have information (at least if we do not compute the explicit perturbation) on the terms of order higher than 6 in the eccentricities. Therefore, we will derive a theorem of BNF that is not optimal.
We are going to perform a scheme by induction on the order of the terms in x, as previously done, and at each step deduce the norm of the polynomial and the vector field we are using by Cauchy's inequality. Another possibility would have been to perform a scheme on all the terms (of any order) we want to remove to obtain a BNF, and to iterate it by evaluating the remainders of this operation; though, this scheme is more appropriate to a small perturbation around H_2, which is not our case. In our scheme, we will lose a lot of analyticity width, and

it is therefore not recommended to iterate it a large amount of times in an explicit application. Let $R > 0$, define by induction for $m > K > 2$ and some $\gamma, \tau, \epsilon > 0$:

$$\begin{cases} R_{(K)} = R \\ R_{(m)} = \frac{1}{2}\left(\frac{\gamma}{2\epsilon}\frac{(2n-1)!}{m^\tau(m+2n-1)!}R_{(m-1)}^m\right)^{\frac{1}{m-2}} \end{cases} \tag{2.22}$$

We have the following corollary:

Corollary 2.8. *Let a Hamiltonian H be of the form* (2.1.3)*, such that for some $R > 0$, $\|H\|_R < \epsilon$. Let $M \geq K$, and assume the frequency vector $\omega \in D(\gamma, \tau)$. If*

$$R^2 \leq \frac{2 \times K^\tau(K+2n-1)!}{\gamma(2n-1)!}\epsilon,$$

then there exists a analytic symplectic map τ such that:

1. *$\tau : B(0, R_{(M)}) \to B(0, R)$,*
2. *$H \circ \tau$ is under BNF up to the order M.*

Proof. Let us start with the first step, *i.e.* for $m = K$. From the formula of the Hamiltonian, we know that for $i \in A_{m,n}$, we have $\omega_{K,i} = \frac{\partial^K H}{\partial x_i}(0)$. Cauchy's inequality gives the following bound:

$$|\omega_{K,i}| \leq \frac{K! \times \|H\|_R}{R^K} \tag{2.23}$$

For the first step, we consider the Hamiltonian H on the ball B_R. In order to remove the terms of order K, we define as usual the following polynomial:

$$P_K(x) = \sum_{i \in B_{K,n}} \beta_{K,i} x^i, \qquad \beta_{K,i} = \frac{-1}{\langle\langle \omega, i \rangle\rangle}\frac{\partial^K H(0)}{\partial x_i}. \tag{2.24}$$

We can derive from this definition a bound on the vector field,

$$\begin{aligned} \|X_{P_K}\| &\leq \frac{(K+2n-1)!}{K!(2n-1)!}\frac{K!\epsilon}{R^K} \sup_{i \in B_{K,n}} \left|\frac{1}{\langle\langle \omega, i \rangle\rangle}\right| \\ &\leq \frac{(K+2n-1)!}{(2n-1)!}\frac{\epsilon}{R^K} \sup_{i \in B_{K,n}} \left|\frac{1}{\langle\langle \omega, i \rangle\rangle}\right|. \end{aligned}$$

Now, since ω is Diophantine, for $i \in A_{K,n}$, if we define $k_l = i_{l+n} - i_l$ for $l \in [\![1, n]\!]$, we have:

$$\langle\langle \omega, i \rangle\rangle = \omega \cdot k,$$

where $k = (k_1, ..., k_n)$. Since $i \in A_{K,n}$, we have $|k|_1 \leq K$. Hence,

$$\sup_{i \in B_{K,n}} \left|\frac{1}{\langle\langle \omega, i \rangle\rangle}\right| \leq \frac{K^\tau}{\gamma}$$

Finally:

$$\|X_{P_K}\| \leq \frac{K^\tau(K+2n-1)!}{(2n-1)!}\frac{\epsilon}{\gamma R^K}$$

Considering the time-1 map associated to this vector field ϕ_K, we wish to find two balls of size R' and R'' such that $R'' < R' < R$ and $\phi_K(B_{R''}) \subset B_{R'}$. The new Hamiltonian $H \circ \phi_K$ will then

be analytic on the ball R''.
With lemma 2.5 applied to the time 1, the condition on R' and R'' is:

$$1 \geq \frac{R' - R''}{||X_{P_K}||(R')^{K-1}}.$$

To simplify, we will fix $R'' = R'/2$, and as well that the inequality is in fact an equality (we do not need more). We then have:

$$\begin{aligned} R'^{K-2} &= \frac{\gamma}{2\epsilon}\frac{(2n-1)!}{K^{\tau}(K+2n-1)!}R^K \\ &= (2R_{(K+1)})^{K-2}, \end{aligned}$$

and therefore $R'' = R_{K+1}$.
The condition $R' \leq R$ still requires verification:

$$\begin{aligned} R' \leq R &\Leftrightarrow \frac{(R')^{K-2}}{R^{K-2}} \leq 1 \\ &\Leftrightarrow R^2 \frac{\gamma}{2\epsilon}\frac{(2n-1)!}{K^{\tau}(K+2n-1)!} \leq 1 \\ &\Leftrightarrow R^2 \leq \frac{2\epsilon}{\gamma}\frac{K^{\tau}(K+2n-1)!}{(2n-1)!}, \end{aligned}$$

which corresponds to the assumption of the corollary. Therefore, the first step is proved, and we have $\phi_K : B_{R_{K+1}} \to B_{R_K}$ with $H \circ \phi_K$ under BNF up to the order $K+1$.

Assume now that we built the application $\tau_m = \phi_m \circ ... \circ \phi_K : B_{R_{m+1}} \to B_{R_m}$ such that $H \circ \tau_m$ is under BNF up to the order $m+1 < M$. There is no much change in the scheme we applied for the first step, though there still exists some subtleties.
First, concerning the coefficient $\omega_{m,i}$ for $i \in A_{m,n}$, the bound we can give by Cauchy's inequality is now:

$$\begin{aligned} |\omega_{m,i}| &\leq \frac{m!}{R^m_{(m+1)}} \left\|H \circ \tau_m\right\|_{R_{(m+1)}} \\ &\leq \frac{m!}{R^m_{(m+1)}} \left\|H\right\|_R \end{aligned}$$

Then, replacing in the computations done in the first step, K by the index of this step $m+1$, we obtain a new time-1 map $\phi_{m+1} : B_{R_{m+2}} \to B_{R_{m+1}}$ without difficulties. The last inequality we have to check is that $2R_{m+2} \leq R_{m+1}$. This is the case if

$$R^2_{m+1} \leq \frac{2\epsilon}{\gamma}\frac{(m+1)^{\tau}(m+2n)!}{(2n-1)!}.$$

Yet, the right term increases with the value of m, and the left term is decreasing with it, therefore it is sufficient to check it at step one only. This finishes the proof of the corollary. □

Let us make some remarks on the corollary and its proof.
Notice that when using Cauchy's inequality, we bounded our term $\omega_{m,i}$ using $\|H\|_R$ even though we could have bounded it using $\|H\|_{R_m}$. Nevertheless, we do not know anything about this specific value. Worse, since the Hamiltonian is bounded as a whole (with H_2) on some ball, decreasing the size of the ball on which we bound the Hamiltonian will, for small R, makes this bound decrease as R^2, due to the term H_2. This quadratic decrease will not be enough to compensate the loss of analyticity width induced by the preceding steps, and even if it could

improve the result by computing at each steps the new size of the Hamiltonian in the wanted ball, the difference would not be much. In the 3-body problem it is even worse since the perturbation is of the form $H = C + H_2 + ...$ with C a non-zero constant, and therefore decreasing the size of the ball has even less impact on the bound of H.
In the case of a frequency vector that is only non-resonant up to the order M, the corollary holds. Though it is easier here to work with the Diophantine condition in order to simplify the computations.

With the last two corollaries, we can put the Hamiltonian under BNF up to any order, by losing some analyticity width, with formulas that are explicit. This work will be enough in our case, and we will not derive other BNF related theorems.

2.2 Secular Hamiltonian, and remainder of the perturbation

In this section, we are concerned with the part of the perturbation that depends on the fast angles, and we wish to "push" its norm to an even smaller value, so as to make the induced perturbation in frequencies much smaller than the frequencies of the slow angles (that are of the order of the perturbation).
We are going to take advantage of the Kepler problem part and its high frequencies, and the analyticity width of the perturbation to deal with this problem. This will be done in a space of dimension n, and we consider a Hamiltonian of the following form:

$$H(I,\theta) \ = H_0(I) + \epsilon H_1(I,\theta), \qquad \epsilon \ll 1$$

First, we divide the Hamiltonian H_1 in two parts, as described to obtain what will become our secular Hamiltonian:

$$\begin{aligned} H_1(I,\theta) &= \bar{H}_1(I) + \tilde{H}_1(I,\theta) \\ \bar{H}_1(I) \quad &= \frac{1}{(2\pi)^n}\int_{\mathbb{T}^n} H_1(I,\theta)d\theta \end{aligned}$$

The aim of this chapter is to find a theorem stating:

Theorem 2.9. *Let a Hamiltonian $H : \mathbb{C}^n \times \mathbb{T}^n_{\mathbb{C}}$ be defined as above and analytic on a non-empty set $\mathcal{D}$. Then if ϵ is sufficiently small, and under non-resonance condition on the Hamiltonian $H_0 + \bar{H}_1$, there exists a symplectomorphism $\varphi : \mathcal{D}' \subset \mathcal{D} \to \mathcal{D}$, such that:*

$$H \circ \varphi = H_0 + \epsilon \bar{H}_1 + \epsilon H_2$$

with $\sup_{\mathcal{D}'} |H_2| \ll \sup_{\mathcal{D}} |H_1|$.

Moreover, we want this theorem to be explicit in the domains $\mathcal{D}$ and $\mathcal{D}'$, on the hypothesis on ϵ, and on the norm of H_2.
This transformation is very close to the KAM step in its structure. Though, we will not adapt the analyticity width and the norm of the remainder of this operation that would allow us to do an infinite number of steps. Therefore, the result will be closer to a step of a Nekhoroshev theorem. Since we will later deal with the KAM theorem, we will do this step slowly, sometimes naively, so as to answer the questions associated with this transformation. The approach we will take here will be the vector fields one, and not directly the Hamiltonian one.

2.2.1 Introduction to the theorem

The main scheme

Let X be a vector field. The flow associated to this vector field is determined by

$$\begin{cases} \dfrac{d\varphi_X^t}{dt}(x) = X(\varphi_X^t(x)) \\ \\ \varphi_X^0 = Id \end{cases}$$

for $x = (I, \theta)$. In the case of a Hamiltonian vector field, the flow φ_X^t is a symplectic transformation. Now consider a Hamiltonian as defined before, analytic on the non-empty set $\mathcal{D}$, and the following unperturbed Hamiltonian:

$$H_{0,1} = H_0 + \bar{H}_1.$$

This Hamiltonian depends only in the actions. For any x on the pre-image of the set $\mathcal{D}$ under the flow φ_X^t, this Hamiltonian takes the form:

$$\begin{aligned} H_{0,1} \circ \varphi_X^t(x) &= H_{0,1}(\varphi_X^0(x)) + t\,\frac{d}{dt}(H_{0,1}(\varphi_X^t(x))\Big|_{t=0} + \int_0^t \frac{d^2}{ds^2}(H_{0,1}(\varphi_X^s(x)))(t-s)ds \\ &= H_{0,1}(x) + t(H'_{0,1} \cdot X)(x) + \int_0^t ((H'_{0,1} \cdot X)' \cdot X) \circ \varphi_X^s(x)(t-s)ds \\ &= H_{0,1}(x) + t(H'_{0,1} \cdot X)(x) + R_2(x,t) \end{aligned}$$

The same computation for the perturbation $\tilde{H}_1$ gives:

$$\begin{aligned} \tilde{H}_1 \circ \varphi_X^t(x) &= \tilde{H}_1(x) + \int_0^t (\tilde{H}'_1 \cdot X) \circ \varphi_X^s(x)ds \\ &= \tilde{H}_1(x) + \tilde{R}_1(x,t) \end{aligned}$$

To succeed demonstrating our theorem, we ask that on a set $\mathcal{D}'$ such that $\varphi_X^t(\mathcal{D}') \subset \mathcal{D}$, our vector field X verifies:

$$tH'_{0,1}(x).X(x) = -\epsilon\tilde{H}_1(x), \tag{2.25}$$

and that the remainders:

$$R_2(x,t) + \epsilon\tilde{R}_1(x,t) \ll \epsilon. \tag{2.26}$$

Under this conditions, the theorem 2.9 holds.
First, observe that the equation (2.25) fixes the vector field X. This vector field being completely determined, we have to verify the conditions on the existence of a set $\mathcal{D}'$ to be non-empty (and possibly of a size and a shape suitable to work with), and the norm of the remainders.
To simplify our calculation, we will consider the flow at time $t = \epsilon$, and therefore we will have to check the conditions on $\mathcal{D}'$ at this special time. The cohomological equation on this set will then be:

$$H'_0(x).X(x) = -\tilde{H}_1(x) \tag{2.27}$$

Besides, for this special flow, it is possible to simplify the remainder as follows:

$$\begin{aligned} R_f(x,\epsilon) &= R_2(x,\epsilon) + \epsilon\tilde{R}_1(x,\epsilon) \\ &= \int_0^\epsilon \Big((\epsilon - s)(H'_{0,1} \cdot X)' \cdot X + \epsilon\tilde{H}'_1.X\Big) \circ \varphi_X^s(x)ds \\ &= \int_0^\epsilon \Big(\epsilon(H'_{0,1} \cdot X + \tilde{H}_1)' \cdot X + s(H'_{0,1} \cdot X)' \cdot X\Big) \circ \varphi_X^s(x)ds \\ &= \int_0^\epsilon s\Big((H'_{0,1} \cdot X)' \cdot X\Big) \circ \varphi_X^s(x)ds \end{aligned} \tag{2.28}$$

There is then only one term to estimate in the remainder.

Under all this conditions, one can see that the vector field X is of order 1 by the equation (2.27). Thus, for the remainder R_f, the presence of the variable s of integration under the integral, and the fact that we are integrating from 0 to ϵ implies that R_f is of order 2 in ϵ. To insure this stays true, one only have to control the term $(H'_{0,1} \cdot X)' \cdot X$ on the set $\mathcal{D}$.

Definition of the initial sets and of the norms

Recall that the coordinates we are using in this part are action-angle coordinates. Moreover, the theorems we are going to apply will not be isotropic in actions and angles, in the sense that we want to separate in our work the actions and the angles, and therefore their associated analyticity width. The precise estimates of the analyticity width of the perturbation among the different variables justifies this choice. Besides, it allows us to not lose artificially some analyticity width while computing an operation on either the action or the angles.
The real coordinates we are considering before the analytic continuation will be of the form $x = (I, \theta) \in \mathbb{R}^n \times \mathbb{T}^n$. More precisely, letting B be an open subset of $\mathbb{R}^n$, the initial set of real variables will be $B \times \mathbb{T}^n$. Let us define the new set of complex coordinates after the analytic continuation of the analytic Hamiltonian. First, we want this set to be included in the set $\mathbf{T}^n_{\mathbb{C}} = \mathbb{C}^n \times \mathbb{T}^n_{\mathbb{C}}$. Secondly, we need to define some distances to the real initial set: for $x = (I, \theta) \in \mathbb{C}^n \times \mathbf{T}^n_{\mathbb{C}}$,

$$\begin{aligned} ||x||^I = d_I(x, B) &= \max_{i \in [\![1,n]\!]} \left(\inf_{y \in B} |I_i - y_i| \right) \\ ||x||^\theta = d_\theta(x, \mathbb{T}^n) &= \max_{i \in [\![n+1,2n]\!]} |\Im x_i| = \max_{i \in [\![1,n]\!]} |\Im \theta_i| \end{aligned}$$

Observe that in the case $B = \{0\}$ the first distance is a semi-norm. In order to be more general, we will keep the set B in our definitions. Define as well the following distance which will be useful: $||x||_\infty = \max(||x||^I, ||x||^\theta)$. We have the equivalence:

$$||x||_\infty = 0 \Leftrightarrow x \in B \times \mathbb{T}^n.$$

With this definition, the polydisc around the set $B \times \mathbb{T}^n$ is expressed, for $r, s > 0$, by:

$$B(r, s) = \left\{ x \in \mathbb{C}^n \times \mathbf{T}^n_{\mathbb{C}}, ||x||^I \leq r, ||x||^\theta \leq s \right\}. \tag{2.29}$$

This notation does not make appear the importance of the set B, though it is necessary to recall that it is completely dependent on it.
Consider now a function $f : B(r, s) \to \mathbb{C}$, define the supremum norm of f over the set $B(r, s)$ with the notation

$$\|f\|_{r,s} = \|f\|_{B(r,s)} = \sup_{x \in B(r,s)} |f(x)|.$$

The tangent space of $\mathbf{T}^n_{\mathbb{C}}$ can be identified with the space $\mathbb{C}^n \times \mathbb{C}^n$. For vectors in the latter space, it is possible to define again two semi-norms related to the tangent space of the actions in one hand, and to the one of the angles on the other hand:

$$\begin{aligned} ||v||^I &= \max_{i \in [\![1,n]\!]} |v_i| \\ ||v||^\theta &= \max_{i \in [\![n+1,2n]\!]} |v_i| \\ ||v||_\infty &= \max(||v||^I, ||v||^\theta) \end{aligned}$$

Applying a linear form $L \in T^\star \mathbf{T}^n_{\mathbb{C}}$ to a vector v of $\mathbb{C}^n \times \mathbb{C}^n$, we have the following inequality:

$$\begin{aligned}|L.v| &= |L_I.v_I + L_\theta.v_\theta|\\ &\leq |L_I.v_I| + |L_\theta.v_\theta|\end{aligned}$$

Considering separately the tangent space on the actions and the one on the angle, we wish to define two norms L_1 on linear forms in the following way:

$$\begin{aligned}|L.v| &\leq \left(\sum_{i=1}^{n} |L_{I,i}|\right) ||v||^I + \left(\sum_{i=1}^{n} |L_{\theta,i}|\right) ||v||^\theta\\ &\leq ||L||_1^I ||x||^I + ||L||_1^\theta ||x||^\theta\end{aligned}$$

Thus, we defined the two following semi-norms on the cotangent space $T^\star \mathbf{T}^n_{\mathbb{C}}$:

$$\begin{aligned}\forall L \in &\mathcal{L}(\mathbb{C}^n, \mathbb{C}) : \quad L = L_1 e_1^\star + ... + L_{2n} e_{2n}^\star\\ ||L||_1^I &= \sum_{i=1}^{n} |L_i|\\ ||L||_1^\theta &= \sum_{i=n+1}^{2n} \sup_{\theta\in\mathbb{T}} |L_i|,\end{aligned}$$

where $(e_1^\star, ..., e_{2n}^\star)$ is a dual base of $\mathbb{C}^n$.
We will also need norms on the vector fields to compute our theorem. Let us simply define

$$\begin{aligned}||X||_D^I &= \sup_{x\in D} \left(\max_{i\in[\![1,n]\!]} |X_i(x)| \right),\\ ||X||_D^\theta &= \sup_{x\in D} \left(\max_{i\in[\![n+1,2n]\!]} |X_i(x)| \right).\end{aligned}$$

Eventually, we will define some more objects, but those were the most important definitions we required to continue.

2.2.2 Proof of the theorem

Solving the cohomological equation

In this part, we will compute the vector field X as a function of the expression of the two Hamiltonians $H_{0,1}$ and $\tilde{H}_1$. In this aim, we will solve the equation (2.27) formally, then naively, and finally, under some conditions, we will give the expression of the vector field and the new domain of analyticity in which X is well defined.

Formal resolution of the cohomological equation
Let X be a Hamiltonian vector field associated to the Hamiltonian G. It can be described as follows:

$$X = J\nabla G, \tag{2.30}$$

where $G : \mathbb{R}^n \times \mathbb{T}^n \to \mathbb{R}$, and

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Assume that G is analytic on some set $B(r, s)$, with $r, s > 0$, and where H is also analytic. We can write $X = (X_1, X_2)$ with $X_1, X_2 \in \mathbb{R}^n$. By definition, we have the equations:

$$\begin{cases} X_1 = \partial_\theta G \\ X_2 = -\partial_I G \end{cases} \tag{2.31}$$

The equation (2.27) we wish to solve is:

$$H'_{0,1}(x) \cdot X(x) = -\tilde{H}_1(x), \ \forall x \in B \times \mathbb{T}^n$$

Thus, using that $H_{0,1}$ does not depend on the angles. It can be expressed in the following way:

$$\partial_I H_{0,1}(x) \cdot \partial_\theta G(x) = -\tilde{H}_1(x), \ \forall x \in B \times \mathbb{T}^n \tag{2.32}$$

Using the fact that the Hamiltonian is analytic and 2π-periodic on the angle, we can develop its different parts in Fourier series:

$$\tilde{H}_1(x) = \sum_{k\in\mathbb{Z}^n\setminus\{0\}} h_k(I) e^{\imath k\cdot\theta}$$
$$G(x) = \sum_{k\in\mathbb{Z}^n\setminus\{0\}} g_k(I) e^{\imath k\cdot\theta}$$

Here we took G with zero mean, this choice does not affect in any way the computation. Let us continue with the Fourier coefficients:

$$\begin{aligned}
\partial_I H_{0,1}(x) \cdot \partial_\theta G(x) &= \sum_{i=1}^n \partial_{I_i} H_{0,1}(x) \partial_{\theta_i} G(x) = \sum_{i=1}^n \omega_i(I) \left(\sum_{k\in\mathbb{Z}^n\setminus\{0\}} k_i g_k(I) e^{\imath k\cdot\theta} \right) \\
&= \sum_{k\in\mathbb{Z}^n\setminus\{0\}} g_k(I) e^{\imath k\cdot\theta} \sum_{i=1}^n \omega_i(I) k_i \\
&= \sum_{k\in\mathbb{Z}^n\setminus\{0\}} \omega(I)\cdot k \times g_k(I) e^{\imath k\cdot\theta} \\
&= -\sum_{k\in\mathbb{Z}^n\setminus\{0\}} h_k(I) e^{\imath k\cdot\theta},
\end{aligned}$$

where $\omega(I)$ is a frequency vector in a point x associated to the secular Hamiltonian. We can now identify each terms of the Fourier series, which gives the formulas on the g_k, for $k \in \mathbb{Z}^n \setminus \{0\}$:

$$\omega(I)\cdot k \times g_k(I) = -h_k(I) \tag{2.33}$$

Formally (so without considering the convergence of the series, or the fact that some terms could be infinite), the solution of the cohomological equation is the following:

$$G(I,\theta) = -\sum_{k\in\mathbb{Z}^n\setminus\{0\}} \frac{h_k(I)}{\omega(I)\cdot k} \exp(\imath k\cdot\theta).$$

G is unique up to a constant.

A naive approach to the convergence of G

Clearly, G can be ill-defined: several problems can occur. The first and most obvious one is the case of a rationally dependent vector $\omega(I)$. Indeed, in this case there exists $k \in \mathbb{Z} \setminus \{0\}$ such that $\omega(I)\cdot k = 0$ and therefore g_k is not defined. In dimension two ($n = 2$) for instance, it corresponds to a rational ratio of frequencies. That is a first problem we need to take care of. Assuming that the vector $\omega(I)$ is not rationally dependent for some I, we have the existence of $g_k(I)$ for all $k \neq 0$. However, it does not imply that the Fourier series of G converges. Indeed, we are confronted to the problem of small denominators. The larger K is, the easier it is to find an element $k \in \mathbb{Z} \setminus \{0\}$ such that $|k|_1 \leq K$ and $\omega(I)\cdot k$ is really close to zero, making the estimations on the norm of g_k deteriorate quickly with $|k|_1$. On the other hand, since H_1 is analytic, we know that the coefficients h_k are decreasing exponentially with $|k|_1$. Indeed, if H_1

is analytic on the set $B(r,s)$, the coefficients satisfy $|h_k(x)| \leq C\exp(-|k|_1 s)$. Therefore, there is a competition between the smallness of these two terms, and to get the convergence of the previous Fourier series, it is sufficient to ask that the factors $1/|\omega(I)\cdot k|$ do not increase more than polynomially with $|k|_1$. In other words, we ask the vector $\omega(I)$ to be Diophantine for some constants (γ,τ). Under this condition, G is well defined.
For a fixed frequency vector ω, different versions of theorems to solve the cohomological equation exists (the Kolmogorov's version can be found in [35, 36, 37]). Here we considered analytic functions, and we will therefore use the theorem of [64], which gives an optimal estimate on the norm of the solution for generic analytic functions. We want to consider functions of n angles that are analytic on some domain of width s, and of zero mean. Define:

$$\begin{aligned}
\mathbb{T}^n_s &= \{\theta \in \mathbb{T}^n_{\mathbb{C}}, \forall i \in [\![1,n]\!] : |\Im\theta_i| < s\} \\
\mathcal{A}^s &= \{f : \mathbb{T}^n_s \to \mathbb{C},\ \text{f } \mathbb{C}\text{ -analytic}\} \\
\mathcal{A}^s_0 &= \{f \in \mathcal{A}^s,\ \text{s.t.} \int_{\mathbb{T}^n} f = 0\}.
\end{aligned}$$

Writing $|f|_s = sup_{\mathbb{T}^n_s}|f|$, we have the following theorem:

Theorem 2.10 (Rüssmann). *Let $\omega \in D_{\gamma,\tau}$ a Diophantine vector, and $g \in \mathcal{A}^s_0$. Then the equation*

$$\partial_\omega f = g \tag{2.34}$$

has a unique solution f in $\bigcup_{0<\sigma<s} \mathcal{A}^{s-\sigma}_0$, and we have the following bound on the norm of f for $0<\sigma<s$:

$$|f|_{s-\sigma} \leq \frac{C_0}{\gamma\sigma^\tau}|g|_s,$$

where $C_0 = \frac{3\pi}{2}6^{\frac{n}{2}}\frac{\sqrt{\tau\Gamma(2\tau)}}{2^\tau}$.

This theorem states that if the frequency vector is Diophantine, then there exists a solution to the cohomological equation. In the case $\omega(I)$ was a constant Diophantine vector, one could then find a solution for G for each I, and by analyticity, a solution on a whole set $B(r',s')$ for some r' and s'.
A new difficulty occurs, coming from the fact that ω is not a constant function of I. Moreover, the set $D(\gamma,\tau)$ is a Cantor set, and ω is continuous. Therefore, G defined by the equation (2.27) can never be solution on some set $B(r',s')$ if $r'>0$.
However, the exponential decrease of the norm of the h_k with $|k|_1$ makes the sum of the terms with large $|k|_1$ small. Salvation will come from the fact that we can truncate our Hamiltonian H_1 in two parts, one with relatively small $|k|_1$, and one with the terms h_k for $|k|_1 > K$, where K as to be fixed. In this case, we do not require to verify a Diophantine condition for $\omega(I)$, but only a non-resonant condition up to the order K.

Truncation and solution of a reduced cohomological equation
Define the truncation of the function H_1 of order $K>0$:

$$\tilde{H}^K_1(I,\theta) = \sum_{k\in\mathbb{Z}^n\setminus\{0\}, |k|_1\leq K} h_k(I)\exp(\imath k\cdot\theta) \tag{2.35}$$

We ask G to verify the reduced cohomological equation:

$$\tilde{H}^K_1(x) + \partial_I H_{0,1}(x)\cdot\partial_\theta G(x) = 0 \tag{2.36}$$

The proof of the theorem 2.10 uses the fact that the frequency vector $\omega(I) = H'_{0,1}(I)$ is Diophantine to control the products $\omega\cdot k$ for $k\in\mathbb{Z}^n\setminus\{0\}$. Though in our case, the proof is much easier since our unknown G has no Fourier coefficient of order k such that $|k|_1 > K$. The previous theorem still holds, and we will use the corollary:

Corollary 2.11. *Let $K \in \mathbb{N}^\star$, $g \in \mathcal{A}_0^s$ such that for all $|k|_1 > K$, the Fourier coefficient of order k of g is zero, and let $\omega \in \mathbb{C}^n$, such that:*

$$|\omega \cdot k| \geq \frac{\gamma}{|k|_1^\tau}, \qquad \forall |k|_1 \leq K.$$

Then the equation

$$\partial_\omega f = g$$

has a unique solution f in $\bigcup_{0<\sigma<s} \mathcal{A}_0^{s-\sigma}$, and we have the following bound on the norm of f for $0 < \sigma < s$:

$$|f|_{s-\sigma} \leq \frac{C_0}{\gamma \sigma^\tau} |g|_s,$$

where C_0 is the same as previously.

With this corollary, the condition on $\omega(I)$ is much less restrictive, and if we can find a set on which it is verified for all I, then we will be able to solve our equation and the bound on G will hold.
Assume that there exists $x_0 = (I_0, \theta) \in B(r,s)$ such that $\omega(I_0) \in D_{\gamma,\tau}$. Then for all $x = (I, \theta') \in B(r,s)$ (here the value of θ does not matter), we have the inequalities:

$$\begin{aligned}
\|\omega(I) \cdot k - \omega(I_0) \cdot k\|_r &= \|(\omega(I) - \omega(I_0)) \cdot k\|_r \\
&= \left\| \sum_{i=1}^n (\partial_{I_i} H_{0,1}(x) - \partial_{I_i} H_{0,1}(x_0)) k_i \right\|_r \\
&\leq K \sup_{1 \leq i \leq n} \|(\partial_{I_i} H_{0,1}(x) - \partial_{I_i} H_{0,1}(x_0))\|_r \\
&\leq K \|x - x_0\|^I \left\| H''_{0,1} \right\|_r \\
&\leq (2r + diam(B)) K \left\| H''_{0,1} \right\|_r
\end{aligned}$$

Yet,

$$\|\omega(I) \cdot k\|_r \geq | \ \|\omega(I_0) \cdot k\|_r - \|\omega(I) \cdot k - \omega(I_0) \cdot k\|_r \ |$$

Thus, in order to verify the condition on $\omega(I)$, we can ask the term $||\omega(I) \cdot k - \omega(I_0) \cdot k||_r$ to be sufficiently small, or equivalently, the set $B(r,s)$ to be small. We will therefore make the following hypothesis:

$$\forall k \in \mathbb{Z}^n \setminus \{0\}, \text{such that } |k|_1 \leq K, \quad ||\omega(I) \cdot k - \omega(I_0) \cdot k||_r \leq \frac{\gamma}{2|k|_1^\tau} \tag{2.37}$$

Under this condition, the following inequality holds:

$$\begin{aligned}
||\omega(I) \cdot k||_r &\geq ||\omega(I_0) \cdot k||_r - ||\omega(I) \cdot k - \omega(I_0) \cdot k||_r \\
&\geq \frac{\gamma}{2|k|_1^\tau},
\end{aligned}$$

which corresponds to the condition of the corollary.
Gathering the two previous computations, the hypothesis we want to verify is the following:

$$(2r + diam(B)) K ||H''_{0,1}||_r \leq \frac{\gamma}{2K^\tau}.$$

Thus, the corollary we will use is:

Corollary 2.12. *Let $r, s, \sigma > 0$, $K \in \mathbb{N}^\star$, g an analytic function on the set $B(r, s+\sigma)$ such that for all $|k|_1 > K$, the Fourier coefficient g_k is the zero function. If there exists $x_0 = (I_0, \theta) \in B(r, s-\sigma)$ such that $\omega(I_0) \in D_{\gamma,\tau}$, and under the hypothesis*

$$2r + diam(B) \leq \frac{\gamma}{2K^{\tau+1}||H''_{0,1}||_r}, \tag{2.38}$$

then the equation

$$\partial_\omega f = g,$$

where ω is a non constant vector field, has a unique analytic solution f on $\bigcup_{0<t<s} B(r,t)$, and we have the following bound on the norm of f for $0 < \sigma < s$:

$$|f|_{r,s-\sigma} \leq \frac{2C_0}{\gamma\sigma^\tau}|g|_{r,s}$$

with $C_0 = \frac{3\pi}{2}6^{\frac{n}{2}}\frac{\sqrt{\tau\Gamma(2\tau)}}{2^\tau}$.

We managed to find a solution of the cohomological equation on the set $B(r,s)$ at the expense of two things: first, as we had to truncate our initial function to be able to solve our problems, it implies that there will be another remainder to bound; secondly, we introduced a new limit on the size of $B(r,s)$, which will limit our analyticity width available for the next steps.

Estimates on the flow of the vector field

In order to give explicit estimates on the remainder R_f of the transformation, we first have to give a precise estimate on the Hamiltonian vector field X. After, we need to determine an estimation on the size of the flow φ_X^t so as to know the loss of analyticity we suffer while doing the transformation.
Let us show the following lemma, that is very close to the lemma shown previously:

Lemma 2.13. *Let $r, s, \rho, \delta > 0$, $\mathcal{D}' = B(r,s)$ and $\mathcal{D} = B(r,s)$. If*

$$t \leq \min\left(\frac{\rho}{||X||^I_{\mathcal{D}}}, \frac{\delta}{||X||^\theta_{\mathcal{D}}}\right), \tag{2.39}$$

then $\varphi_X^t : \mathcal{D}' \to \mathcal{D} = B(r+\rho, s+\delta)$.

Proof. To show this lemma, define the minimum escape time $\bar{t}$ of φ_X^t from $B(r+\rho, s+\delta)$ relatively to $B(r,s)$:

$$\bar{t} = \bar{t}^{(\rho,\delta)}_{(r,s)} = \inf_{x\in B(r,s)} \sup\{t > 0, \varphi_X^s(x) \in B(r+\rho, s+\delta), \forall |s| < t\}$$

By definition, there exists $\bar{x} \in B(r,s)$ such that $||\varphi_X^{\pm\bar{t}}(\bar{x})||^I = r+\rho$ or $||\varphi_X^{\pm\bar{t}}(\bar{x})||^\theta = s+\delta$. Assume that $\bar{x}$ verify this equality for $t = +\bar{t}$ (the negative time case works exactly the same). Assume again that the we have the equality $||\varphi_X^{\bar{t}}(\bar{x})||^I = r+\rho$, then there exists $i \in [\![1,n]\!]$ such that $||(\varphi_X^{\bar{t}})_i(\bar{x})||^I = r+\rho$. In the case our initial set $B = \{0\}$, we can write:

$$\begin{aligned}
\left|(\varphi_X^{\bar{t}})_i(\bar{x})\right| &= \left|(\varphi_X^0)_i(\bar{x}) + \int_0^{\bar{t}} \frac{d(\varphi_X^s)_i}{ds}(\bar{x})ds\right| \\
&\leq \left|(\varphi_X^0)_i(\bar{x})\right| + \int_0^{\bar{t}} |X_i(\varphi_X^s(\bar{x}))|\,ds \\
&\leq r + \int_0^{\bar{t}} \sup_{x\in\mathcal{D}} |X_i(x)|\,ds \\
&\leq r + \bar{t} \max_{i\in[\![1,n]\!]} \sup_{x\in\mathcal{D}} |X_i(x)| \\
&\leq r + \bar{t}\,\|X\|^I_{\mathcal{D}}.
\end{aligned}$$

So finally:

$$r + \rho \le r + \bar{t}||X||^I_{\mathcal{D}},$$

which implies $\bar{t} \ge \dfrac{\rho}{||X||^I_{\mathcal{D}}}$. The calculation is exactly the same for the angles, thus the result. In the general case where $B \neq \{0\}$, one has to change the modulus by distances to the set B, but the proof would still hold. □

With a bound on the norm of the Hamiltonian G determined in the last section, we can deduce again a bound on the norm of the vector field X. We can then derive a precise condition related to losses of analyticity so that we can consider the flow φ^t_X at a time ϵ.

Lemma 2.14. *Let $r, s, \rho, \rho', \delta, \delta' > 0$, $K \in \mathbb{N} \setminus \{0\}$ and $\mathcal{D}' = B(r,s)$. Assume H is analytic on the set $B(r+\rho+\rho', s+\delta+\delta'+2\sigma)$ and that there exists $I_0 \in B(r,s)$ such that $H'_{0,1}(I_0) \in D(\gamma,\tau)$. Under the assumptions*

$$\begin{cases} 2r + diam(B) \le \dfrac{\gamma}{2K^{\tau+1}||H''_{0,1}||_r} \\ \\ \epsilon \times \|H_1\|_{r+\rho,s+\delta+2\sigma+\delta'} \le \gamma\dfrac{\rho\sigma^\tau\delta'}{4C_0} \\ \\ \epsilon \times \|H_1\|_{r+\rho+\rho',s+\delta+2\sigma} \le \gamma\dfrac{\rho'\delta\sigma^\tau}{4C_0} \\ \\ 4^n n! K^n \exp(-K\sigma) \le 1 \end{cases} \tag{2.40}$$

where C_0 is defined as in corollary 2.12, we have

$$\varphi^\epsilon_X : \mathcal{D}' \to \mathcal{D} = B(r+\rho, s+\delta).$$

In this lemma, the variables ρ, δ are associated with the size of the transformation φ^t_X, while ρ', δ' are associated with the loss of analyticity when we use Cauchy estimates on the vector field X, and finally σ relates to the loss of analyticity when applying Cauchy estimates on H_1. It would be possible to give a shorter lemma making them equal, but here we try to keep track of the different losses induced when proving the lemma (if one knows for example the exact expression of G, then estimates on G do not rely on losing the analyticity width ρ' and δ').

Proof. Under the assumptions of the lemma, corollary 2.12 holds and we have the inequation

$$\|G\|_{r,s} \le \frac{2C_0}{\gamma\sigma^\tau}\left\|\tilde{H}^K_1\right\|_{r,s+\sigma} \tag{2.41}$$

We have as well

$$\begin{aligned} ||\tilde{H}^K_1||_{r,s+\sigma} &\le ||H_1||_{r,s+\sigma} + ||H_1 - \tilde{H}^K_1||_{r,s+\sigma} \\ &\le ||H_1||_{r,s+\sigma} + ||R_K||_{r,s+\sigma} \end{aligned}$$

where

$$R_K(x) = \sum_{k\in\mathbb{Z}^n\setminus\{0\}, |k|_1 > K} h_k(I)\exp(\imath k\cdot\theta)$$

A simple application of lemma B.1 gives the following inequality on the norm of R_K:

$$||R_K||_{r,s} \le 4^n n! K^n \exp(-K\sigma)||H_1||_{r,s+\sigma}. \tag{2.42}$$

Define $C_1(\sigma) = 4^n n! K^n \exp(-K\sigma)$, the previous inequation becomes

$$||R_K||_{r,s} \leq C_1(\sigma)||H_1||_{r,s+\sigma},$$

and therefore, using the fourth assumption:

$$||\tilde{H}_1^K||_{r,s+\sigma} \leq (1 + C_1(\sigma))||H_1||_{r,s+2\sigma} \leq 2||H_1||_{r,s+2\sigma}.$$

Let us go back to our vector field $X = (X_1, X_2)$ obtained by derivation of G among the angles and the actions. A bound on the norm of the two components of X can be derived by the Cauchy formula:

$$\begin{aligned} ||X||^I_{r,s} = ||X_1||^I_{r,s} = ||\partial_\theta G||_{r,s} \leq \frac{1}{\delta'}||G||_{r,s+\delta'}, \\ ||X||^\theta_{r,s} = ||X_1||^\theta_{r,s} = ||-\partial_I G||_{r,s} \leq \frac{1}{\rho'}||G||_{r+\rho',s}, \end{aligned}$$

and so:

$$\begin{cases} ||X||^I_{r,s} \leq \dfrac{4C_0}{\gamma\sigma^\tau\delta'}||H_1||_{r,s+2\sigma+\delta'} \\ \\ ||X||^\theta_{r,s} \leq \dfrac{4C_0}{\gamma\rho'\sigma^\tau}||H_1||_{r+\rho',s+2\sigma} \end{cases} \tag{2.43}$$

With the result of lemma 2.13, and under the assumptions of the lemma, we have $\epsilon \leq \bar{t}$, which implies the result. □

With this transformation, we can now apply the scheme we developed before.

Estimates on the remainders

We are interested in this part by the size of the different remainders coming from the transformation we did. Let us write again our Hamiltonian:

$$H = H_{0,1} + \tilde{H}_1^K + R_K.$$

Therefore, under the assumptions of the last lemma, we can apply φ_X^ϵ, and we get:

$$\begin{aligned} H \circ \varphi_X^\epsilon &= H_{0,1} \circ \varphi_X^\epsilon + \tilde{H}_1^K \circ \varphi_X^\epsilon + R_K \circ \varphi_X^\epsilon \\ &= H_{0,1} + R_f + R_K \circ \varphi_X^\epsilon \end{aligned}$$

There are two remainders to estimate: R_f and $R_K \circ \varphi_X^\epsilon$.
The second one has already been studied in the previous part, and we have:

$$\begin{aligned} \|R_K \circ \varphi_X^\epsilon\|_{r,s} &\leq \|R_K\|_{r+\rho,s+\delta} \\ &\leq C_1(\sigma)\,\|H_1\|_{r+\rho,s+\sigma+\delta} \end{aligned}$$

Let us now work on R_f. Again, we assume that the assumptions of the previous lemma are verified. Recall that:

$$R_f(x,\epsilon) = \int_0^\epsilon t\left((H'_{0,1} \cdot X)' \cdot X\right) \circ \varphi_X^t(x)dt.$$

We are interested by the term under the integral

$$(H'_{0,1} \cdot X)' \cdot X) = H''_{0,1} \cdot X^2 + H'_{0,1} \cdot X' \cdot X.$$

Our transformation $\varphi_X^\epsilon : \ B(r,s) \to B(r+\rho, s+\delta)$, and so:

$$\begin{aligned}
\|R_f(x,t)\|_{r,s} &\le \int_0^\epsilon t \ \left\|(H'_{0,1}\cdot X)'\cdot X)\circ \varphi_X^t(x)\right\|_{r,s} \ dt \\
&\le \int_0^\epsilon t \ \left\|(H'_{0,1}\cdot X)'\cdot X)\right\|_{r+\rho,s+\delta} \ dt \\
&\le \int_0^\epsilon t \ \left\|(H''_{0,1}\cdot X^2)+(H'_{0,1}\cdot X'\cdot X)\right\|_{r+\rho,s+\delta} \ dt \\
&\le \frac{\epsilon^2}{2}\left(\left\|H''_{0,1}\cdot X^2\right\|_{r+\rho,s+\delta}+\left\|H'_{0,1}\cdot X'\cdot X\right\|_{r+\rho,s+\delta}\right) \qquad (2.44)
\end{aligned}$$

The term $H_{0,1}$ depends only in the action, therefore its derivative only has partial derivative for the first n variables. The second derivative has only non zero terms in the first square of size $n \times n$ of its matrix of size $2n \times 2n$. This implies:

$$\left\|H''_{0,1}\cdot X^2\right\|_{r+\rho,s+\delta} \le \left\|H''_{0,1}\right\|_{r+\rho,s+\delta}\left(\|X\|^I_{r+\rho,s+\delta}\right)^2,$$

where we defined:

$$||H''_{0,1}||_{r+\rho,s+\delta} = \sup_{x\in B(r+\rho,s+\delta)}\left(\sum_{(i,j)\in[\![1,n]\!]^2}\left|\partial_{I_i,I_j}H_{0,1}(x)\right|\right).$$

With our previous notations, we also have

$$\left\|H'_{0,1}\cdot X'\cdot X\right\|_{r+\rho,s+\delta} \le \left\|H'_{0,1}\right\|^I_{1,r+\rho,s+\delta}\|X'\cdot X\|^I_{r+\rho,s+\delta}$$

Let now $x \in B(r+\rho, s+\rho)$, $i \in [\![1,n]\!]$:

$$\left|(X'(x)\cdot X(x))_i\right| = \left|\sum_{j=1}^{n}(X'(x))_{i,j}X_j(x) + \sum_{j=n+1}^{2n}(X'(x))_{i,j}X_j(x)\right|$$

$$\begin{aligned}
&\le n\times \max_{j\in[\![1,n]\!]}|(X'(x))_{i,j}| \max_{j\in[\![1,n]\!]}|X_j(x)| + \\
&\quad n\times \max_{j\in[\![n+1,2n]\!]}|(X'(x))_{i,j}| \max_{j\in[\![n+1,2n]\!]}|X_j(x)|
\end{aligned}$$

$$\begin{aligned}
&\le n\times\left(\sup_{x\in B(r+\rho,s+\delta)}\max_{(i,j)\in[\![1,n]\!]^2}\left|(X'(x))_{i,j}\right|\right)||X||^I_{r+\rho,s+\delta} \ + \\
&\quad n\times\left(\sup_{x\in B(r+\rho,s+\delta)}\max_{(i,j)\in[\![1,n]\!]\times[\![n+1,2n]\!]}\left|(X'(x))_{i,j}\right|\right)||X||^\theta_{r+\rho,s+\delta}
\end{aligned}$$

To estimate these two terms, we use Cauchy's inequality on the vector field X. The Jacobian of X can be written as follows, for $i,j \in [\![1,n]\!]$, and $i' \in [\![n+1,2n]\!]$:

$$J_X(I,\theta) = \begin{pmatrix} \partial_{I_j}X_i(I,\theta) & \partial_{\theta_j}X_i(I,\theta) \\ \partial_{I_j}X_{i'}(I,\theta) & \partial_{\theta_j}X_{i'}(I,\theta) \end{pmatrix}$$

$$J_X(I,\theta) = \begin{pmatrix} \partial_{I_j,\theta_i}G(I,\theta) & \partial_{\theta_j,\theta_i}G(I,\theta) \\ -\partial_{I_j,I_i}G(I,\theta) & -\partial_{\theta_j,I_i}G(I,\theta) \end{pmatrix}$$

This allows us to compute the results in the following way:

$$\begin{aligned}\sup_{x\in B(r,s)}\left(\max_{(i,j)\in[\![1,n]\!]^2}|(X'(x))_{i,j}|\right) &= \sup_{x\in B(r,s)}\left(\max_{(i,j)\in[\![1,n]\!]^2}|\partial_{I_j,\theta_i}G(I,\theta)|\right)\\ &\leq \max_{(i,j)\in[\![1,n]\!]^2}\left(\frac{1}{\rho'\delta'}||G||_{r+\rho',s+\delta'}\right)\\ &\leq \frac{4C_0}{\gamma\rho'\sigma^\tau\delta'}||H_1||_{r+\rho',s+2\sigma+\delta'},\end{aligned}$$

and

$$\begin{aligned}\sup_{x\in B(r,s)}\left(\max_{(i,j)\in[\![1,n]\!]\times[\![n+1,2n]\!]}|(X'(x))_{i,j}|\right) &= \sup_{x\in B(r,s)}\left(\max_{(i,j)\in[\![1,n]\!]\times[\![n+1,2n]\!]}|\partial_{\theta_j,\theta_i}G(I,\theta)|\right)\\ &\leq \max_{(i,j)\in[\![1,n]\!]\times[\![n+1,2n]\!]}\left(\frac{1}{\delta'^2}||G||_{r,s+2\delta'}\right)\\ &\leq \frac{4C_0}{\gamma\sigma^\tau\delta'^2}||H_1||_{r,s+2\sigma+2\delta'}\end{aligned}$$

We can now give a bound on the two terms of R_f. First,

$$\left\|H''_{0,1}\cdot X^2\right\|_{r+\rho,s+\delta} \leq \frac{16C_0^2}{\gamma^2\sigma^{2\tau}\delta'^2}\left\|H''_{0,1}\right\|_{r+\rho,s+\delta}\|H_1\|^2_{r+\rho,s+2\sigma+\delta+\delta'},$$

secondly,

$$\left\|H'_{0,1}\cdot X'\cdot X\right\|_{r+\rho,s+\delta} \leq n\left\|H'_{0,1}\right\|^I_{1,r+\rho,s+\delta}\frac{32C_0^2}{\gamma^2\rho'\sigma^{2\tau}\delta'^2}\|H_1\|^2_{r+\rho+\rho',s+2\sigma+\delta+2\delta'}.$$

Finally, we have:

$$\|R_f\|_{r,s} \leq \frac{8nC_0^2\epsilon^2}{\gamma^2\sigma^{2\tau}\delta'^2}\Bigg(\left\|H''_{0,1}\right\|_{r+\rho,s+\delta}\|H_1\|^2_{r+\rho,s+2\sigma+\delta+\delta'} + \frac{2}{\rho'}\left\|H'_{0,1}\right\|^I_{1,r+\rho,s+\delta}\|H_1\|^2_{r+\rho+\rho',s+2\sigma+\delta+2\delta'}\Bigg). \tag{2.45}$$

2.2.3 Explicit statement and application to the three-body problem

In this section, we will gather all the information from the lemmas and the estimates on the remainder to give an explicit theorem such as theorem 2.9. First we will give a general statement in dimension n, and secondly, we will apply it to the case of the plane planetary three-body problem. In the latter case, we will get rid of some constants by estimating them, in order to have a simple and applicable theorem that we can use easily.

In this aim, we will now lose track of the different kind of analyticity losses, by considering $\rho = \rho'$ and $\sigma = \delta = \delta'$.

Theorem 2.15. *Let $H(I,\theta) = H_0(I) + \epsilon H_1(I,\theta)$ be a Hamiltonian analytic on the set $B(r,s)$ with $r, s > 0$. Let $\rho < r$, $\delta < s$, $K \in \mathbb{N}^\star$, and $H_{0,1} = H_0 + \epsilon/(2\pi)^n \int_{\mathbb{T}^n} H_1$. Assume that there exists $I_0 \in B(r-\rho, s-\sigma)$ such that $H'_{0,1} \in D(\gamma,\tau)$.*
Under the assumptions:

$$\begin{cases} 2r + diam(B) \leq \dfrac{\gamma}{2K^{\tau+1}\left\|H''_{0,1}\right\|_{r,s}}, \\[2ex] \dfrac{8\times 5^{\tau+1}C_0\epsilon}{\gamma}\times\|H_1\|_{r,s} \leq \rho\sigma^{\tau+1}, \\[2ex] 4^n n! K^n \exp\left(\dfrac{-4K\sigma}{5}\right) \leq 1, \end{cases}$$

where $C_0 = \frac{3\pi}{2} 6^{\frac{n}{2}} \frac{\sqrt{\tau\Gamma(2\tau)}}{2^\tau}$*, then there exists a symplectic transformation* $\varphi_X^\epsilon : B(r-\rho, s-\sigma) \to B(r,s)$ *such that on the set* $B(r-\rho, s-\sigma)$ *we have*

$$H \circ \varphi_X^\epsilon(I,\theta) = H_{0,1}(I) + \epsilon^2 H_2(I,\theta).$$

Moreover, the following bound on H_2 *holds:*

$$\begin{aligned}\|H_2\|_{r-\rho,s-\sigma} \leq & \frac{5^{2(\tau+1)} 8nC_0^2}{\gamma^2\sigma^{2(\tau+1)}} \|H_1\|_{r,s}^2 \left(\left\| H_{0,1}'' \right\|_{r,s} + \frac{4}{\rho} \left\| H_{0,1}' \right\|_{1,r,s}^I \right) + \\ & \frac{4^n n! K^n}{\epsilon} \exp\left(-\frac{4K\sigma}{5}\right) \|H_1\|_{r,s} .\end{aligned} \tag{2.46}$$

The proof is direct with the scheme we described, and with the results of lemma 2.14 and the estimates on the remainders.

We wish now to apply this theorem to the plane planetary three-body problem. We are interested here in the coordinates (Λ, λ) and will consider the other variables as parameters. In this case, we have $n = 2$. We will consider as well that $B = \{0\}$, and that there exists a I_0 in some $B(r-\rho, s-\sigma) \in D(\gamma, n) = D(\gamma, 2)$. We want to push at a higher order the part of the Hamiltonian that depends on the fast angles λ. We will therefore apply the previous theorem in this case.
First, we can bound C_0 in the case $\tau = n = 2$ by:

$$C_0 = \frac{3\pi}{2} 6^{\frac{n}{2}} \frac{\sqrt{\tau\Gamma(2\tau)}}{2^\tau} \leq 25 = 5^2$$

As well, we will consider that we lose half of our analyticity width during the operation, that corresponds to taking $r = 2\rho$ and $s = 2\sigma$. We have the following corollary:

Corollary 2.16. *Let* $H : \mathbb{R}^2 \times \mathbb{T}^2 \to \mathbb{R}$ *be a Hamiltonian of the form* $H(I,\theta) = H_0(I) + H_1(I,\theta)$*. Assume that it is analytic on the set* $B(r,s)$ *with* $r, s > 0$*. Let* $K \in \mathbb{N}^\star$*, and* $H_{0,1} = H_0 + \epsilon/(2\pi)^2 \int_{\mathbb{T}^2} H_1$*. Assume that there exists* $I_0 \in B(r/2, s/2)$ *such that* $H_{0,1}'(I_0) \in D(\gamma, 2)$*. Under the assumptions*

$$\begin{cases} K \leq \left(\dfrac{\gamma}{2r \left\| H_{0,1}'' \right\|_{r,s}} \right)^{\frac{1}{3}}, \\ \\ \|H_1\|_{r,s} \leq \dfrac{\gamma r s^3}{4.10^5}, \\ \\ 32K^2 \exp\left(\dfrac{-2Ks}{5}\right) \leq 1, \end{cases}$$

then there exists a symplectic transformation $\varphi_X^\epsilon : B(r/2, s/2) \to B(r,s)$ *such that on the set* $B(r/2, s/2)$ *we have*

$$H \circ \varphi_X^\epsilon(I,\theta) = H_{0,1}(I) + H_2(I,\theta).$$

Moreover, the following bound on H_2 *holds:*

$$\begin{aligned}\|H_2\|_{r/2,s/2} \leq & \frac{10^{10}}{\gamma^2 s^6} \|H_1\|_{r,s}^2 \left(\left\| H_{0,1}'' \right\|_{r,s} + \frac{8}{r} \left\| H_{0,1}' \right\|_{1,r,s}^I \right) + \\ & 32K^2 \exp\left(-\frac{2Ks}{5}\right) \|H_1\|_{r,s} .\end{aligned}$$

With this corollary, we have a symplectic change of variables which (if the norm of the perturbation is small enough) decreases the norm of the perturbation at the expense of a loss of half of our analyticity width. Releasing the constraints on the parameters, the new unperturbed Hamiltonian $H_{0,1}$ depends this time on the 4 action variables, and the perturbation of the Hamiltonian is now much smaller, as wanted.

2.2.4 Iterating the scheme

In the case the new perturbation H_2 is not small enough to apply the KAM theorem, we have to iterate this scheme to push again the perturbation to a higher order. We are therefore going to iterate the following lemma two more times (this is actually the exact number that will be necessary to apply the KAM theorem, but another argument is that we will show one step in order to see the changes we have to do to the previous corollary, and a second one to show the recurrence factors that arose from a further iteration).
An important difference between the previous computations and an iteration of this computation is that we do not know anything about H_2 but the estimate on its bound. Whereas we gave an estimate on H_2 that depended on $H'_{0,1}$, if we consider the Hamiltonian $H_{0,2} = H_{0,1} + \bar{H}_2$, we do not know this time the exact expression of $H_{0,2}$. It is easily possible to overcome any difficulty related to the size of H'_2 by using Cauchy estimates, though solving the cohomological equation will imply to be careful of the change of frequencies induced by H_2.
It is necessary to make one more assumption each time we iterate to be able to obtain the result of the corollary. Here we will assume that for some $K \in \mathbb{N} \setminus \{0\}$ and $r', s' > 0$:

$$\|H'_2\|_{r',s'} \leq \frac{\gamma}{4K^\tau}$$

This will allow us to have an inequality of the form:

$$\begin{aligned}\|\omega_2(I) \cdot k\|_{r'} &= \left\|H'_{0,1}(I) \cdot k + H'_2(I) \cdot k\right\|_{r'} \\ &\geq H'_{0,1}(I_0) \cdot k - \frac{\gamma}{2K^\tau} - \frac{\gamma}{4K^\tau} \\ &\geq \frac{\gamma}{4|k|_1^\tau}\end{aligned}$$

Therefore, under this condition, we have a non-resonance condition as before for the constant $\gamma/2$. We can moreover compute the size of $\|H'_2\|_{r',s'} \leq \frac{\gamma}{4K^\tau}$ with the help of a Cauchy estimate. With that assumption, we can consider the same lemma as before with analyticity width divided by 2. Therefore, for the next order, we have the corollary of theorem 2.15:

Corollary 2.17. *Let $H : \mathbb{R}^2 \times \mathbb{T}^2 \to \mathbb{R}$ be a Hamiltonian of the form $H(I,\theta) = H_{0,1}(I) + H_2(I,\theta)$ with H_2 obtained by the corollary 2.16. Assume that it is analytic on the set $B(r/2, s/2)$ with $r, s > 0$. Let $K \in \mathbb{N}^\star$, and $H_{0,2} = H_{0,1} + \epsilon/(2\pi)^n \int_{\mathbb{T}^n} H_2$. Assume that there exists $I_0 \in B(r/8, s/4)$ such that $H'_{0,1}(I_0) \in D(\gamma, 2)$. Under the assumptions*

$$\begin{cases} K^3 \leq \min\left(\dfrac{2\gamma}{r\left\|H''_{0,1}\right\|_{r,s}}, \dfrac{r\gamma}{16\left\|H_2\right\|_{r/2,s/2}}\right), \\ \left\|H_2\right\|_{r/2,s/2} \leq \dfrac{\gamma r s^3}{2^8.10^5}, \\ 32K^2 \exp\left(\dfrac{-Ks}{5}\right) \leq 1, \end{cases}$$

there exists a symplectic transformation $\varphi_2^\epsilon : B(r/8, s/4) \to B(r/2, s/2)$ *such that on the set* $B(r/8, s/4)$ *we have*

$$H \circ \varphi_2^\epsilon(I, \theta) = H_{0,2}(I) + H_3(I, \theta).$$

Moreover, the following bound on H_3 *holds:*

$$\|H_3\|_{r/8,s/4} \leq \frac{2^8.10^{10}}{\gamma^2 s^6} \|H_2\|^2_{r/2,s/2} \left(\left\| H''_{0,1} \right\|_{r,s} + \frac{2^5}{r} \left\| H'_{0,1} \right\|^I_{1,r,s} + \frac{2^8}{r^2} \|H_2\|_{r/2,s/2} \right) + 32K^2 \exp\left(-\frac{Ks}{5}\right) \|H_2\|_{r/2,s/2} .$$

The result is straightforward under the additional assumption. One only changes γ, r, s to $\gamma/2, r/4, s/2$ to compute the new assumptions and the bound on H_3. The fact that we have to divide not by 2 but by 4 the analyticity width among the angles comes from the estimate on the derivative of H_2 by Cauchy's inequality, that makes us lose some more width.

While iterating another time, we need again to add an assumption on H_3, and we proceed as precedes. We will not give the precise corollary, just the new hypothesis necessary and the bound on H_3.
Assume that there exists $I_0 \in B(r/32, s/8)$, such that $H'_{0,1}(I_0) \in D(\gamma, 2)$, and consider that we have already applied corollaries 2.16 and 2.17. Then the assumptions to iterate a third time are:

$$\begin{cases} K^3 \leq \min\left(\dfrac{8\gamma}{r \left\| H''_{0,1} \right\|_{r,s}}, \dfrac{r\gamma}{2^6 \|H_2\|_{r/2,s/2}}, \dfrac{r\gamma}{2^7 \|H_3\|_{r/8,s/4}} \right), \\ \|H_3\|_{r/8,s/4} \leq \dfrac{\gamma r s^3}{2^{14}.10^5}, \\ 32K^2 \exp\left(\dfrac{-Ks}{10}\right) \leq 1 \end{cases}$$

We then have our symplectic application $\varphi_3^\epsilon : B(r/32, s/8) \to B(r/8, s/4)$, and the bound on H_4:

$$\begin{aligned} \|H_4\|_{r/32,s/8} \leq & \frac{2^{16}.10^{10}}{\gamma^2 s^6} \|H_3\|^2_{r/8,s/4} \times \\ & \left(\left\| H''_{0,1} \right\|_{r,s} + \frac{2^7}{r} \left\| H'_{0,1} \right\|^I_{1,r,s} + \frac{2^{10}}{r^2} \left(\|H_2\|_{r/2,s/2} + \|H_3\|_{r/8,s/4} \right) \right) + \\ & 32K^2 \exp\left(-\frac{Ks}{10}\right) \|H_3\|_{r/2,s/2} . \end{aligned}$$

Observe that for further iterating the scheme, one would have to add an assumption that is close to the new ones we have made, but with a factor 1/2 appearing between these assumptions. Moreover, one can see that we cannot iterate indefinitely this scheme, and produce directly a KAM theorem with it. Though, it was not the goal here, and this work will be enough towards our aim of applying the KAM theorem to the three-body problem.

Chapter 3

Explicit KAM theorem with parameters

This chapter is dedicated to the statement of a version of the KAM theorem with parameters in the case the Hamiltonian is analytic, and with explicit constants. In the previous chapters, we first estimated the analyticity width of the Hamiltonian of the plane planetary three-body problem, and then we showed that we could separate it into a secular part, depending only in the action variables, and a smaller perturbation depending in the action-angle variables. We want now to derive a classical KAM theorem for an analytic Hamiltonian. This theorem will require further assumptions (such as the non-degeneracy of the unperturbed Hamiltonian that will be proved later).

As of now, there exists different KAM theorems, related to different points of view. We chose the version of Pöschel [62] for several reasons. First, the intelligibility in the statement of the theorem and its implications makes it easy to understand the problem we are dealing with. Secondly, the KAM step we used in the section 2.2 is very close to Pöschel's. Finally, Pöschel's proof shows all the dependencies in the different variables, the only non-explicit constants being those depending on the dimension n and the Diophantine constant τ. Therefore, we "only" have to explicit these constants (and to rebuild the statement at the same time in order to make it more precise). We will follow Pöschel's proof in order to make it easier to read, and compute the constants appearing in each calculation. We encourage the reader to consult it in order to find a great amount of references for closely related problems.

The theorem we are going to use is a parameterized version of the KAM theorem. This version induces a great loss of analyticity considering the Hamiltonian we are using (for some reason we will explain further), though it is convenient for our use. In this approach, one considers the perturbation of a family of linear Hamiltonians, parameterized by the frequencies ω of the unperturbed Hamiltonian. The non-degeneracy condition is then contained in the dependency of the frequency parameter ω with respect to the initial actions. This approach dates back to the work of Moser [49], and relies on adding the quadratic part of the unperturbed Hamiltonian to the perturbation, leaving only linear terms in the new unperturbed Hamiltonian.

3.1 General scheme of the KAM theorem with parameters

In this section, we will first describe the operation of parameterization of the Hamiltonian with respect to the frequencies. Then we will give the general statement of Pöschel and discuss it. Finally, we will quickly give the outline of the proof, that resembles strongly to the one we have developed in section 2.2, but taking care of the frequency variations at each step (we will therefore not detail too much the explanations).

3.1.1 Parameterizing the Hamiltonian

Consider, for $n \geq 2$ the following analytic Hamiltonian:

$$H(p,q) = h(p) + \epsilon f(p,q,\epsilon), \qquad (p,q) \in D \times \mathbb{T}^n, \ \epsilon \ll 1, \tag{3.1}$$

where $D \subset \mathbb{R}^n$, and $|f| \sim |h|$.
This Hamiltonian is divided into two terms, the unperturbed Hamiltonian h, leading to a quasi-periodic motion of frequency $\omega(p) = h'(p) \in \mathbb{R}^n$, and the perturbation $\epsilon f(p,q,\epsilon)$ of small norm compared to the Hamiltonian.
Consider now that the unperturbed Hamiltonian is non-degenerate on the set D:

$$det(h''(p)) = det\left(\frac{\partial\omega}{\partial p}(p)\right) \neq 0.$$

Then the frequency map $h' : D \to \Omega$ is a local diffeomorphism between D and the frequency set $\Omega \subset \mathbb{R}^n$. The approach of Moser consists in expanding the Hamiltonian h around one particular frequency ω, and work with a linear Hamiltonian parameterized with that frequency.
Let $p = p_0 + I$, with $I \in B = D - p_0$, the Hamiltonian h then becomes:

$$h(p) = h(p_0) + \langle h'(p_0), I\rangle + \int_0^1 (1-t)\langle h''(p_0 + tI)I, I\rangle dt \tag{3.2}$$

The frequency map being a local diffeomorphism, it is equivalent to work the coordinates ω (which is one-to-one with the variable p_0) and I, or with a variable p. Working with the derivative and its inverse can be simplified by introducing the Legendre transform of h, that we will call g, and defined by

$$g(\omega) = \sup_{p\in D}(\langle p, \omega\rangle - h(p))$$

We then have the relation $g'(\omega) = (h'(p_0))^{-1}$ where $\omega = h'(p_0)$. Fixing the action p_0 (and therefore ω), one can write the equation (3.2):

$$h(p) = e(\omega) + \langle\omega, I\rangle + P_h(I;\omega),$$

with $P_h(I;\omega) = \int_0^1 (1-t)\langle h''(g'(\omega) + tI)I, I\rangle dt$. From the action-angle coordinates, we defined new coordinates $(\omega, I, \theta) \in \Omega \times B \times \mathbb{T}^n$, where we wrote θ instead of q not to be mistaken.
Now we can consider the term P_h to be part of the perturbation, since a bound on its norm can be made as small as wanted by looking at a sufficiently small ball in the action I around the origin.
One can then write $H = N + P$, where $N = e(\omega) + \langle\omega, I\rangle$, N being called the normal form, and

$$P = P_h(I;\omega) + P_\epsilon(I, \theta; \omega),$$

with $P_\epsilon(I,\theta;\omega) = \epsilon f(g'(\omega) + I, \theta, \epsilon)$.
The family of Hamiltonians under normal form N have equations of motions that are easy to compute. Indeed, the vector field associated to it is

$$X_N = \sum_{j=1}^n \omega_j \frac{\partial}{\partial\theta_j}, \qquad \omega \in \Omega.$$

The motion is quasi-periodic, and takes place on a specific torus $\{0\} \times \mathbb{T}^n$ for every $\omega \in \Omega$. These tori can be seen as a trivial embedding of $\mathbb{T}^n$ over the set Ω in the phase space given by the function

$$\begin{aligned} \Phi_0 : \mathbb{T}^n \times \Omega &\to B \times \mathbb{T}^n \\ (\theta, \omega) &\mapsto (0, \theta) \end{aligned}$$

For a generic Hamiltonian, the perturbation P will limit the existence of these tori. However, under several hypotheses, one can show that almost all of these tori (in the sense of the Lebesgue measure) survive a perturbation, this is the main result of the KAM theorem.
The goal of this KAM theorem with parameters is to eliminate the whole perturbation P instead of removing only the perturbation f, while modifying slightly the normal form N. Observe that in order for the perturbation P to verify a smallness condition, we have to restrict ourselves to a set of actions close to the origin. This implies that we will lose some analyticity width artificially to make P_h of the order of f.

3.1.2 Domains of analyticity and other definitions

Before stating the KAM theorem, we have to make explicit the sets we are working on.
Let Ω be the set of initial frequencies we are considering. Let $\tau > n-1$, $\gamma > 0$ and consider $D(\gamma, \tau)$ the Cantor set of real numbers verifying a Diophantine condition for the constants γ and τ. Let $\Gamma_\gamma = \Omega \cap D(\gamma, \tau)$, it is as well a Cantor set. Finally, for $\beta > 0$, let

$$\Omega_\gamma^\beta = \Gamma_\gamma \setminus \{\omega \in \Gamma_\gamma : \exists\, \omega' \in \mathbb{R}^n \setminus \Omega, |\omega - \omega'| < \beta\}.$$

The last set is therefore composed of the points of the set Γ_γ that are at least at a distance β to the boundary of Ω. We will fix later the needed value for the constant β.
Now let us define the various domains we will use in the theorem. These sets will be polydiscs around some set. For the frequencies, define

$$O_h = \{\omega \in \mathbb{C}^n, |\omega - \Omega_\gamma^\beta| < h\}.$$

For the action-angle variable, let

$$D_{r,s} = \{I \in \mathbb{C}^n, |I| < r\} \times \{\theta \in \mathbb{T}^n_{\mathbb{C}}, |\Im(\theta)| < s\}.$$

Define the norms with indices as follows, for $f : \mathbb{C}^n \times \mathbb{T}^n \to \mathbb{C}$:

$$||f||_{r,s} = \sup_{D_{r,s}} |f|\ , \quad ||f||_h = \sup_{O_h} |f|,$$
$$||f||_{r,s,h} = \sup_{D_{r,s}\times O_h} |f|.$$

For vector valued functions, we have:

$$||f||_{r,s} = \sup_{D_{r,s}} ||f||,$$

where $||\cdot||$ is the sup norm. When considering the Diophantine condition, we will always consider the norm $|k|_1 = |k_1| + ... + |k_n|$ for the vectors $k \in \mathbb{Z}^n$. Finally, to state the theorem we will need the following Lipschitz norm on the frequencies:

$$|f|_L = \sup_{\omega \neq \omega'} \frac{|f(\omega) - f(\omega')|}{|\omega - \omega'|},$$

where $|\cdot|$ represent the supremum norm.

3.1.3 Statement of the KAM theorem with parameters

The precise statement of the KAM theorem in Pöschel's paper is the following:

Theorem 3.1. *Let $H = N + P$. Suppose P is real analytic on $D_{r,s} \times O_h$ with*

$$||P||_{r,s,h} \leq c\gamma r s^{\nu}, \qquad \gamma s^{\nu} \leq h,$$

where $\nu = \tau + 1$ and c is a small constant depending only on n and τ. Suppose also that $r, s, h \leq 1$. Then there exists a Lipschitz continuous map $\varphi : \Gamma_\gamma \to \Omega$ close to the identity and a Lipschitz continuous family of real analytic torus embeddings $\Phi : \mathbb{T}^n \times \Gamma_\gamma \to B \times \mathbb{T}^n$ close to Φ_0 such that for each $\omega \in \Gamma_\gamma$ the embedded tori are Lagrangian and

$$X_{H\,|\varphi(\omega)} \circ \Phi = \Phi' \cdot X_N.$$

Moreover, Φ is real analytic on $T_\star = \{\theta : |\Im\theta| < s/2\}$ for each ω, and

$$\|W(\Phi - \Phi_0)\|, \gamma s^{\nu}\|W(\Phi - \Phi_0)\|_L \leq \frac{c}{\gamma r s^{\nu}}||P||_{r,s,h},$$
$$\|\varphi - Id\|, \gamma s^{\nu}\|\varphi - Id\|_L \leq \frac{c}{r}||P||_{r,s,h},$$

uniformly on $T_\star \times \Gamma_\gamma$ and Γ_γ, respectively, where c is a large constant depending only on n and τ, and $W = diag(r^{-1}Id, s^{-1}Id)$.

Qualitatively, the statement may be described as follows: first, as said before, while removing the perturbation, we are slightly changing the frequencies; secondly, the Lipschitz estimates allow us to control the size of the the set $\varphi(\Gamma_\gamma)$. Indeed, with these estimates, one can prove that the complement of this set is of size $O(\alpha)$; finally, every embedded torus is Lagrangian and is close to its associated unperturbed torus.
In the next section, we will first make explicit all the constants that appear in the theorem. Secondly, we will gather the hypotheses in order to have a simpler smallness condition.

3.1.4 Sketch of the proof

Let us describe the general scheme of the proof, that will consist in an iteration of a KAM step. Since the KAM step will be close to the theorem we proved in 2.2, we will be quick on the main operation, and will describe more precisely the change in frequencies.
At each step, we consider a Hamiltonian $H = N + P$ with N under normal form: $N = e + \langle \omega, I \rangle$. We want to find a transformation $\mathcal{F}$ such that $H \circ \mathcal{F} = N_+ + P_+$, where N_+ is again under normal form and P_+ verifies $||P_+|| \leq C||P||^{\kappa}$, for some $\kappa > 1$. This transformation will make us lose some analyticity width related to the norm of P, though, the constant κ will ensure that the scheme will rapidly converge and that we can eliminate the initial perturbation if it is small enough.
To build the transformation $\mathcal{F}$, instead of considering P, we will consider only the linear part in the action (by truncating the Taylor expansion in the action at the order 2). Next, as done in the previous chapter, we truncate the Fourier series in the angle at some order K. Let us call the new Hamiltonian after this two steps R. The remainder $P - R$ will be either of order 2 in the actions, and therefore small looking close to the origin, or it will be part of the remainder of the Fourier series, which will be small as well if K is high enough.
Assume ω is fixed. Let F be a Hamiltonian affine in the actions, and $X_F = X$ its associated Hamiltonian vector field. Call $\Phi^t = \Phi^t_X$ the flow associated to the previous vector field, and $\Phi = \Phi^t|_{t=1}$ the time-1 map of this flow.
Call $\bar{H} = N + R$, we have:

$$\begin{aligned}\bar{H} \circ \Phi &= N + \{N, F\} + \int_0^1 (1-t)\{\{N, F\}, F\} \circ \Phi^t dt + R + \int_0^1 \{R, F\} \circ \Phi^t dt \\ &= N + \{N, F\} + R + \int_0^1 \{(1-t)\{N, F\} + R, F\} \circ \Phi^t dt\end{aligned}$$

The term under the integral will constitute another part of the perturbation, in sum with $P-R$. We want F such that

$$N + \{N, F\} + R = N_+ \tag{3.3}$$

The main difference from the previous scheme, where we solved the cohomological equation, is that R is not of zero-mean, and therefore we need to divide it into two parts: $R = \bar{R} + \tilde{R}$ where

$$\bar{R} = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} R d\theta$$

We can therefore define F by the formula

$$F = \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{R_k}{\imath \langle k, \omega \rangle} \exp(\imath k \cdot \theta),$$

where the R_k are the Fourier coefficients of the Hamiltonian R. As for the average of R with respect to the angles, since it is not possible to remove this term with the symplectic transformation we built, we simply add it to the unperturbed Hamiltonian. Observe that since R was affine in the actions, $N_+ = N + \bar{R}$ remains affine in the actions ($\bar{R}$ is obviously independent of the angles). We can then write:

$$N_+ = e_+(\omega) + \langle \omega + v(\omega), I \rangle$$

The new frequency vector is given by $\omega_+ = \omega + v(\omega)$, and if v is small, then there exists a map φ close to the identity such that $\varphi(\omega_+) = \omega$.
The map N_+ can then be written $N_+ = (N + \bar{R}) \circ \varphi$. As for the perturbation, computing the remainders of the transformation gives:

$$P_+ = \int_0^1 \left\{ (1-t)\bar{R} + tR, F \right\} \circ \Phi^t dt + (P - R) \circ \Phi$$

The total transformation $\mathcal{F}$ then corresponds to the map (Φ, φ). Moreover, observe that Φ can be written $\Phi = (U, V)$, both function depending on the parameter ω, U depending on the action and the angle, but V depending only on the angles by construction. With these definitions, this scheme can be iterated under specific assumptions on the analyticity width of the Hamiltonian.

The loss of analyticity related to one iteration has its roots in different parts of the scheme. Estimating the bound of R requires already to lose some analyticity width. Then, as we saw before, solving the cohomological equation will induce a new loss, as well as estimating the derivatives with Cauchy's inequality. The size of the ball for the frequencies will also decrease while computing the function φ, that is the inverse of the function $Id + v$ on some domain. All these variables need to be controlled while iterating the scheme.
The difference between the previous scheme scheme and its iterations we developed in section 2.2 is that we were going further and further from a vector verifying a Diophantine condition, whereas in this scheme, the iteration requires to get closer and closer to one (in order to be able to solve the cohomological equation with a perturbation truncated at an increasing order). One of the difficulties is therefore to control the frequencies, and instead of perturbing the initial frequency, we consider our initial frequency to be the perturbation of a Diophantine vector.

3.2 Quantitative KAM step and its proof

Before giving the full explicit KAM theorem statement, we will work on the KAM step, that will be iterated an infinite amount of times so as to obtain the main statement. This will allow

us to explain the different facets of this theorem. Again, we follow almost completely the work of Pöschel in his proof, although we will give first a precise statement, with all the necessary constants.
It is important to say that we chose the case $\tau = n$ as Diophantine constant for the sake of simplicity. Keeping τ as an unknown should not be much harder to compute with the work here, since all the constants are computed explicitly.

3.2.1 Statement of the KAM step

Preliminary definition

We consider a Hamiltonian H real analytic over the set $D_{r,s} \times O_h$, for some $r, s, h > 0$, such that H can be written as the sum of a normal form and a small perturbation, *i.e.* $H = N + P$.
Define $\nu = \tau + 1 = n + 1$. Let $\sigma < s/5$, and $\delta \geq 40$, a constant that will appear later while performing the transformation.
Define the following sets:

$$A = \{K \in \mathbb{N} : K\sigma \geq (2n+\nu) log\ 2\} \tag{3.4}$$

$$B = \left\{K \in \mathbb{N} : 2K^{n+\nu}\sigma^{\nu}e^{-K\sigma} \leq \frac{1}{\delta}\right\} \tag{3.5}$$

Now B consists of two integer intervals, one for small K (for instance $K = 0$ belongs to B,) and one for large K (when K goes to infinity, the left term tends to zero). We are interested in the second one, hence we define

$$B_+ = \{K \in B : \ \forall m \in \mathbb{N},\ K + m \in B\} \tag{3.6}$$

Now let $K' = \min(A \cap B_+) > 0$. This minimum is well defined, this two sets having non-empty intersection (K sufficiently large verifies the two assumptions). Recall the definition of the constant appearing in Rüssmann's theorem while solving the cohomological equation, $C_0 = 3\pi 6^{\frac{n}{2}}\frac{\sqrt{n(2n)!}}{2^n}$ (we choose to multiply it by a factor 2 for a reason explained later). We can finally define:

$$\epsilon = \min\left(\frac{\gamma r \sigma^{\nu}}{4^{\nu}C_1}, \frac{hr}{\delta}, \frac{\gamma r}{2K'^{\nu}\delta}\right) \tag{3.7}$$

with $C_1 = 4^{\nu}(200nC_0 + 32 + 8^{\nu}n!)^2$.

The KAM step can then be stated as follows:

Proposition 3.2. *Assume that $\|P\|_{r,s,h} \leq \epsilon$. Then there exists a real analytic transformation*

$$\mathcal{F} = (\Phi, \varphi) : \ D_{\eta r, s-5\sigma} \times O_{h/4} \to D_{r,s} \times O_h$$

with $\eta = \sqrt{\frac{\epsilon}{\gamma r \sigma^{\nu}}}$ such that $H \circ \mathcal{F} = N_+ + P_+$ with

$$\|P_+\|_{\eta r, s-5\sigma, h/4} \leq 200nC_0\frac{\epsilon^2}{\gamma r \sigma^{\nu}} + (32\eta^2 + 4^{\nu}n!K^n e^{-K\sigma})\epsilon.$$

Moreover,

$$2\|W(\Phi - Id)\|, \|W(D\Phi - Id)W^{-1}\| \leq \frac{40C_0\epsilon}{\gamma r \sigma^{\nu}}$$

$$||\varphi - Id||, 4h||D\varphi - Id|| \leq \frac{10\epsilon}{r}$$

uniformly on $D_{\eta r, s-5\sigma} \times O_h$ and $O_{h/4}$ respectively, with the weight matrix

$$W = diag(r^{-1}Id, \sigma^{-1}Id).$$

We will prove this proposition in the next section. Let us first discuss the definitions we have made.

Observations on the statement

The different conditions on ϵ arise from different parts of the proof. The first condition in the minimum is a limit due to the analyticity width we have on the actions and on the angles. This limit is necessary to obtain an exponential decrease of the bound on the norm of the perturbation at each step. The second condition is related to the transformation on the frequency vector. To be able to invert the map giving the new frequency $\omega + v(\omega)$, it is essential to have enough analyticity width compared to the size of v. The third condition is a condition on K'. K' needs to be big enough to allow the remainder of the Fourier series of Q to be small enough, and then for this remainder to decrease exponentially while iterating our scheme. However, we want as well K' to be small enough so that all the frequencies in O_h verify a non-resonance condition of order K'.
In the KAM step, the factor δ appears always at the denominator, and therefore always seems to add a stronger constraint when increasing it. Though, when iterating the scheme, we will see that it allows to make the norm of the transformation smaller while increasing this factor. This is the reason we keep it in the computation.

3.2.2 Proof of the proposition

In this section, we make the hypotheses of proposition 3.2.

Implication of the hypotheses

Let $\epsilon_0 \leq \epsilon$. Define $h_0 = \dfrac{\delta\epsilon_0}{r}$ and $K_0 = \left\lfloor \sqrt[\nu]{\dfrac{\gamma}{2h_0}} \right\rfloor$. These two constants satisfy $h_0 < h$ and $K_0 > K'$. Indeed, the first inequality is clear, and for the second one, using the fact that both K_0 and K' are integers, and the definition of ϵ:

$$\begin{aligned} K'^{\nu} &\leq \frac{\gamma r}{2\delta\epsilon} \leq \frac{\gamma r}{2\delta\epsilon_0} \\ &\leq \frac{\gamma}{2h_0}. \end{aligned}$$

The definition of h_0 means that if we consider a smaller perturbation, we will not use all the available analyticity h corresponding to the frequencies.
These definitions allow us to compute a central inequality, that will prove useful later on in the proof:

$$\frac{1}{2^{\nu}} K_0^n \sigma^{\nu} \exp(-K_0\sigma) \leq \frac{\epsilon_0}{\gamma r} \leq \frac{h_0}{\gamma\delta} \leq \frac{1}{2\delta K_0^{\nu}}. \tag{3.8}$$

The last two inequalities are straightforward given the definition of h_0 and K_0. Regarding the first one, using the definition of B_+:

$$\begin{aligned} \frac{\epsilon_0}{\gamma r \sigma^{\nu}} = \frac{h_0}{\gamma\sigma^{\nu}\delta} &\geq \frac{2K_0^{n+\nu} e^{-K\sigma}}{\gamma} h_0 \\ &\geq K_0^n e^{-K_0\sigma} \left(\sqrt[\nu]{\frac{2h_0}{\gamma}} \times \left\lfloor \sqrt[\nu]{\frac{\gamma}{2h_0}} \right\rfloor \right)^{\nu} \\ &\geq K_0^n e^{-K_0\sigma} \left(\left\lceil \sqrt[\nu]{\frac{\gamma}{2h_0}} \right\rceil^{-1} \times \left\lfloor \sqrt[\nu]{\frac{\gamma}{2h_0}} \right\rfloor \right)^{\nu} \geq \frac{1}{2^{\nu}} K_0^n e^{-K_0\sigma}. \end{aligned}$$

We will now use the definition of K_0 to show the non-resonance condition that the frequency vectors of O_{h_0} must verify. Indeed, let k such that $0 < |k| \leq K_0$, and let $\omega \in O_{h_0}$, there exists $\omega^\star \in \Omega_\gamma^\beta$ such that $|\omega - \omega^\star| < h_0$, and therefore, the following inequalities hold:

$$|\langle k, \omega - \omega^\star \rangle| \leq |k||\omega - \omega^\star| \leq K_0 h_0 \leq \frac{\gamma}{2K_0^\tau} \leq \frac{\gamma}{2|k|^\tau}.$$

Since $\omega^\star$ verify a Diophantine condition for the constant γ and $\tau = n$, we get:

$$|\langle k, \omega \rangle| \geq \frac{\gamma}{2|k|^n}, \quad \forall 0 < |k| \leq K_0$$

To remove the factor 2 arising from this inequality, we modified the constant C_0 of Rüssman theorem, multiplying it by 2.

The goal at each step of the KAM theorem is to make a change of variables that will decrease the norm of the perturbation. However, instead of trying to make the value of ϵ_0 decrease directly, we will consider the ratio $E_0 = \frac{\epsilon_0}{\gamma r \sigma^\nu}$. If we let r, and σ decrease in a polynomial way, but that we managed to make E decrease exponentially, then ϵ_0 will as well decrease exponentially.

First estimates

As introduced in the outline of the proof, we will switch from the perturbation P to another perturbation R in two steps. First, consider the linearization Q of P in the actions around the origin, secondly, truncate the Fourier series of Q at order K_0. R is a trigonometric polynomial on the angles and affine in the action.
The size of the perturbation, by our assumptions, is smaller than ϵ_0 on the set $D_{r,s} \times O_{h_0}$. Let us start by bounding the linearized function Q:

$$\begin{aligned}
\|Q\|_{\frac{3r}{4},s,h_0} &\leq \|P\|_{\frac{3r}{4},s,h_0} + \frac{3r}{4}\|P_I\|_{\frac{3r}{4},s,h_0} \\
&\leq \|P\|_{r,s,h_0} + \frac{3r}{4}\frac{\|P\|_{r,s,h_0}}{r/4} \\
&\leq 4\epsilon_0.
\end{aligned}$$

Now let $\eta = \sqrt{\frac{\epsilon_0}{\gamma r \sigma^\nu}}$. Using the definition of ϵ and the fact that $\epsilon_0 \leq \epsilon$, we obtain $8\eta \leq 1$. From the Taylor expansion formula and Cauchy's inequality, we get:

$$\begin{aligned}
\|P - Q\|_{2\eta r,s,h_0} &\leq (2\eta r)^2 \frac{4\|P\|_{r,s,h_0}}{(1-2\eta)^2 r^2} \\
&\leq 32\eta^2\epsilon_0.
\end{aligned}$$

With lemma B.1, we obtain the following estimates on the difference between Q and its truncation at the order K_0:

$$\begin{aligned}
\|R - Q\|_{\frac{3r}{4},s-\sigma,h_0} &\leq 4^n n! K_0^n \exp(-K_0\sigma)\|Q\|_{\frac{3r}{4},s,h_0} \\
&\leq 4^\nu n! K_0^n \exp(-K_0\sigma)\epsilon_0.
\end{aligned}$$

Hence:

$$\begin{aligned}
\|R\|_{\frac{3r}{4},s-\sigma,h_0} &\leq \|R - Q\|_{\frac{3r}{4},s-\sigma,h_0} + \|Q\|_{\frac{3r}{4},s,h_0} \\
&\leq (4 + 4^\nu n! K_0^n e^{-K_0\sigma})\epsilon_0.
\end{aligned}$$

Since $K_0 \in A \cap B_+$, we have the following inequality:

$$\begin{aligned} 4^\nu n! K_0^n e^{-K_0\sigma} &= 4^\nu n! \frac{2K_0^{n+\nu}\sigma^\nu e^{-K_0\sigma}}{2K_0^\nu\sigma^\nu} \\ &\leq \frac{1}{\delta}\frac{4^\nu n!}{2((2n+\nu)\log 2)^\nu} \leq 1. \end{aligned}$$

Indeed, the term on the right depends only on n, and is decreasing with this variable. Since it takes a value less than 1 for $n = 1$, the result follows directly. Hence, we have

$$\|R\|_{\frac{3r}{4},s-\sigma,h_0} \leq 5\epsilon_0.$$

Solving the cohomological equation

We would now like to solve the equation (3.3) in F. Letting $\hat{N} = N_+ - N$, we can write:

$$\{F, N\} + \hat{N} = R.$$

Recall that in the outline, we wanted $\hat{N} = \bar{R} = \frac{1}{(2\pi)^n}\int_{\mathbb{T}^n} R d\theta$. With the work done previously, we get the following bound on $\hat{N}$:

$$\|\hat{N}\|_{\frac{3r}{4},h_0} \leq \|R\|_{\frac{3r}{4},s-\sigma,h_0} \leq 5\epsilon_0.$$

Since the Fourier series of R only contains terms of indices k such that $|k| \leq K_0$, we can apply the corollary 2.12 of Rüssman's theorem and solve the remainder of the cohomological equation (3.3). The norm of the Hamiltonian F solving this equation therefore satisfies:

$$\|F\|_{\frac{3r}{4},s-2\sigma,h_0} \leq \frac{C_0\|R\|_{\frac{3r}{4},s-\sigma,h_0}}{\gamma\sigma^n} \leq \frac{5C_0\epsilon_0}{\gamma\sigma^n}.$$

Hence, with Cauchy's inequality, we get:

$$\begin{aligned} \|F_\theta\|_{\frac{r}{2},s-3\sigma,h_0} &\leq \frac{5C_0\epsilon_0}{\gamma\sigma^\nu}, \\ \|F_I\|_{\frac{r}{2},s-3\sigma,h_0} &\leq \frac{20C_0\epsilon_0}{\gamma r\sigma^n}. \end{aligned}$$

Estimates on the transformation Φ

After obtaining the estimates on the derivatives of F, we can deduce some estimates on the vector field associated to F, and then on the time-1 map Φ. On the domain $D_{\frac{r}{2},s-3\sigma,h_0}$, using the value of ϵ_0 and η, we get:

$$\begin{aligned} \|F_\theta\|_{\frac{r}{2},s-3\sigma,h_0} &\leq \sqrt{\frac{\epsilon_0}{\gamma r\sigma^\nu}}\sqrt{\frac{\epsilon_0}{\gamma r\sigma^\nu}}5C_0 r \\ &\leq \frac{5C_0}{4^\nu(200nC_0 + 32 + 4^\nu n!)}\eta r \\ &\leq \eta r \leq \frac{r}{8}, \\ \|F_I\|_{\frac{r}{2},s-3\sigma,h_0} &\leq \frac{20C_0}{4^{2\nu}(200nC_0 + 32 + 4^\nu n!)^2}\sigma \leq \sigma. \end{aligned}$$

With these two inequalities, the time-1 map Φ is well-defined on the domains:

$$\begin{aligned} \Phi = \Phi^t|_{t=1} &: D_{\frac{r}{4},s-4\sigma,h_0} \longrightarrow D_{\frac{r}{2},s-3\sigma,h_0}, \\ \Phi = \Phi^t|_{t=1} &: D_{\eta r,s-5\sigma,h_0} \longrightarrow D_{2\eta r,s-4\sigma,h_0}. \end{aligned}$$

Only considering the first domain is not enough to prove the KAM step. Indeed, the estimates of the difference $P - Q$ require to lose a lot of analyticity on the actions to keep this term small. Writing $\Phi = (U, V)$, since F is linear in the actions, V is independent of I. The Jacobian of F is therefore

$$\Phi' = \begin{pmatrix} U_I & U_\theta \\ 0 & V_\theta \end{pmatrix}.$$

On the set $D_{\frac{r}{8},s-5\sigma,h_0}$, and hence on the set $D_{\eta r,s-5\sigma,h_0}$, the following inequalities are verified

$$\|U - Id\| \leq \|F_\theta\| \leq \frac{5C_0\epsilon_0}{\gamma\sigma^\nu}, \quad \|V - Id\| \leq \|F_I\| \leq \frac{20C_0\epsilon_0}{\gamma r\sigma^n},$$

$$\|U_I - Id\| \leq \frac{40C_0\epsilon_0}{\gamma r\sigma^\nu}, \qquad \|U_\theta\| \leq \frac{5C_0\epsilon_0}{\gamma\sigma^{\nu+1}},$$

$$\|V_\theta - Id\| \leq \frac{20c\epsilon_0}{\gamma r\sigma^\nu},$$

whence the estimates on Φ in the proposition.

Estimates on the new perturbation

After the transformation, the Hamiltonian takes the form $H = N^+ + P^+$. We will now work on a bound on the norm of P^+. First, consider $\{R, F\}$:

$$\begin{aligned}
\|\{R,F\}\|_{\frac{r}{2},s-3\sigma,h_0} &\leq n(\|R_I\|_{\frac{r}{2},s-3\sigma,h_0}\|F_\theta\|_{\frac{r}{2},s-3\sigma,h_0} + \|F_I\|_{\frac{r}{2},s-3\sigma,h_0}\|R_\theta\|_{\frac{r}{2},s-3\sigma,h_0}) \\
&\leq n\left(\frac{20\epsilon_0}{r}\frac{5C_0\epsilon_0}{\gamma\sigma^\nu} + \frac{5\epsilon_0}{\sigma}\frac{20C_0\epsilon_0}{\gamma r\sigma^n}\right) \\
&\leq 200nC_0\frac{\epsilon_0^2}{\gamma r\sigma^\nu}
\end{aligned}$$

The same inequality remains true for $\{\hat{N}, F\}$. Hence:

$$\begin{aligned}
\left\|\int_0^1 \{(1-t)\hat{N} + tR, F\} \circ X_F^t dt\right\|_{\eta r,s-5\sigma,h_0} &\leq \left\|\{(1-t)\hat{N} + tR, F\}\right\|_{\frac{r}{2},s-4\sigma,h_0} \\
&\leq 200nC_0\frac{\epsilon_0^2}{\gamma r\sigma^\nu}.
\end{aligned}$$

It remains to find the bound of the term induced by $P - R$:

$$\begin{aligned}
\|(P-R)\circ\Phi\|_{\eta r,s-5\sigma,h_0} &\leq \|P-R\|_{2\eta r,s-4\sigma,h_0} \\
&\leq \|P-Q\|_{2\eta r,s-4\sigma,h_0} + \|Q-R\|_{2\eta r,s-4\sigma,h_0} \\
&\leq (32\eta^2 + 4^\nu n! K^n e^{-K\sigma})\epsilon_0.
\end{aligned}$$

Finally, the estimate on the norm of P^+ is the following:

$$\|P^+\|_{\eta r,s-5\sigma,h_0} \leq 200nC_0\frac{\epsilon_0^2}{\gamma r\sigma^\nu} + (32\eta^2 + 4^\nu n! K^n e^{-K\sigma})\epsilon_0 = \epsilon_0^+. \tag{3.9}$$

Exponential decrease

As we explained before, we are interested in the exponential decrease of the ratio $E = \frac{\epsilon_0}{\gamma r\sigma^\nu}$. Let us already think about the iterative step, and choose the variables we will use at the next step.

First, let $\sigma^+ = \sigma/2$, $h^+ = h_0/4^\nu$ and $K^+ = 4K_0$; regarding the actions, since we have to lose much more analyticity width and we let $r^+ = \eta r$. With these definitions, let us compute E^+:

$$\begin{aligned}
E^+ &= \frac{\epsilon_0^+}{\gamma r^+ \sigma^{+\nu}} = \frac{2^\nu \epsilon_0^+}{\gamma \eta r \sigma^\nu} \\
&\leq 200nC_0 \frac{2^\nu}{\eta} \frac{\epsilon_0^2}{(\gamma r \sigma^\nu)^2} + 2^\nu(32\eta^2 + 4^\nu n! K_0^n e^{-K_0\sigma}) \frac{\epsilon_0}{\gamma \eta r \sigma^\nu} \\
&\leq 2^\nu 200nC_0 \frac{E^2}{\eta} + (32\eta^2 + 8^\nu n! K_0^n e^{-K_0\sigma}) \frac{E}{\eta}.
\end{aligned}$$

Using the fundamental inequality (3.8), we have $2^\nu E_0 \geq K_0^n e^{-K_0\sigma}$. Observe as well that $E = \eta^2$, hence:

$$E^+ \leq 2^\nu(200nC_0 + 32 + 8^\nu n!)E^{\frac{3}{2}} = \sqrt{C_1} E^{\frac{3}{2}}, \tag{3.10}$$

i.e. $C_1 E^+ \leq (C_1 E)^{\frac{3}{2}}$. The scheme converges exponentially fast if $E < C_1^{-1}$, therefore if $\epsilon_0 \leq \frac{\gamma r \sigma^\nu}{C_1}$. The initial condition on ϵ, and on $\epsilon_0 \leq \epsilon$ shows that we have even better: $C_1 E \leq \frac{1}{4^\nu}$, hence the exponential decrease of E.

Changing the frequencies

It remains to deal with the function φ, which controls the frequency shift when adding the mean of the linearized perturbation over the angles. We will use lemma B.2 to make explicit the domain on which this map is well-defined. Let $v = \hat{N}_I = [R_I]$, the new frequency vector is $\omega^+ = \omega + v(\omega)$. Computing the norm of v gives:

$$\begin{aligned}
\|v\|_{h_0} = \|N_I\|_{h_0} &\leq \frac{5}{\frac{3r}{4} - \eta} \epsilon_0 \leq 10 \frac{\epsilon_0}{r} \\
&\leq \frac{10 h_0}{\delta} \\
&\leq \frac{h_0}{4}
\end{aligned}$$

Applying lemma B.2, we obtain the inverse map $\varphi : O_{\frac{h_0}{4}} \to O_{h_0}$, $\omega^+ \to \omega$, satisfying:

$$\begin{aligned}
\|\varphi - Id\|_{\frac{h_0}{4}} &\leq \frac{10\epsilon_0}{r} \\
\|D\varphi - Id\|_{\frac{h_0}{4}} &\leq \frac{5\epsilon_0}{2h_0 r}
\end{aligned}$$

In this configuration, we let $N^+ = (N + \hat{N}) \circ \varphi$, and we obtained the new Hamiltonian $H^+ = N^+ + P^+$.

This ends the proof of the KAM step, that we now have to iterate.

3.3 Quantitative KAM Theorem with parameters

In this section, we give the explicit statement of the KAM theorem. Recall that we chose $\tau = n$ in the computation. We will quickly remind all the definitions we did to gather everything we need here.

3.3.1 Statement of the quantitative KAM theorem

Let $r, s, h > 0$, $\delta \geq 40$, $\sigma = s/20$, and $K' = \min(A \cap B^+)$, where

$$A = \{K \in \mathbb{N} : K\sigma \geq (2n+\nu)log\ 2\}, \tag{3.11}$$

$$B = \left\{K \in \mathbb{N} : 2K^{n+\nu}\sigma^\nu e^{-K\sigma} \leq \frac{1}{\delta}\right\}, \tag{3.12}$$

$$B_+ = \{K \in B : \ \forall m \in \mathbb{N},\ K+m \in B\}. \tag{3.13}$$

Call $\nu = n+1$, $C_1 = 4^\nu(200nC_0 + 32 + 8^\nu n!)^2$ with $C_0 = 3\pi 6^{\frac{n}{2}}\frac{\sqrt{n(2n)!}}{2^n}$. Recall the definition of ϵ in equation (3.7):

$$\epsilon = \min\left(\frac{\gamma r\sigma^\nu}{4^\nu C_1}, \frac{hr}{\delta}, \frac{\gamma r}{2K'^\nu\delta}\right). \tag{3.14}$$

Besides, define the following (exponentially convergent) series for $\nu \geq 1$:

$$S_\nu = \sum_{i=0}^{\infty} 2^{\nu\left(3i+2-\left(\frac{3}{2}\right)^i\right)}, \tag{3.15}$$

$$T_\nu = \sum_{i=0}^{\infty} 2^{(2\nu+1)i-\nu\left(\frac{3}{2}\right)^i}. \tag{3.16}$$

Finally, define

$$\mu = \exp\left(\frac{5}{\delta}\right), \tag{3.17}$$

$$\xi = \exp\left(10C_0\frac{\epsilon_0}{\gamma r_0\sigma_0^\nu}\right). \tag{3.18}$$

Theorem 3.3. *Let $H = N + P$ a Hamiltonian, such that P is real analytic on the set $D_{r,s} \times O_h$ and $\|P\|_{r,s,h} = \epsilon_0 \leq \epsilon$. Then there exists a Lipschitz continuous map $\varphi : \Gamma_\gamma \to \Omega_\gamma^{h_0}$, with $h_0 = \frac{\delta\epsilon_0}{r}$, and a Lipschitz continuous family of real analytic torus embeddings $\Phi : \mathbb{T}^n \times \Gamma_\gamma \to B \times \mathbb{T}^n$ close to Φ_0 such that for each $\omega \in \Gamma_\gamma$, the embedded tori are Lagrangian and*

$$X_H|_{\varphi(\omega)} \circ \Phi = \Phi' \cdot X_N.$$

Φ is real analytic on the set $\{\theta : |\Im\theta| < s/2\}$ for each ω, and the following inequalities on Φ and φ hold:

$$\|W(\Phi - \Phi_0)\| \leq 6n\xi \log\xi \tag{3.19}$$

$$\|\varphi - Id\| \leq 4h_0 n\mu \log\mu, \tag{3.20}$$

$$\tag{3.21}$$

where $W = diag(r^{-1}Id, s^{-1}Id)$. As for their Lipschitz constant, we have:

$$\|W_0(\Phi - \Phi_0)\|_L \leq 4^\nu 80\frac{nC_0}{\gamma\sigma^\nu\delta}T_\nu \tag{3.22}$$

$$\|\varphi - Id\|_L \leq 4nS_\nu\mu\log\mu. \tag{3.23}$$

3.3.2 Proof of the theorem

Soundness of the iteration

To iterate the KAM step, the hypotheses at step $j+1$ need to be fulfilled knowing that they are at a step j. We will therefore use the new value of each variables obtained after one KAM step, and check if they satisfy the hypotheses of the KAM step. Recall that after a step j, we have: $K_{j+1} = 4K_j$, $\sigma_{j+1} = \sigma_j/2$, $\eta_j = \sqrt{\frac{\epsilon_j}{\gamma r_j \sigma_j^\nu}}$, $r_{j+1} = \eta_j r_j$, $h_{j+1} = h_j/4^\nu$. The size of the new perturbation has to decrease much faster than the analyticity width in order to obtain the KAM theorem.
It is necessary to notice that the constant δ correlating the value of ϵ_{j+1} to h_{j+1} is not conserved after one step, though it is increasing and still satisfies $\delta \geq 40$. This variable δ fixes the size of the transformation: the larger it is, the smaller are the norms of $\varphi - Id$ and $\Phi - \Phi_0$. Obviously, after one step, since the new perturbation is much smaller, the new transformation will also be smaller and therefore we can let δ increase.
Now let us see for each variable the inequalities that need to be fulfilled:

- The equality $K_{j+1}\sigma_{j+1} = 2K_j\sigma_j$ shows that K_{j+1} belongs again to the set A_{j+1}.

- Let us verify that K_{j+1} belongs to $B_{+,j+1}$ as well:

$$
\begin{aligned}
2K_{j+1}^{n+\nu}\sigma_{j+1}^{\nu}\exp(-K_{j+1}\sigma_{j+1}) &= 2(4^{n+\nu}K_j^{n+\nu})\left(\frac{\sigma_j^\nu}{2^\nu}\right)\exp(-2K_j\sigma_j) \\
&\leq \frac{4^{n+\nu}}{2^\nu}\frac{\exp(-K_j\sigma_j)}{\delta} \\
&\leq 2^{2n+\nu}\frac{\exp(-K_j\sigma_j)}{\delta}.
\end{aligned}
$$

The condition to have $K_{j+1} \in B_{+,j+1}$ is therefore the following:

$$2^{2n+\nu}\exp(-K_j\sigma_j) \leq 1.$$

Since $K_j \in A_j$ it is verified.

- $K_j \in A_j \cap B_j$, we have $K_{j+1} \in A_{j+1} \cap B_{+,j+1}$.

- In the KAM step, we defined the limit value $\epsilon = \epsilon^-$. Define now

$$\epsilon^+ = \min\left(\frac{\gamma r_{j+1}\sigma_{j+1}^\nu}{4^\nu C_1}, \frac{h_{j+1}r_{j+1}}{\delta}, \frac{\gamma}{2K_{j+1}^\nu\delta}\right).$$

With the hypotheses on the variables we made, we have in fact

$$\epsilon^+ = \min\left(\frac{\eta_j}{2^\nu}, \frac{\eta_j}{4^\nu}, \frac{1}{4^\nu}\right)\epsilon^- = \frac{\eta_j\epsilon^-}{4^\nu}$$

ϵ^+ is therefore the new limit of the application of the KAM step. We have to check the condition $\epsilon_{j+1} \leq \epsilon^+$.

$$
\begin{aligned}
\epsilon_{j+1} &= \gamma r_{j+1}\sigma_{j+1}^\nu E_{j+1} \\
&\leq \frac{\eta_j}{2^\nu}\gamma r_j\sigma_j^\nu\sqrt{C_1}E_j^{\frac{3}{2}} \\
&\leq \frac{\eta_j}{2^\nu}\gamma r_j\sigma_j^\nu\sqrt{C_1}\frac{\epsilon_j^{\frac{3}{2}}}{(\gamma r_j\sigma_j^\nu)^{\frac{3}{2}}} \\
&\leq \frac{\eta_j\epsilon_j\sqrt{C_1}}{2^\nu}\sqrt{\frac{\epsilon_j}{\gamma r_j\sigma_j^\nu}}.
\end{aligned}
$$

By assumption, $\epsilon_j \leq \epsilon^- \leq \frac{\gamma r_j \sigma_j^\nu}{4^\nu C_1}$. Hence:

$$\epsilon_{j+1} \leq \frac{\eta_j \epsilon^-}{4^\nu} = \epsilon^+$$

The condition $\epsilon \leq \frac{\gamma r \sigma^\nu}{4^\nu C_1}$, that we imposed in the definition of ϵ, and corresponding to the control of the transformation among the actions and angles, takes its root here, and allows us to iterate the KAM step. The values ϵ^+, ϵ^- represent the worst case possible. At each step, we consider, in a sense, this worst case, and not the real perturbation. Indeed, it could happen that for some perturbation, and some transformation we have $E_{j+1} \ll E_j$ (using the effective value of ϵ_{j+1} and not its estimate), and therefore the inequality $K_{j+1}^n \exp(-K_{j+1}\sigma_{j+1}) \leq 2^\nu E_{j+1}$ would not be fulfilled. In this case, one could skip some steps of the KAM theorem, and would obtain better values for the estimates than in the worst case. That is why in an explicit application, it can be interesting to compute the first steps by hand before applying the KAM theorem.
Going back to the fundamental inequality (3.8), it holds at the step $j+1$, in particular $K_{j+1}^n \exp(-K_{j+1}\sigma_{j+1}) \leq 2^\nu E_{j+1}$. Let us express now E_{j+1} using ϵ^+.

$$\begin{aligned} K_{j+1}^n \exp(-K_{j+1}\sigma_{j+1}) &= 4^n (K_j^n \exp(-K_j\sigma_j)) \exp(-K_j\sigma_j) \\ &\leq 2^\nu \frac{\epsilon^-}{\gamma r_j \sigma_j^\nu} 4^n \exp(-K_j\sigma_j) \\ &\leq 2^\nu \frac{4^\nu \epsilon^+}{\eta_j} \frac{\eta_j}{2^\nu \gamma r_{j+1} \sigma_{j+1}^\nu} 4^n \exp(-K_j\sigma_j) \\ &\leq 2^\nu E_{j+1} 2^{2n+\nu} \exp(-K_j\sigma_j). \end{aligned}$$

Since $K_j \in A_j$, we have: $K_{j+1}^n \exp(-K_{j+1}\sigma_{j+1}) \leq 2^\nu E_{j+1}$.

- It remains to verify that $\frac{\epsilon_j}{h_j r_j}$ is decreasing:

$$\begin{aligned} \frac{\epsilon_{j+1}}{h_{j+1} r_{j+1}} &= \frac{\gamma \sigma_{j+1}^\nu E_{j+1}}{h_{j+1}} \\ &= \frac{2^\nu \gamma \sigma_j^\nu E_{j+1}}{h_j} \\ &= 2^\nu \frac{E_{j+1}}{E_j} \frac{\epsilon_j}{h_j r_j} \\ &\leq 2^\nu \sqrt{C_1 E_j} \frac{\epsilon_j}{h_j r_j} \leq \frac{\epsilon_j}{h_j r_j}. \end{aligned}$$

We checked all the inequalities that needed to be checked in order to iterate the KAM step an indefinite time. The scheme is well defined, and we now need to compute the size of the transformations.

Transformations involved and their estimates

The initial Hamiltonian was under the form $H = N + P$. At each KAM step, we defined two transformations: Φ_j which modifies the action-angle coordinates, and φ_j which modifies the frequencies. We let $s_{j+1} = s_j - 5\sigma_j$, with $s_0 = s$, $r_0 = r$, $\eta_0 = \eta$ and $\eta_j = \sqrt{\frac{\epsilon_j}{\gamma r_j \sigma_j^\nu}} = \sqrt{E_j}$.
Define $\mathcal{F}_0 = Id$, and for $j > 0$:

$$\begin{aligned} \mathcal{F}_{j+1} : D_{j+1} \times O_{j+1} &\to D_j \times O_j \\ (I, \theta, \omega) &\mapsto (\Phi_{j+1}(I, \theta, \omega), \varphi_{j+1}(\omega)), \end{aligned}$$

with

$$D_j = \{I \in \mathbb{C}^n : |I| < r_j\} \times \{\theta \in \mathbb{T}^n : |\Im(\theta)| \leq s_j\},$$
$$O_j = \left\{\omega \in \mathbb{R}^n : |\omega - \Omega_\gamma^\beta| < h_j\right\}.$$

Call $\mathcal{F}^j = \mathcal{F}_0 \circ ... \circ \mathcal{F}_{j-1}$. We then have:

$$\mathcal{F}^j : D_j \times O_j \to D_0 \times O_0$$

Thereafter, we will prove the estimates on the transformation $\mathcal{F}^j$, and show its convergence when j goes to infinity.

Preliminaries: The map $\mathcal{F}$ transforms a torus associated to a frequency vector belonging to the set Γ_γ to a deformed torus where the motion has frequencies belonging to the set Ω_γ^β. The action p_0 on the first torus is entirely and uniquely determined by the frequency vector, using the Legendre transform to relate p_0 and ω. The uniqueness comes from the hypothesis of non-degeneracy of the unperturbed Hamiltonian. In order to be precise, we define the following mapping:

$$\Psi : \mathbb{T}^n \times \Gamma_\gamma \to D \times \mathbb{T}^n.$$

First, define the map:

$$\begin{aligned} \Xi : B \times \mathbb{T}^n \times \Gamma_\gamma &\to D \times \mathbb{T}^n \\ (I, \theta, \omega) &\mapsto (h_p^{-1}(\omega) + I, \theta). \end{aligned}$$

Assume $\Phi : \{0\} \times \mathbb{T}^n \to B \times \mathbb{T}^n$ and $\varphi : \Omega_\gamma^\beta \to \Gamma_\gamma$ exist as a limit of Φ_j and φ_j.
Then, one can define Ψ as follows:

$$\begin{aligned} \Psi : \mathbb{T}^n \times \Gamma_\gamma &\to D \times \mathbb{T}^n \\ (\theta, \omega) &\mapsto \Xi(\Phi(0, \theta), \varphi(\omega)) \end{aligned}$$

The KAM theorem shows that, on $T^\star \times \Gamma_\gamma$, $H \circ \Psi = N'$, where $N' = \lim_{j\to\infty} N_j$.

Estimates on the transformations: In order to simplify the formulas, we introduce the weight matrix $W_j = diag(r_j^{-1} Id, \sigma_j^{-1} Id)$. Recall the size of the transformation obtained previously on the set $D_j \times O_j$:

$$\begin{aligned} \|W_j(\Phi_j - Id)\| &\leq \frac{20 C_0 \epsilon_j}{\gamma r_j \sigma_j^\nu}, \\ \left\|W_j(\Phi_j' - Id) W_j^{-1}\right\| &\leq \frac{40 C_0 \epsilon_j}{\gamma r_j \sigma_j^\nu}, \\ \|\varphi_j - Id\| &\leq \frac{10 \epsilon_j}{r_j}, \\ \left\|\varphi_j' - Id\right\| &\leq \frac{5 \epsilon_j}{2 h_j r_j}. \end{aligned}$$

We can estimate the norm of the difference between to consecutive transformations $\mathcal{F}^j$.

$$\begin{aligned} \left\|W_0(\Phi^{j+1} - \Phi^j)\right\| &= \left\|W_0(\Phi^j \circ \Phi_j - \Phi^j)\right\| \\ &\leq 2n \left\|W_0 (\Phi^j)' W_j^{-1}\right\| \|W_j(\Phi_j - Id)\| \\ &\leq 2n \xi_j \|W_j(\Phi_j - Id)\| \\ &\leq \xi_j \frac{40 n C_0 \epsilon_j}{\gamma r_j \sigma_j^\nu}, \end{aligned}$$

it is well-defined when j goes to infinity if the variable $\xi_j = \|W_0 D\Phi^j W_j^{-1}\|$ does not increase too fast on D_j. Observe that the factor $2n$ comes from the estimates of the matrix product using the supremum norm. In the same way, we compute:

$$\begin{aligned}\left\|h_0^{-1}(\varphi^{j+1}-\varphi^j)\right\| &= \left\|h_0^{-1}(\varphi^j\circ\varphi_j-\varphi^j)\right\| \\ &\leq n\left\|h_0^{-1}(\varphi^j)'h_j^{-1}\right\|\|h_j(\varphi_j-Id)\| \\ &\leq n\mu_j\|h_j(\varphi_j-Id)\| \\ &\leq \mu_j\frac{10n\epsilon_j}{h_jr_j},\end{aligned}$$

where again it is necessary to check the increase of $\mu_j = \left\|h_0^{-1}(\varphi^j)'h_j^{-1}\right\|$ on O_j.

On $D_j \times O_j$, we have in fact $(\Phi^j) = (\Phi_0)' \cdot ... \cdot (\Phi_{j-1})'$, where the differentials are estimated at different points, that are not important to make explicit as we have a bound on their whole set of definition. With the decrease of the variables r and σ, we get $\left\|W_jW_{j+1}^{-1}\right\| \leq 1/2$ (the sup coming from the factor σ_{j+1}/σ_j). Hence,

$$\begin{aligned}\xi_j &= \left\|W_0(\Phi^j)'W_j^{-1}\right\| = \left\|W_0(\Phi_0)'\cdot ... \cdot(\Phi_{j-1})'W_j^{-1}\right\| \\ &\leq \left\|W_0(\Phi_0)'W_0^{-1}\right\|2n\left\|W_0W_1^{-1}\right\|2n\times ...\times 2n\left\|W_{j-1}(\Phi_{j-1})'W_{j-1}^{-1}\right\|2n\left\|W_{j-1}W_j^{-1}\right\| \\ &\leq (2n)^{2j}\left(\frac{1}{2}\right)^j\prod_{i=1}^{j}\left(1+\frac{40C_0\epsilon_j}{\gamma r_j\sigma_j^\nu}\right) \\ &\leq (2n^2)^j\prod_{i=1}^{j}\left(1+\frac{40C_0\epsilon_j}{\gamma r_j\sigma_j^\nu}\right).\end{aligned}$$

Indeed, we have a product of $2j$ matrices, and the presence of j matrices of the form $W_jW_{j+1}^{-1}$, hence the factor $(2n^2)^j$. Likewise, for μ_j:

$$\begin{aligned}\mu_j &\leq \left(\frac{n^2}{4^\nu}\right)^j\prod_{i=1}^{j}\left(1+\frac{5\epsilon_j}{2h_jr_j}\right) \\ &\leq \prod_{i=1}^{j}\left(1+\frac{5\epsilon_j}{2h_jr_j}\right).\end{aligned}$$

This time we can get rid of the factor depending on n because of the factor 4^ν between h_j and h_{j+1}.

Since the variables ϵ_j decrease exponentially fast towards 0, and that the terms h_j and r_j do not decrease as fast, the products in the formulas will converge when j tends to infinity. We can bound them using the estimates we obtained in the KAM step.

First, recall that $\dfrac{40C_0\epsilon_j}{\gamma r_j\sigma_j^\nu} < 1$, whence, using the logarithm for $j \geq 1$:

$$\begin{aligned}\log\left((2n^2)^{-j}\xi_j\right) &\leq \sum_{i=1}^{j}\log\left(1+\frac{40C_0\epsilon_i}{\gamma r_i\sigma_i^\nu}\right) \\ &\leq \sum_{i=1}^{j}\frac{40C_0\epsilon_i}{\gamma r_i\sigma_i^\nu} \\ &\leq \sum_{i=1}^{j}40C_0E_i.\end{aligned}$$

Using the exponential decrease of E_j:

$$\begin{aligned} E_j &\leq \sqrt{C_1} E_{j-1}^{\frac{3}{2}} \leq \ldots \leq C_1^{\frac{1}{2}\sum_{i=0}^{j-1}\left(\frac{3}{2}\right)^i} E_0^{\left(\frac{3}{2}\right)^j} \\ &\leq (C_1 E_0)^{\left(\frac{3}{2}\right)^j - 1} E_0 \\ &\leq 4^{-\nu\left(\frac{3}{2}\right)^j + \nu} E_0. \end{aligned}$$

Finally:

$$\begin{aligned} (2n^2)^{-j}\xi_j &\leq \exp\left(\sum_{i=1}^{\infty} 40 C_0 E_0 4^{\nu - \nu\left(\frac{3}{2}\right)^i}\right) \\ &\leq \exp\left(40 C_0 E_0 \sum_{i=1}^{\infty} 4^{\nu-\nu\left(\frac{3}{2}\right)^i}\right) \\ &\leq \exp(10 C_0 E_0) = \exp\left(10 C_0 \frac{\epsilon_0}{\gamma r_0 \sigma_0^{\nu}}\right) \equiv \xi. \end{aligned}$$

In the same way, we get for μ_j:

$$\begin{aligned} \mu_j &\leq \exp\left(\frac{5}{2}\sum_{i=1}^{\infty}\frac{\epsilon_i}{r_i h_i}\right) = \exp\left(\frac{5}{2}\sum_{i=1}^{\infty}\frac{\gamma E_i \sigma_i^{\nu}}{h_i}\right) \\ &\leq \exp\left(\frac{5}{2}\sum_{i=1}^{\infty}\frac{\gamma E_i \sigma_0^{\nu} 2^{\nu i}}{h_0}\right) \\ &\leq \exp\left(\frac{5}{2}\frac{\gamma\sigma_0^{\nu}}{h_0}\sum_{i=1}^{\infty} E_0 2^{\nu i - 2\nu\left(\frac{3}{2}\right)^i + 2\nu}\right) \\ &\leq \exp\left(5\frac{\epsilon_0}{r_0 h_0}\right) \leq \exp\left(\frac{5}{\delta}\right) \equiv \mu. \end{aligned}$$

With this computation, we can continue towards our aim of estimating $\mathcal{F}^j$ for all $j \geq 1$.

$$\begin{aligned} \left\| W_0(\Phi^j - \Phi_0)\right\| &\leq \sum_{i=0}^{j-1}\left\| W_0(\Phi^{i+1} - \Phi^i)\right\| \\ &\leq \sum_{i=0}^{j-1} \xi_i \frac{40 n C_0 \epsilon_i}{\gamma r_i \sigma_i^{\nu}} \leq 40 n C_0 \xi \sum_{i=0}^{\infty}\left((2n^2)^i E_i\right) \\ &\leq 40 n C_0 E_0 \xi \sum_{i=0}^{\infty}\left((2n^2)^i 4^{-\nu\left(\frac{3}{2}\right)^i + \nu}\right) \\ &\leq 60 n C_0 E_0 \xi = 6 n \xi \log \xi \end{aligned}$$

As well, for all $j \geq 1$,

$$\begin{aligned} \left\| h_0^{-1}(\varphi^j - Id)\right\| &\leq \sum_{i=0}^{j-1}\left\| h_0^{-1}(\varphi^{i+1} - \varphi^i)\right\| \\ &\leq \sum_{i=0}^{j-1}\mu_i \frac{10 n \epsilon_i}{h_i r_i} \\ &\leq 4 n \mu \log \mu. \end{aligned}$$

Therefore, with these uniform bounds, we can let j go to infinity. The transformation $\mathcal{F}$ is well-defined on $T^\star \times \Gamma_\gamma$.
The set Ω_γ^β, defined while constructing φ^j, depends on Γ_γ and on h_0. More precisely, recall that for all $\omega' \in \Omega_\gamma^\beta$, there exists $\omega \in \Gamma_\gamma$ such that $|\omega - \Omega_\gamma^\beta| < h_0$. Therefore, we can let $\beta = h_0$ so that the set $O_h \subset \Omega$.

First conclusion on the transformation: Before obtaining the Lipschitz norm of the transformation, we are going to draw some conclusion on the transformation we built. First, we have the relation $H \circ \Xi \circ \mathcal{F}^j - N^j = P^j$ on the set $D_j \times O_j$ for all $j \geq 1$. With this equality, we can describe the difference between the vector field associated to H and the one associated to N^j. Formally, by derivation on the action-angle coordinates, we have

$$^t(H \circ \Xi \circ \mathcal{F}^j - N^j)' = {}^t(\Xi \circ \Phi^j)' \cdot \nabla H \circ \Xi \circ \mathcal{F}^j - \nabla N^j.$$

For a fixed ω, the map Ξ is constant and linear in the actions and the angles, which simplifies the computation. Therefore, using Cauchy's inequality on the action coordinates, we can bound the previous derivative.

$$\left\| {}^t(\Phi^j)' \cdot \nabla H \circ \Xi \circ \mathcal{F}^j - \nabla N^j \right\|_{0,s_j-\sigma_j} \leq \max\left(\frac{\epsilon_j}{r_j}, \frac{\epsilon_j}{\sigma_j}\right).$$

Using the weighted matrix W_j which was useful to give the estimates on $(\Phi^j)'$, and the symplectic matrix J,

$$J = \begin{pmatrix} 0 & Id \\ -Id & 0 \end{pmatrix},$$
$$W_j \times J = \begin{pmatrix} 0 & \sigma_j^{-1} \\ -r_j^{-1} & 0 \end{pmatrix},$$

and finally multiplying on the left by the latter matrix our relation, we obtain:

$$\|W_j \times J \times {}^t(\Phi^j)' \cdot \nabla H \circ \Xi \circ \mathcal{F}^j - W_j \times J \times \nabla N^j\|_{0,s_j-\sigma_j} \leq \frac{\epsilon_j}{r_j \sigma_j}$$

The map Φ^j being symplectic, it satisfies $J^t\Phi' = (\Phi')^{-1}J$. Hence, multiplying the last inequality by $W_0\Phi' W_j^{-1}$, we get:

$$\begin{aligned} \left\| W_0 \left(J\nabla H \circ \Xi \circ \mathcal{F}^j - \Phi' J \nabla N^j \right) \right\|_{0,s_j-\sigma_j} &\leq \|W_0 \Phi' W_j^{-1}\|_{0,s_j-\sigma_j} \frac{\epsilon_j}{r_j\sigma_j} \\ &\leq (2n^2)^j \xi \frac{\epsilon_j}{r_j \sigma_j} \end{aligned}$$

Looking at the vector field, this inequality becomes:

$$\|W_0\| \times \|X_H \circ \Xi \circ \mathcal{F}^j - (\Phi^j)' \cdot X_N\|_{0,s_j-\sigma_j} \leq (2n^2)^j \xi \frac{\epsilon_j}{r_j\sigma_j}$$

We can let j tend to infinity, and finally, on the set $\mathbb{T}^\star \times \Gamma_\gamma$, we get the equality of these two vector fields, *i.e.*, for some $\omega \in \Gamma_\gamma$, with $N = (e(\omega) + \langle \omega, I \rangle)$,

$$X_H \circ \Psi = \Phi' \cdot X_N \tag{3.24}$$

Observe that we wrote Ψ instead of Φ, it expresses the fact that we consider the "origin" of the action I at the point $\varphi(\omega)$.

Lipschitz norm of the transformation On the Cantor set $\Omega_\gamma^{h_0} \subset \Omega$, the formulas of the derivatives of $(\varphi^j - Id)$ converge. Indeed, the existence of a uniform constant for $\xi_j/(2n^2)^j$ and μ_j, and the exponentially fast convergence of the terms $\frac{\epsilon_j}{r_j h_j^m}$ for all m implies that the norm of every derivative of the mapping $\mathcal{F}^j$ (more precisely its difference to (Φ_0, Id)) converges, whatever the order of the derivative.

We will compute the Lipschitz norm of the transformation $\mathcal{F}$ to complete the proof of the theorem. First we will evaluate the derivative of the map $\varphi - Id$ with respect to ω. Although the map φ is defined on a Cantor set, it is possible to extend it in a Lipschitz way, and even in a C^1 map, using Whitney's extension theorem (see article [70] or the statement of the theorem B.16 in appendix B). We will not dwell further on these notions, and will just compute an estimate on its norm.

$$\begin{aligned}\|(\varphi^j - Id)'\|_{\frac{h_j}{2}} &\leq \sum_{i=0}^{j-1} \|(\varphi^{i+1})' - (\varphi^i)'\|_{\frac{h_i}{2}} \\ &\leq \sum_{i=0}^{j-1} \|(\varphi^{i+1} - \varphi^i)'\|_{\frac{h_i}{2}} \\ &\leq \sum_{i=0}^{j-1} \frac{2\|\varphi^{i+1} - \varphi^i\|_{h_i}}{h_i} \\ &\leq h_0 \times 20n\mu \sum_{i=0}^{i-1} \frac{\epsilon_i}{r_i h_i^2}\end{aligned}$$

As done before, we can compute this sum and we obtain

$$\begin{aligned}\|(\varphi^j)' - I_n\|_{\frac{h_j}{2}} &\leq \frac{20nh_0\mu}{\delta h_0} \sum_{i=0}^{j-1} 2^{\nu\left(3i+2-\left(\frac{3}{2}\right)^i\right)} \\ &\leq \frac{20n\mu}{\delta} S_\nu.\end{aligned}$$

This function S_ν increases with the value of n (or ν), and therefore we cannot bound it uniformly for all n. Letting j go to infinity, we obtain the Lipschitz norm:

$$\|\varphi - Id\|_L \leq \frac{20n\mu}{\delta} S_\nu \tag{3.25}$$

As for $\Phi - \Phi_0$, computing the Lipschitz estimate exactly in the same way, we get:

$$\|W_0(\Phi - \Phi_0)\|_L \leq 4^\nu 80 \frac{nC_0}{\gamma\sigma^\nu\delta} T_\nu$$

These computations end the proof of the explicit KAM theorem.

3.3.3 Estimates in the initial actions

After giving the estimates on the map $\mathcal{F} = (\Phi, \varphi)$, we estimate the map $\Psi = \Xi \circ \mathcal{F}$. Indeed, these estimates will shift the torus back to its original place around the action p_0.
Consider the difference $\Psi - \Psi_0$. We have:

$$\begin{aligned}(\Psi - \Psi_0)(0, \theta, \omega) = (g'(\varphi(\omega)) - g'(\omega) + (\Phi_1(0, \theta, \varphi(\omega)) - \Phi_{0,1}(0, \theta, \omega)), \\ \Phi_2(0, \theta, \varphi(\omega)) - \Phi_{0,2}(0, \theta, \omega)),\end{aligned}$$

where $\Phi(I, \theta) = (\Phi_1(I, \theta), \Phi_2(I, \theta))$. Hence, the norm of $\Psi - \Psi_0$ on the set $T^\star \times \Gamma_\gamma$ satisfies:

$$\begin{aligned}\|W_0(\Psi - \Psi_0)\| &\leq \frac{1}{r}\|g' \circ \varphi - g'\| + \|W_0(\Phi - Id)\| \\ &\leq n \sup_{\Gamma_\gamma} \|g''\| \frac{\|\varphi - Id\|}{r} + \|W_0(\Phi - Id)\|\end{aligned}$$

$$\|W_0(\Psi - \Psi_0)\| \leq \sup_{\Gamma_\gamma} \|g''\| \times \frac{4n^2 h_0}{r} \mu \log \mu + 6n\xi \log \xi \tag{3.26}$$

We can also compute an estimate of the Lipschitz norm of Ψ with respect to ω. We will only compute one for the first coordinate of Ψ, and the Lipschitz norm will then hold for the second coordinate given the shape of this map. The estimate on the Lipschitz norm of $\Phi - \Phi_0$ being known, we will be interested more precisely in the map $\Upsilon(\omega) = g'(\varphi(\omega)) - g'(\omega)$. Let $\omega, \omega' \in \Gamma_\gamma$, we have:

$$\begin{aligned}|\Upsilon(\omega) - \Upsilon(\omega')| = \Big| \int_0^1 & \left[g''(\omega + t(\varphi - Id)(\omega)) \cdot (\varphi - Id)(\omega)\right] dt - \\ & \int_0^1 \left[g''(\omega' + t(\varphi - Id)(\omega')) \cdot (\varphi - Id)(\omega')\right] dt \Big|.\end{aligned}$$

In order to compute the value of this term, we need to add some intermediate terms under the integral. For simplicity, we use $v = \varphi - Id$, we have

$$\begin{aligned}& g''(\omega + tv(\omega)) \cdot v(\omega) - g''(\omega + tv(\omega)) \cdot v(\omega) = \\ & g''(\omega + tv(\omega)) \cdot \left[v(\omega) - v(\omega')\right] + \left[g''(\omega + tv(\omega)) - g''(\omega' + tv(\omega'))\right] \cdot v(\omega').\end{aligned}$$

The first term of this sum is bounded by

$$|g''(\omega + tv(\omega)) \cdot \left[v(\omega) - v(\omega')\right] | \leq n \sup_{\Gamma_\gamma} |g''| \times |\varphi - Id|_L \times |\omega - \omega'|.$$

For the second term, we write:

$$\begin{aligned}& |g''(\omega + tv(\omega)) - g''(\omega' + tv(\omega'))| \\ & = \left| \int_0^1 g^{(3)}((1-s)(\omega + tv(\omega)) - s(\omega' + tv(\omega'))ds \right| \cdot |(\omega + tv(\omega)) - (\omega' + tv)| \\ & \leq n \sup_{\Gamma_\gamma} \left|g^{(3)}\right| \times (|\omega - \omega'| + t|v(\omega) - v(\omega')|) \\ & \leq n \sup_{\Gamma_\gamma} \left|g^{(3)}\right| \times (1 + t|\varphi - Id|_L) \, |\omega - \omega'|.\end{aligned}$$

Injecting these bounds in the previous inequality, we obtain

$$\begin{aligned}& |\Upsilon - \Upsilon_0|_L \leq \\ & \qquad n \sup_{\Gamma_\gamma} |g''| \times |\varphi - Id|_L + n^2 \sup_{\Gamma_\gamma} \left|g^{(3)}\right| \times |\varphi - Id|_L \times \left(1 + \frac{|\varphi - Id|_L}{2}\right).\end{aligned} \tag{3.27}$$

Finally,

$$|\Psi - \Psi_0|_L \leq |\Upsilon - \Upsilon_0|_L + |\Phi - \Phi_0|_L. \tag{3.28}$$

Chapter 4

Normal form of the secular Hamiltonian

Following our work in the three previous chapter, we can now try to determine the expression of the unperturbed Hamiltonian to which we want to apply the KAM theorem. The degeneracy of the Kepler Hamiltonian requires to include some terms of the perturbation into the unperturbed Hamiltonian, to allow the frequency map to be a diffeomorphism.
In this aim, the first step consists in considering the mean of the perturbation over the fast angles. This idea dates back to Lagrange [16] and Laplace [17], when trying to understand the slow motion of the ellipses in the solar system. The development in power of the eccentricities of the perturbation depends only in terms of even powers; however it depends on the angle of the periapses. Different computations of this series can be found in [58, 41, 44, 18, 13, 45]... Applying a rotation on the Cartesian coordinates of Poincaré variables, one can remove the dependency in g in the term of order two in eccentricities. Then, using the Birkhoff Normal Form theorem of chapter 2, one can put the Hamiltonian under a normal form up to the order four in these variables. Choosing to truncate this Hamiltonian at order 5 in eccentricities, the terms of higher order can be considered as part of the perturbation for small enough eccentricities. The Kepler Hamiltonian plus the part of the perturbation obtained by this scheme is now non-degenerate on a specific domain (yet to be determined). A last averaging is then necessary to make the perturbation even smaller. With this work, we are able to quantify the bound on the norm of the final perturbation. Lastly, we determine the analyticity width for the frequencies, which is essential to apply the KAM theorem we developed in chapter 3.

4.1 The secular part of the perturbation expressed in eccentricities

In this section, the goal is to determine a Birkhoff's normal form up to the order four in eccentricities, so as to obtain a normal form for the unperturbed Hamiltonian of order 2 in the action variables.

4.1.1 Expression in term of the true anomaly

In order to compute the normal form of the Hamiltonian, we need to integrate the perturbation over the fast angles, which are the mean longitudes. Therefore, we will first try to exchange the expression of our integral over the mean anomalies l_i for an integral over the true anomalies v_i, so as to make it easier to compute.

Indeed, in terms of the true anomaly v, the distance to the star $|Q_i|$ has the following form:

$$|Q_i| = \frac{a_i(1-e_i^2)}{1+e_i\cos v_i} = \frac{\Lambda_i^2}{G_{grav}M_i\mu_i^2} \times \frac{1-e_i^2}{1+e_i\cos v_i}$$

The simplicity of this expression makes it relevant to try to integrate over the true anomalies. However, we need to be careful with the change of variables during the integration.
Recall the relation between the true anomaly v and the eccentric anomaly u:

$$\begin{cases} v &= 2\arctan\left(\sqrt{\frac{1+e}{1-e}}\tan\left(\frac{u}{2}\right)\right), \\ u &= 2\arctan\left(\sqrt{\frac{1-e}{1+e}}\tan\left(\frac{v}{2}\right)\right). \end{cases}$$

Since we aim to a BNF up to the order four in eccentricities, we need to expand every formula on these variables. The Taylor expansion to the fourth order of the preceding formulas are:

$$v = u + e\sin u + \frac{e^2}{4}\sin(2u) + \frac{e^3}{12}(3\sin u + \sin(3u)) + \frac{e^4}{32}(4\sin(2u) + \sin(4u)) + o(e^5)$$
$$u = v - e\sin v + \frac{e^2}{4}\sin(2v) - \frac{e^3}{12}(3\sin v + \sin(3v)) + \frac{e^4}{32}(4\sin(2v) + \sin(4v)) + o(e^5)$$

We had chosen the Poincaré variables and hence the mean anomalies, in order to determine a bound on the perturbation. Let us continue with this choice, and call the change of variables between the true and the mean anomaly Φ:

$$\begin{aligned} \Phi : \mathbb{T}^2 &\to \mathbb{T}^2 \\ (v_1, v_2) &\to (l_1, l_2) \end{aligned}$$

Using the relation $l = u - e\sin u$, we can give the expansion of the mean anomaly as a function of the true one, again to the order 4 in eccentricity:

$$l = v - 2e\sin v + \frac{3e^2}{4}\sin(2v) - \frac{e^3}{3}\sin(3v) + \frac{e^4}{32}(4\sin(2v) + 5\sin(4v)) + o(e^5)$$

We can now give the relation between the integral over the mean longitude and the true anomaly. Using the function Φ, the fact that the function we want to integrate is 2π-periodic, and that we have a diffeomorphism between the mean anomaly and the true anomaly ($e < 1$), we can write:

$$\begin{aligned} \int_{\mathbb{T}^2} f(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2 &= \int_{\mathbb{T}^2} f(l_1, l_2) dl_1 dl_2 \\ &= \int_{\mathbb{T}^2} f(\Phi(v_1, v_2)) \left|\det \Phi'(v_1, v_2)\right| dv_1 dv_2 \end{aligned}$$

The expression of the determinant of Φ can be computed directly, and its Taylor expansion to the order 4 as well:

$$\begin{aligned} \det \Phi'(v_1, v_2) &= \frac{\sqrt{(1-e_1^2)}\sqrt{(1-e_2^2)}(1+e_1^2)^2(1+e_2^2)^2}{(1+e_1\cos v_1)^2(1+e_2\cos v2)^2} \\ &= \prod_{i=1,2}\Big(1 - 2e_i\cos v_i + \frac{3e_i^2}{2}\sin(2v_i) - e_i^3\cos(3v_i) + \\ &\qquad \frac{e_i^4}{8}(2\sin(2v_i) + 5\sin(4v_i) + o(e_i^5))\Big) \end{aligned}$$

We can now write the expression of the secular part of the perturbation:

$$\frac{1}{4\pi^2}\int_{\mathbb{T}^2} H_{pert}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,e_1,e_2,g_1,g_2)d\lambda_1 d\lambda_2 =$$
$$G_{grav}\mu_1 m_2 \int_{\mathbb{T}^2}\left(\sum_{n\geq 2}\sigma_n P_n(\cos S)\left(\frac{\Lambda_1^2}{G_{grav}M_1\mu_1^2}\frac{1-e_1^2}{1+e_1\cos v_1}\right)^n \times \left(\frac{\Lambda_2^2}{G_{grav}M_2\mu_2^2}\frac{1-e_2^2}{1+e_2\cos v_2}\right)^{-(n+1)}\right)|\det\Phi'(v_1,v_2)|dv_1dv_2$$

Rearranging the terms, we obtain the formula we will be working with:

$$\begin{aligned}\frac{1}{4\pi^2}\int_{\mathbb{T}^2} H_{pert}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,e_1,e_2,g_1,g_2)d\lambda_1 d\lambda_2 =\\ G_{grav}^2\sum_{n\geq 2}\sigma_n\frac{(1-e_1^2)^{n+1/2}}{(1-e_2^2)^{(n+1/2)}}(1+e_1^2)^2(1+e_2^2)^2\frac{\Lambda_1^{2n}}{\Lambda_2^{2(n+1)}}\frac{(m_0+m_1)^{3n+1}}{(m_0+m_1+m_2)^{n+1}}\frac{m_2^{2n+3}}{(m_0m_1)^{2n-1}}\\ \times\frac{1}{4\pi^2}\int_{\mathbb{T}^2}P_n(\cos(v_2-v_1+g_2-g_1))\frac{(1+e_2\cos v_2)^{n-1}}{(1+e_1\cos v_1)^{n+2}}dv_1dv_2+o(e^5)\end{aligned} \tag{4.1}$$

Depending on the context, it can be easier to work with or without the Taylor expansion in eccentricity.

In the case of a small ratio of the semi-major axes, it can be enough to work with the first terms of the development of the perturbation in powers of the semi-major axes. In this case, it is simpler to integrate over other angles, namely (u_1, v_2), so as to work with trigonometric polynomials of the angles only, and not with algebraic fractions. For $n \leq 5$, the computation (consult Appendix A of [43]) gives:

$$\frac{1}{4\pi^2}\int_{\mathbb{T}^2} H_{pert}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,e_1,e_2,g_1,g_2)d\lambda_1 d\lambda_2 = G_{grav}^2\sum_{n\geq 2}A_n$$

The formulas for $n \leq 5$ are the following:

$$\begin{cases} A_2 = \dfrac{\Lambda_1^4}{\Lambda_2^6}\dfrac{(m_0+m_1)^7}{(m_0+m_1+m_2)^3}\dfrac{m_2^7}{(m_0m_1)^3}\dfrac{2+3e_1^2}{8(1-e_2^2)^{3/2}}\\[2ex] A_3 = -\dfrac{\Lambda_1^6}{\Lambda_2^8}\dfrac{(m_0+m_1)^9}{(m_0+m_1+m_2)^4}\dfrac{m_2^9}{(m_0m_1)^5}(m_0-m_1)\dfrac{15}{64}\dfrac{e_1e_2(4+3e_1^2)}{(1-e_2^2)^{5/2}}\cos g\\[2ex] A_4 = \dfrac{\Lambda_1^8}{\Lambda_2^{10}}\dfrac{(m_0+m_1)^{10}}{(m_0+m_1+m_2)^5}\dfrac{m_2^{11}}{(m_0m_1)^7}(m_0^3+m_1^3)\\ \qquad\times\dfrac{9}{1024}\dfrac{(8+40e_1^2+15e_1^4)(2+3e_2^2)+70e_1^2e_2^2(2+e_1^2)\cos(2g)}{(1-e_2)^{7/2}}\\[2ex] A_5 = -\dfrac{\Lambda_1^{10}}{\Lambda_2^{12}}\dfrac{(m_0+m_1)^{13}}{(m_0+m_1+m_2)^6}\dfrac{m_2^{13}}{(m_0m_1)^9}(m_0-m_1)(m_0^2+m_1^2)\times\dfrac{105}{4096}e_1e_2\cos g\\ \qquad\times\dfrac{8(8+20e_1^2+5e_1^4)+(48+64e_1^2+9e_1^4)e_2^2+14e1^2e_2^2(8+3e_1^2)\cos(2g)}{(1-e_2^2)^{9/2}} \end{cases}$$

It is a known fact that for the even power of semi-major axis (and therefore A_n for even n), only squares of the eccentricities will appear. In case n is odd, we can see that mixed terms appears multiplied by the cosine of the difference between the angles of the perihelion for each bodies, i.e. terms in $e_1^k e_2^l \cos(g)$, where $g = g_2 - g_1$. Since we want an unperturbed Hamiltonian depending only on the action variables, it will be necessary to make a first change of variables to remove this angle in the frequencies (which appears already for $n = 3$), and then to use a BNF theorem to remove all the dependency in the angle of the unperturbed Hamiltonian we want to consider.

4.1.2 Development in eccentricities

In case we need more terms than just the first five in the development in semi major axis, we can expand the formula (4.1) in terms of eccentricity using the Taylor expansion of Φ written above. The part of the development we are interested in is:

$$
\begin{aligned}
\frac{1}{4\pi^2}\int_{\mathbb{T}^2} & H_{pert}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,e_1,e_2,g_1,g_2)d\lambda_1 d\lambda_2 = \\
& G_{grav}^2 \sum_{n\geq 2} \sigma_n \frac{\Lambda_1^{2n}}{\Lambda_2^{2(n+1)}} \frac{(m_0+m_1)^{3n+1}}{(m_0+m_1+m_2)^{n+1}} \frac{m_2^{2n+3}}{(m_0 m_1)^{2n-1}} \\
& \times \Big(b_{0,0,n} + b_{2,0,n} e_1^2 + b_{0,2,n} e_2^2 + b_{1,1,n} e_1 e_2 \cos g + \\
& b_{4,0,n} e_1^4 + b_{0,4,n} e_2^4 + b_{2,2,0,n} e_1^2 e_2^2 + b_{2,2,2,n} e_1^2 e_2^2 \cos^2 g + \\
& b_{3,1,n} e_1^3 e_2 \cos g + b_{1,3,n} e_1 e_2^3 \cos g + o((e_1,e_2)^6) \Big)
\end{aligned}
$$

To simplify the notation, let us write:

$$
\mathcal{M}_n = G_{grav}^2 \sigma_n \frac{(m_0+m_1)^{3n+1}}{(m_0+m_1+m_2)^{n+1}} \frac{m_2^{2n+3}}{(m_0 m_1)^{2n-1}}
$$

$$
c_{i,j} = \sum_{n\geq 2} \mathcal{M}_n \frac{\Lambda_1^{2n}}{\Lambda_2^{2(n+1)}} \times b_{i,j,n}
$$

We then obtain:

$$
\begin{aligned}
\frac{1}{4\pi^2}\int_{\mathbb{T}^2} & H_{pert}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,e_1,e_2,g_1,g_2)d\lambda_1 d\lambda_2 = \\
& c_{0,0} + c_{2,0} e_1^2 + c_{0,2} e_2^2 + c_{1,1} e_1 e_2 \cos g + \\
& c_{4,0} e_1^4 + c_{0,4} e_2^4 + c_{2,2,0} e_1^2 e_2^2 + c_{2,2,2} e_1^2 e_2^2 \cos^2 g + \\
& c_{3,1} e_1^3 e_2 \cos g + c_{1,3} e_1 e_2^3 \cos g + o((e_1,e_2)^6).
\end{aligned}
$$

Using the expansion we gave before, we can give the exact expression of the $b_{i,j,n}$. Each one corresponds to a specific term in the expansion in eccentricity of the secular Hamiltonian. The expressions before the integration of each one of those terms for $n \geq 2$ are gathered in appendix A.1.1. Each of this expression is a trigonometric polynomials depending on the angles (v_1, v_2, g). These terms can be computed in the following way: using each times changes of variables to remove the dependency in g in the Legendre polynomials, and then recurrence relations, we can put these terms under the form $t_{i,j,n} \times \int_0^{2\pi} P_n(\cos x)dx$, where $t_{i,j,n}$ is a known constant. There only remains to find the value of the latter integral, which is:

$$
\int_0^{2\pi} P_n(\cos x)dx = \begin{cases} 0 \text{ if n is odd} \\ \dfrac{2\pi}{2^{2n}} \dbinom{n}{\frac{n}{2}}^2 \text{ if n is even} \end{cases}
$$

The new formulas after integration are given by formulas in appendix A.1.2, and now the coefficients $c_{i,j}$ are completely known. Their formula is given by a power series of explicit coefficients $b_{i,j}$ in product with a variable $\mathcal{M}_n$ which depends on the masses and the Λ_i. These series are hyper-geometric functions, and one could express their expression with special notations relative to these functions.

4.2 Expression in symplectic coordinates

In this section, we will switch from our previous variables that are not symplectic to Poincaré coordinates. Then, since KAM theorem can be applied to a Hamiltonian in action-angle variables, we will need again to perform a change of variables.

4.2.1 From eccentricities to Poincaré Cartesian coordinates

Recall the formulas for the Poincaré coordinates ξ and η, in term of the eccentricity, the Λ_i, and the angle of the perihelion :

$$\xi_i = \sqrt{2\Lambda_i\left(1-\sqrt{1-e_i^2}\right)}\cos(-g_i),$$
$$\eta_i = \sqrt{2\Lambda_i\left(1-\sqrt{1-e_i^2}\right)}\sin(-g_i).$$

The expression of the eccentricity as a function of these two variables is therefore:

$$e_i^2 = \frac{\xi_i^2+\eta_i^2}{\Lambda_i}\left(1-\frac{\xi_i^2+\eta_i^2}{4\Lambda_i}\right) = 2\frac{\Gamma_i}{\Lambda_i} - \left(\frac{\Gamma_i}{\Lambda_i}\right)^2,$$

where $\Gamma_i = \frac{\xi_i^2+\eta_i^2}{2}$. Thus:

$$e_i = \sqrt{2\frac{\Gamma_i}{\Lambda_i}}\left(1-\frac{\Gamma_i}{4\Lambda_i}+o(\Gamma_i)\right),$$
$$e_1e_2 = 2\sqrt{\frac{\Gamma_1\Gamma_2}{\Lambda_1\Lambda_2}} - \frac{\Gamma_1^{\frac{3}{2}}\Gamma_2^{\frac{1}{2}}}{2\Lambda_1^{\frac{3}{2}}\Lambda_2^{\frac{1}{2}}} - \frac{\Gamma_1^{\frac{1}{2}}\Gamma_2^{\frac{3}{2}}}{2\Lambda_1^{\frac{1}{2}}\Lambda_2^{\frac{3}{2}}} + o(\Gamma^2).$$

where $\Gamma = (\Gamma_1, \Gamma_2)$. We will not need terms of order more than 2 in Γ to obtain a suitable normal form in order to apply KAM theorem, and therefore we will not express any terms beyond this order here.
We did not yet express the term $\cos g$ present in the Hamiltonian. Let us write:

$$\Gamma_0 = \frac{\xi_1\xi_2+\eta_1\eta_2}{2} \qquad \text{and} \qquad \cos g = \frac{\Gamma_0}{\sqrt{\Gamma_1\Gamma_2}}.$$

Hence:

$$e_1e_2\cos g = 2\frac{\Gamma_0}{\sqrt{\Lambda_1\Lambda_2}} - \frac{\Gamma_0\Gamma_1}{2\Lambda_1^{\frac{3}{2}}\Lambda_2^{\frac{1}{2}}} - \frac{\Gamma_0\Gamma_2}{2\Lambda_1^{\frac{1}{2}}\Lambda_2^{\frac{3}{2}}} + o(\Gamma^2),$$

where $\Gamma = (\Gamma_1, \Gamma_2, \Gamma_0)$.The Hamiltonian can then be rewritten:

$$\frac{1}{4\pi^2}\int_{\mathbb{T}^2} H_{pert}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,e_1,e_2,g_1,g_2)d\lambda_1 d\lambda_2 =$$
$$d_{0,0,0}+d_{1,0,0}\Gamma_1+d_{0,1,0}\Gamma_2+d_{0,0,1}\Gamma_0+$$
$$d_{2,0,0}\Gamma_1^2+d_{0,2,0}\Gamma_2^2+d_{0,0,2}\Gamma_0^2+$$
$$d_{1,1,0}\Gamma_1\Gamma_2+d_{1,0,1}\Gamma_1\Gamma_0+d_{0,1,1}\Gamma_2\Gamma_0+o((\Gamma)^2),$$

with explicit coefficient $d_{i,j,k}$, depending on the $c_{i,j}$. Their expression is given in appendix A.2.1, as well as their expansion up to the second order in $\left(\frac{\Lambda_1}{\Lambda_2}\right)$ in appendix A.2.2.
The latter formulas make obvious the dependency of the different terms with respect to powers of Λ_1/Λ_2. In case the ratio of the semi-major axes is small, we can see that the main term of each of these formulas will be a lot larger than the remainder. As well the smaller is this ratio, the smaller will be the terms involving Γ_0, and therefore the difference between the angle of the perihelia will have less impact on the motion.

4.2.2 First normal form of the secular Hamiltonian

As we would like to keep our coordinates Γ_1 and Γ_2 as action coordinates in our problem, it is mandatory to remove the dependency in the angle g of the secular Hamiltonian. The first step in this aim is to obtain the first normal form of the secular Hamiltonian. This step corresponds to carrying out a rotation among the two sets of coordinates (ξ_1, ξ_2) and (η_1, η_2) (rotation of the same angle to have a symplectic transformation), in order to remove the linear term in Γ_0, and hence to put the Hamiltonian into a normal form of order 1 (in Γ). When performing this transformation, we will need to be careful on the complex domains we are considering, and of a possible loss of analyticity. Since we did not yet speak about the domains of analyticity, we will keep this discussion for the next section.
Define the part of the Hamiltonian that interests us:

$$H_{2,P}(\xi, \eta) = d_{0,0,0} + d_{1,0,0}\Gamma_1 + d_{0,1,0}\Gamma_2 + d_{0,0,1}\Gamma_0$$

Let us use the same lemma as Arnold [4]:

Lemma 4.1. *Let R_φ be the rotation of angle $-\varphi$ on the couple of coordinates (ξ_1, ξ_2) and (η_1, η_2), i.e.:*

$$\begin{cases} \xi_1 = \xi_1' \cos\varphi + \xi_2' \sin\varphi \\ \xi_2 = -\xi_1' \sin\varphi + \xi_2' \cos\varphi \end{cases}$$

$$\begin{cases} \eta_1 = \eta_1' \cos\varphi + \eta_2' \sin\varphi \\ \eta_2 = -\eta_1' \sin\varphi + \eta_2' \cos\varphi \end{cases}$$

This transformation is symplectic, and if φ satisfies the equation:

$$(d_{1,0,0} - d_{0,1,0}) \sin 2\varphi + d_{0,0,1} \cos 2\varphi = 0, \tag{4.2}$$

then one can write $H_{2,D}(\xi', \eta') := H_{2,P} \circ R_\varphi(\xi', \eta') = d'_{0,0,0} + d'_{1,0,0}\Gamma'_1 + d_{0,1,0}\Gamma'_2$, with:

$$\begin{aligned} \Gamma'_i &= \frac{\xi_i'^2 + \eta_i'^2}{2} \\ d'_{1,0,0} &= \frac{d_{1,0,0} + d_{0,1,0}}{2} + \frac{d_{1,0,0} - d_{0,1,0}}{2} \cos(2\varphi) - \frac{d_{0,0,1}}{2} \sin(2\varphi) \\ d'_{0,1,0} &= \frac{d_{1,0,0} + d_{0,1,0}}{2} - \frac{d_{1,0,0} - d_{0,1,0}}{2} \cos(2\varphi) + \frac{d_{0,0,1}}{2} \sin(2\varphi) \end{aligned}$$

Proof. Calling $\Gamma'_0 = \dfrac{\xi'_1\xi'_2 + \eta'_1\eta'_2}{2}$, we can compute the values Γ'_i for $i = 1, 2, 0$ in the new variables ξ'_i and η'_i:

$$\begin{aligned} \Gamma_1 &= \Gamma'_1 \cos^2\varphi + \Gamma'_2 \sin^2\varphi + \Gamma'_0 \sin 2\varphi \\ \Gamma_2 &= \Gamma'_1 \sin^2\varphi + \Gamma'_2 \cos^2\varphi - \Gamma'_0 \sin 2\varphi \\ \Gamma_0 &= -\frac{\Gamma'_1}{2} \sin 2\varphi + \frac{\Gamma'_2}{2} \sin 2\varphi + \Gamma'_0 \cos 2\varphi \end{aligned}$$

The new Hamiltonian has the form:

$$H_{2,D}(\xi',\eta') = d'_{1,0,0}\Gamma'_1 + d_{0,1,0}\Gamma'_2 + d'_{0,0,1}\Gamma'_0$$

The coefficient $d'_{0,0,1}$ is then equal to $(d_{1,0,0} - d_{0,1,0})\sin 2\varphi + d_{0,0,1}\cos 2\varphi$. Choosing φ such that equation (4.2) is verified implies $d'_{0,0,1} = 0$. The other coefficients are direct to compute. □

Let us now come back to our case. We have $d_{1,0,0} - d_{0,1,0} = d_{1,0,0}(1 - \Lambda_1/\Lambda_2)$. Under the assumption $\Lambda_1 < \Lambda_2$, this difference is strictly positive. We can see as well that $d_{0,0,1}$ is strictly negative (recall that $b_{1,1,2p+1}$ is always negative). We are now interested in the expression of the secular Hamiltonian that we expanded to the order 2 in Γ', and more precisely in the coefficients $d'_{i,j,k}$ with $i+j+k=2$, the new coefficients after putting the Hamiltonian under normal form of order 1.
To this end, define:

$$\frac{1}{4\pi^2}\int_{\mathbb{T}^2} H_{pert}(\Lambda_1,\Lambda_2,\lambda_1,\lambda_2,e_1,e_2,g_1,g_2)d\lambda_1 d\lambda_2 = H^{\Lambda}_{4,P}(\xi,\eta) + o((\Gamma)^2).$$

We have the following lemma:

Lemma 4.2. *Under the assumption $\Lambda_1 < \Lambda_2$, call*

$$\upsilon = -\frac{d_{0,1,0} - d_{1,0,0}}{d_{0,0,1}} + \sqrt{1 + \left(\frac{d_{0,1,0} - d_{1,0,0}}{d_{0,0,1}}\right)^2} \tag{4.3}$$

we have:

$$\begin{aligned} H^{\Lambda}_{4,D}(\xi',\eta') :&= H^{\Lambda}_{4,P} \circ R_{\varphi}(\xi',\eta') \\ &= d'_{0,0,0} + d'_{1,0,0}\Gamma'_1 + d'_{0,1,0}\Gamma'_2 \\ &\quad + d'_{2,0,0}\Gamma_1'^2 + d'_{0,2,0}\Gamma_2'^2 + d'_{0,0,2}\Gamma_0'^2 \\ &\quad + d'_{1,1,0}\Gamma'_1\Gamma'_2 + d'_{1,0,1}\Gamma'_1\Gamma'_0 + d'_{0,1,1}\Gamma'_2\Gamma'_0. \end{aligned} \tag{4.4}$$

with, for $0 \leq i+j+k \leq 2$*:*

$$
\begin{aligned}
d'_{0,0,0} &= d_{0,0,0}, \\
d'_{1,0,0} &= \frac{1}{(1+\upsilon^2)} \left(d_{1,0,0} + \upsilon^2 d_{0,1,0} - \upsilon d_{0,0,1}\right), \\
d'_{0,1,0} &= \frac{1}{(1+\upsilon^2)} \left(\upsilon^2 d_{1,0,0} + d_{0,1,0} + \upsilon d_{0,0,1}\right), \\
d'_{2,0,0} &= \frac{1}{(1+\upsilon^2)^2} \left(d_{2,0,0} + \upsilon^4 d_{0,2,0} + \upsilon^2 d_{0,0,2} + \upsilon^2 d_{1,1,0} - \upsilon d_{1,0,1} - \upsilon^3 d_{0,1,1}\right), \\
d'_{0,2,0} &= \frac{1}{(1+\upsilon^2)^2} \left(\upsilon^4 d_{2,0,0} + d_{0,2,0} + \upsilon^2 d_{0,0,2} + \upsilon^2 d_{1,1,0} + \upsilon^3 d_{1,0,1} + \upsilon d_{0,1,1}\right), \\
d'_{0,0,2} &= \frac{1}{(1+\upsilon^2)^2} \Big(4\upsilon^2 d_{2,0,0} + 4\upsilon^2 d_{0,2,0} + (1-\upsilon^2)^2 d_{0,0,2} \\
&\qquad -4\upsilon^2 d_{1,1,0} + 2\upsilon(1-\upsilon^2) d_{1,0,1} - 2\upsilon(1-\upsilon^2) d_{0,1,1}\Big), \\
d'_{1,1,0} &= \frac{1}{(1+\upsilon^2)^2} \Big(2\upsilon^2 d_{2,0,0} + 2\upsilon^2 d_{0,2,0} - 2\upsilon^2 d_{0,0,2} \\
&\qquad +(1+\upsilon^4) d_{1,1,0} + \upsilon(1-\upsilon^2) d_{1,0,1} - \upsilon(1-\upsilon^2) d_{0,1,1}\Big), \\
d'_{1,0,1} &= \frac{1}{(1+\upsilon^2)^2} \Big(4\upsilon d_{2,0,0} - 4\upsilon^3 d_{0,2,0} - 2\upsilon(1-\upsilon^2) d_{0,0,2} \\
&\qquad - 2\upsilon(1-\upsilon^2) d_{1,1,0} + (1-3\upsilon^2) d_{1,0,1} + (3\upsilon^2-\upsilon^4) d_{0,1,1}\Big), \\
d'_{0,1,1} &= \frac{1}{(1+\upsilon^2)^2} \Big(4\upsilon^3 d_{2,0,0} - 4\upsilon d_{0,2,0} + 2\upsilon(1-\upsilon^2) d_{0,0,2} \\
&\qquad +2\upsilon(1-\upsilon^2) d_{1,1,0} + (3\upsilon^2-\upsilon^4) d_{1,0,1} + (1-3\upsilon^2) d_{0,1,1}\Big).
\end{aligned}
$$

Proof. Calling

$$\frac{-d_{0,0,1}}{d_{1,0,0} - d_{0,1,0}} := \frac{1}{\eta},$$

we can write $\tan(2\varphi) = 1/\eta$. It implies $tan(2\varphi) > 0$. Hence, we can choose 2φ in $[0, \frac{\pi}{2}[$. Besides we have the formula:

$$\tan 2\varphi = \frac{2\tan\varphi}{1-\tan^2\varphi} = \frac{1}{\eta}$$

Solving this second degree equation in $\tan\varphi$, we find the admissible values $\tan\varphi = -\eta \pm \sqrt{1+\eta^2}$. Since we determined that $\tan(2\varphi)$ is positive, we keep the positive solution

$$\tan\varphi = -\eta + \sqrt{1+\eta^2} = -\frac{d_{1,0,0} - d_{0,1,0}}{-d_{0,0,1}} + \sqrt{1 + \left(\frac{d_{1,0,0} - d_{0,1,0}}{-d_{0,0,1}}\right)^2} = \upsilon$$

Therefore:

$$
\begin{aligned}
\cos\varphi &= \frac{1}{\sqrt{1+\upsilon^2}}, \qquad & \sin\varphi &= \frac{\upsilon}{\sqrt{1+\upsilon^2}}, \\
\cos 2\varphi &= \frac{1-\upsilon^2}{1+\upsilon^2}, \qquad & \sin 2\varphi &= \frac{2\upsilon}{1+\upsilon^2}.
\end{aligned}
$$

With the previous work, we can find define the Γ_i with υ and the Γ'_i :

$$\begin{cases} \Gamma_1 &= +\dfrac{1}{1+\upsilon^2}\,\Gamma'_1 + \dfrac{\upsilon^2}{1+\upsilon^2}\Gamma'_2 + \dfrac{2\upsilon}{1+\upsilon^2}\Gamma'_0 \\ \Gamma_2 &= +\dfrac{\upsilon^2}{1+\upsilon^2}\,\Gamma'_1 + \dfrac{1}{1+\upsilon^2}\Gamma'_2 - \dfrac{2\upsilon}{1+\upsilon^2}\Gamma'_0 \\ \Gamma_0 &= -\dfrac{2\upsilon}{1+\upsilon^2}\dfrac{\Gamma'_1}{2} + \dfrac{2\upsilon}{1+\upsilon^2}\dfrac{\Gamma'_2}{2} + \dfrac{1-\upsilon^2}{1+\upsilon^2}\Gamma'_0 \end{cases}$$

Injecting these formulas in the Hamiltonian $H^{\Lambda}_{4,P}$ one obtains the result of the lemma. □

After removing the dependency in the angles of the linear part of the secular Hamiltonian, we still have to perform the BNF. However, we will need to talk about analyticity widths to do so. Since we did not expressed the initial sets yet, and we did not consider the loss of analyticity width while rotating our variables, it is therefore essential to deal with these right now.

4.2.3 Domain of analyticity in terms of the new actions

In this section, we fix the initial set of the different systems of coordinates necessary in the study of the secular Hamiltonian, and to apply KAM theorem.
The KAM theorem applies to a Hamiltonian expressed in action-anlgles coordinates, which is the perturbation of a Hamiltonian depending only on the actions. We started with the elliptic coordinates, and changed to Poincaré coordinates which are symplectic. The latter are composed of 2 actions, 2 angles, and 4 Cartesian coordinates. After, we expressed our secular part of the perturbation as a function of two actions, Γ_1 and Γ_2. These 2 actions will be the ones we will be working with, together with their associate angles.
First, we deal with the change from polar coordinates to Cartesian coordinates. Secondly, we consider the complex action-angle coordinates and study the transformation of the domains of analyticity under this change of coordinates. Finally, we do the same work for the step herein-before.

From real polar coordinates to real Cartesian coordinates

First, if we recall our bound on the norm of the perturbation: it gives us a domain for the two Cartesian Poincaré coordinates. This domain is therefore the shape of our target set. The change of coordinates, from polar coordinates $(I_3, I_4, \theta_3, \theta_4)$ to the Poincaré variables $(\xi_1, \xi_2, \eta_1\eta_2)$, is the following: for $i = 1, 2$,

$$\begin{cases} \xi_i = \sqrt{2I_{i+2}}\cos\theta_{i+2} \\ \eta_i = \sqrt{2I_{i+2}}\sin\theta_{i+2} \end{cases}$$

In the coordinates $(\xi_1, \xi_2, \eta_1, \eta_2)$, the target set is of the form $\xi_i + \imath\eta_i \in B(0, \rho)$, for $\rho > 0$, and for $i = 1, 2$.
Secondly, this change of variables needs to be a diffeomorphism. It is therefore compulsory to avoid the singularity in 0 of this operation. Our domain of origin will be chosen to be an annulus centered in 0, and of size (m, M), more precisely:

$$D_{pol} = \Big\{(I_3, I_4, \theta_3, \theta_4) \in \mathbb{R}^2 \times \mathbb{T}^2,\ 0 < m < I_3, I_4 < M\Big\}$$

Let us now come back to our transformation, and see how this annulus is transformed under the polar to Cartesian coordinates. Call:

$$\begin{aligned} \psi_1 : D_{pol} &\to \mathbb{R}^4 \\ (I_3, I_4, \theta_3, \theta_4) &\mapsto (\xi_1, \eta_1, \xi_2, \eta_2) \end{aligned}$$

We deduce from these two remarks that $\psi_1(D_{pol}) \subset B_2(0, \sqrt{2M}) \times B_2(0, \sqrt{2M})$, where we grouped the coordinates (ξ_i, η_i) and B_2 is the ball for the euclidean norm.

From polar to Cartesian coordinates in complex variables

We wish now to consider the action-angle variables after our step of analytic continuation. Again, the target set will be fixed, and we want to find a specified domain for our coordinates such that after the change of variables to Cartesian coordinates and the rotation, we belong to the target set.
The set we will consider for the action-angle coordinates is the following polydisc, for $s > 0$, and $0 < r < m$:

$$\tilde{D}_{pol,r,s} = \Big\{(\tilde{I}_3, \tilde{I}_4, \tilde{\theta}_3, \tilde{\theta}_4) \in \mathbb{C}^2 \times \mathbb{T}^2_{\mathbb{C}},\ \exists\, l \in \mathbb{R}, (I_3, I_4, \theta_3, \theta_4) \in D_{pol} : \\ \tilde{I}_i \in B(I_i, r),\ \tilde{\theta}_i - \theta_i = \imath l,\ |l| < s\Big\}$$

After doing our changes, we want the Cartesian coordinates to be of the form:

$$\begin{cases} \tilde{\xi}'_i = \mu_0 \cos\theta_i + \mu_1 \exp(\imath\alpha_i) \\ \tilde{\eta}'_i = \mu_0 \sin\theta_i + \mu_2 \exp(\imath\beta_i) \end{cases}$$

with $0 \le \mu_0 < \rho$, $0 \le \mu_i < \rho'$, $\alpha_i, \beta_i \in \mathbb{T}$, for $i = 1, 2$.
Let us first consider the image of $\tilde{D}_{pol,r,s}$ by the function $\tilde{\psi}_1$, the analytic continuation of the function ψ_1. The target set we want to consider is of the form:

$$\tilde{D}_{cart,\rho,\rho'} = \Big\{(\xi_1, \eta_1, \xi_2, \eta_2) \in \mathbb{C}^4,\ \exists\ (\xi_{1,0}, \eta_{1,0}, \xi_{2,0}, \eta_{2,0}) \in \mathbb{R}^4\ s.t. \\ \xi_{i,0} + \imath\eta_{i,0} \in B(0, \rho),\ |\xi_i - \xi_{i,0}|, |\eta_i - \eta_{i,0}| < \rho'\Big\} \tag{4.5}$$

We have the following lemma:

Lemma 4.3. *$\tilde{\psi}_1(\tilde{D}_{pol,r,s}) \subset \tilde{D}_{cart,\rho,\rho'}$, with*

$$\begin{cases} \rho = \sqrt{2M} \\ \rho' = \max\left(\dfrac{r}{2\sqrt{2m}}\cosh s + \sqrt{2m}(\cosh s - 1), \dfrac{r}{2\sqrt{2M}}\cosh s + \sqrt{2M}(\cosh s - 1)\right) \end{cases} \tag{4.6}$$

Proof. In the real case, we determined in the previous section that we could choose $\rho = 2\sqrt{M}$. We now have to determine ρ'. Let $(\tilde{I}_3, \tilde{I}_4, \tilde{\theta}_3, \tilde{\theta}_4) \in \tilde{D}_{pol,r,s}$, there exists $0 \le l_1, l_2 < s$, $0 \le \alpha_3, \alpha_4 < r$, $(I_3, I_4, \theta_3, \theta_4) \in D_{pol}$ such that $\tilde{I}_i = I_i + \alpha_i \exp(\imath\sigma_i)$, and $\tilde{\theta}_i = \theta_i + \imath l_i$. For $i = 3, 4$:

$$\begin{aligned} |\tilde{\xi}_{i-2} - \xi_{i-2}| &= \left|\sqrt{2I_i + \alpha_i \exp(\imath\sigma_i)}\cos(\theta_i + \imath l_i) - \sqrt{2I_i}\cos(\theta_i)\right| \\ &\le \left|\sqrt{2I_i + \alpha_i \exp(\imath\sigma_i)}\cos(\theta_i + \imath l_i) - \sqrt{2I_i}\cos(\theta_i + \imath l_i)\right| \\ &\quad + \left|\sqrt{2I_i}\cos(\theta_i + \imath l_i) - \sqrt{2I_i}\cos(\theta_i)\right| \\ &\le \sqrt{2I_i}\cosh l_i \left|\frac{\alpha_i \exp(\imath\sigma_i)}{4I_i}\right| + \sqrt{2I_i}(\cosh l_i - 1) \\ &< \frac{r}{2\sqrt{2I_i}}\cosh s + \sqrt{2I_i}(\cosh s - 1) \end{aligned}$$

The variation with I_i of the right-hand side is negative for I_i sufficiently small, and then becomes positive. Since we are interested in an upper-bound of $|\tilde{\xi}_{i-2} - \xi_{i-2}|$, we fix the upper bound as the maximum of the values of the right member for $I_i = m$ and $I_i = M$, thus the choice for ρ'. The result is the same for $|\tilde{\eta}_{i-2} - \eta_{i-2}|$, since converting the cosine to a sine does not affect the calculation. □

First normal form in real coordinates

We would like to perform the rotation of the coordinates $(\xi_1, \eta_1, \xi_2, \eta_2)$, as done in the previous section. Let $\varphi \in [0, 2\pi]$:

$$\begin{cases} \xi_1 = +\xi_1' \cos\varphi + \xi_2' \sin\varphi \\ \xi_2 = -\xi_1' \sin\varphi + \xi_2' \cos\varphi \end{cases}$$

$$\begin{cases} \eta_1 = +\eta_1' \cos\varphi + \eta_2' \sin\varphi \\ \eta_2 = -\eta_1' \sin\varphi + \eta_2' \cos\varphi \end{cases}$$

Let us call $\psi_2 : \mathbb{R}^4 \to \mathbb{R}^4$ this application. Our aim is to find a condition on an initial set of the form $B_2(0, \rho_1) \times B_2(0, \rho_1)$ such that its image by ψ_2 is contained into a set $B_2(0, \rho) \times B_2(0, \rho)$.

Lemma 4.4. *Let $\rho_1 > 0$ and $\rho = \sqrt{2}\rho_1$. We have:*

$$\psi_2(B_2(0, \rho_1) \times B_2(0, \rho_1)) \subset B_2(0, \rho) \times B_2(0, \rho).$$

Proof. Indeed:

$$\begin{aligned} |\xi_1' + \imath\eta_1'|^2 &= |\xi_1 \cos\varphi - \xi_2 \sin\varphi + \imath(\eta_1 \cos\varphi - \eta_2 \sin\varphi)|^2 \\ &= \left|(\xi_1^2 + \eta_1^2)\cos^2\varphi + (\xi_2^2 + \eta_2^2)\sin^2\varphi - 2(\xi_1\xi_2 + \eta_1\eta_2)\cos\varphi\sin\varphi\right| \\ &\le \left|\rho_1^2 \cos^2\varphi + \rho_1^2 \sin^2\varphi + 2\rho_1^2 \cos\varphi \sin\varphi\right| \\ &\le |\rho_1 \cos\varphi + \rho_1 sin\varphi|^2 \\ &< \rho_1^2 |\cos\varphi + sin\varphi|^2 < 2\rho_1^2 \end{aligned}$$

□

The rotation we made was not a classical rotation on the conjugated variables, that is why there is a loss of information while performing it. It concerns only the two couples of coordinates (ξ_1, ξ_2) and (η_1, η_2). However, it could be possible to obtain a better estimate if we were to know precisely the value of φ. As an example, for $\varphi = 0$, the rotation has no effect and our estimation would not be optimal. A loss of a factor $\sqrt{2}$ will not worry us here, and will be sufficient in our case.

First normal form in complex variables

We now have to perform the rotation in order to obtain the new coordinates $(\xi_1', \eta_1', \xi_2', \eta_2')$, and to find a value of ρ' depending on the value of some ρ_1' defining the initial set. Define $\tilde{\psi}_2$ the analytic continuation of ψ_2. We have the following lemma:

Lemma 4.5. *Let $\rho_1, \rho_1' > 0$, and $\rho = \sqrt{2}\rho_1$, $\rho' = \sqrt{2}\rho_1'$. Then:*

$$\tilde{\psi}_2(\tilde{D}_{cart,\rho_1,\rho_1'}) \subset \tilde{D}_{cart,\rho,\rho'}$$

Proof. The value of ρ follows from the previous lemma in real coordinates. Now consider $(\xi_1, \eta_1, \xi_2, \eta_2)$ of the following form:

$$\begin{cases} \xi_i = \xi_{i,0} + \beta_i \exp(\imath\theta_i) \\ \eta_i = \eta_{i,0} + \gamma_i \exp(\imath\sigma_i) \end{cases}$$

with $\xi_{i,0} + \imath\eta_{i,0} \in B(0, \rho_1)$, and $0 \le \beta_i, \gamma_i < \rho_0'$. We have, for $\varphi \in [0, 2\pi]$:

$$\begin{aligned} \xi_1' &= \xi_{1,0} \cos\varphi + \xi_{2,0} \sin\varphi + \beta_1 \exp(\imath\theta_1) \cos\varphi - \beta_2 \exp(\imath\theta_2) \sin\varphi \\ &= \xi_{1,0}' + \beta_1 \exp(\imath\theta_1) \cos\varphi - \beta_2 \exp(\imath\theta_2) \sin\varphi, \end{aligned}$$

where $\xi'_{1,0} \in B(0,\rho)$. Hence:

$$\begin{aligned} |\xi'_1 - \xi'_{1,0}| &< \rho'_1(|\cos\varphi| + |\sin\varphi|) \\ &< \sqrt{2}\rho'_1. \end{aligned}$$

The upper bound holds for ξ_2, η_1, η_2, and we fix $\rho' = \sqrt{2}\rho'_1$, thus the lemma. □

The remark at the end of the previous section still holds here. Depending on the value of φ it is possible to be optimal, but we will not try to further improve this constant.

4.3 Secular Hamiltonian in Birkhoff's normal form

In the previous section, we removed the dependency in the angles of the secular Hamiltonian up to the order 2 in eccentricities (or up to the order 1 in Γ'_i). Moreover, in order to apply the KAM theorem, it is necessary removing this dependency up to the order 4. This will be done by putting the Hamiltonian in a BNF at the required order.
Here, we perform one step of the transformation described in chapter 2 to eliminate the terms in $\Gamma_i\Gamma_j$ depending on the angles with the help of a symplectomorphism. For the terms of order more than 4 in eccentricities, we show later on that one can consider them as part of the perturbation, and therefore we do not try to put our Hamiltonian under BNF at a higher order.

4.3.1 Isolation of the removable terms

To start our reflection, we need to go back to the formalism of Birkhoff's normal form. Recall the definition of the variables $(x_1, x_2, x_3, x_4) \in \mathbb{C}^4$, we have for $i = 1, 2$:

$$\begin{cases} x_i &= \dfrac{1}{\sqrt{2}}(\xi'_i + \imath\eta'_i) \\ \\ x_{i+2} &= \dfrac{1}{\sqrt{2}}(\eta'_i + \imath\xi'_i) \end{cases}$$

Expressing the variables (ξ'_i, η'_i) in terms of the x_j gives:

$$\begin{cases} \xi'_i = \dfrac{1}{\sqrt{2}}(x_i - \imath x_{i+2}) \\ \\ \eta'_i = \dfrac{1}{\sqrt{2}}(x_{i+2} - \imath x_i) \end{cases}$$

We deduce the following formulas for $\Gamma'_0, \Gamma'_1, \Gamma'_2$, and for their products:

$$\begin{aligned} &\Gamma'_1 = -\imath x_1x_3, && \Gamma'^2_1 = -x_1^2x_3^2 \\ &\Gamma'_2 = -\imath x_2x_4, && \Gamma'^2_2 = -x_2^2x_4^2 \\ &\Gamma'_0 = -\imath(x_3x_2 + x_1x_4), && \Gamma'^2_0 = -x_2^2x_3^2 - x_1^2x_4^2 - 2x_1x_2x_3x_4 \\ &\Gamma'_0\Gamma'_1 = -x_1x_2x_3^2 - x_1^2x_3x_4, && \Gamma'_0\Gamma'_2 = -x_2^2x_3x_4 - x_1x_2x_4^2 \\ &\Gamma'_1\Gamma'_2 = -x_1x_2x_3x_4 \end{aligned}$$

Note that we can write $\Gamma'^2_0 = -x_2^2x_3^2 - x_1^2x_4^2 + 2\Gamma'_1\Gamma'_2$. The secular Hamiltonian up to the order in eccentricities can then be written as follows:

$$\begin{aligned} H^{\Lambda}_{4,D}(\xi',\eta') = d'_{0,0,0} &- \imath d'_{1,0,0}x_1x_3 - \imath d'_{0,1,0}x_2x_4 \\ &- d'_{2,0,0}x_1^2x_3^2 + d'_{0,2,0}x_2^2x_4^2 - (d'_{1,1,0} + 2d'_{0,0,2})x_1x_2x_3x_4 \\ &- d'_{0,0,2}x_2^2x_3^2 - d'_{0,0,2}x_1^2x_4^2 - d'_{1,0,1}x_1x_2x_3^2 - d'_{1,0,1}x_1^2x_3x_4 \\ &- d'_{0,1,1}x_2^2x_3x_4 - d'_{0,1,1}x_1x_2x_4^2 + o(x^5), \end{aligned} \tag{4.7}$$

where the first line corresponds to the linear terms, the second to the terms that are not removable with the Birkhoff's normal form, and the third and fourth line are the terms we are able to remove with our transformation.
Indeed, recall that the set of powers for $x = (x_1, x_2, x_3, x_4)$ that are removable with our transformation is the following:

$$B_{4,4} = \Big\{(i_1, i_2, i_3, i_4) \in \mathbb{N}^4,\ i_1 + i_2 + i_3 + i_4 = 4,\ (i_1, i_2) \neq (i_3, i_4)\Big\}$$

The last notation that we will need was $\langle\langle(\alpha_1, \alpha_2), (i_1, i_2, i_3, i_4)\rangle\rangle = \alpha_1(i_3 - i_1) + \alpha_2(i_4 - i_2)$.
Calling the different parts of our Hamiltonian in the following way:

$$\begin{cases} H_2(x) = -\imath d'_{1,0,0} x_1 x_3 - \imath d'_{0,1,0} x_2 x_4, \\ \\ H^{\perp}(x) = d'_{2,0,0} x_1^2 x_3^2 + d'_{0,2,0} x_2^2 x_4^2 - (d'_{1,1,0} + 2d'_{0,0,2}) x_1 x_2 x_3 x_4, \\ \\ H^{\parallel}(x) = -d'_{0,0,2} x_2^2 x_3^2 - d'_{0,0,2} x_1^2 x_4^2 - d'_{1,0,1} x_1 x_2 x_3^2 - d'_{1,0,1} x_1^2 x_3 x_4 \\ \qquad\qquad -d'_{0,1,1} x_2^2 x_3 x_4 - d'_{0,1,1} x_1 x_2 x_4^2, \end{cases} \tag{4.8}$$

the secular Hamiltonian is given by

$$H^{\Lambda}_{4,D}(\xi', \eta') = d'_{0,0,0} + H_2(x) + H^{\perp}(x) + H^{\parallel}(x) + o(x^5). \tag{4.9}$$

We can now apply the corollary 2.7 since we know the expression of the Hamiltonian we are working on.
Let us call $d_1 = (d'_{1,0,0}, d'_{0,1,0})$, we need to consider the supremum of the different values $|d'_{j,k,l}/\langle\langle d_1, i\rangle\rangle|$ for i in the set $B_{4,4}$.

$i \in B_{4,4}$	$\langle\langle d_1, i\rangle\rangle$	Coefficient to remove
$(0,2,2,0)$	$2(d'_{1,0,0} - d'_{0,1,0})$	$d'_{0,0,2}$
$(2,0,0,2)$	$-2(d'_{1,0,0} - d'_{0,1,0})$	$d'_{0,0,2}$
$(1,1,2,0)$	$(d'_{1,0,0} - d'_{0,1,0})$	$d'_{1,0,1}$
$(2,0,1,1)$	$-\ (d'_{1,0,0} - d'_{0,1,0})$	$d'_{1,0,1}$
$(0,2,1,1)$	$(d'_{1,0,0} - d'_{0,1,0})$	$d'_{0,1,1}$
$(1,1,0,2)$	$-\ (d'_{1,0,0} - d'_{0,1,0})$	$d'_{0,1,1}$

As we can see in the previous table, there will be only three coefficients to take into account.
Let us call α the maximum of these coefficients, *i.e.*:

$$\alpha = \max\left(\frac{1}{2}\left|\frac{d'_{0,0,2}}{d'_{1,0,0} - d'_{0,1,0}}\right|, \left|\frac{d'_{1,0,1}}{d'_{1,0,0} - d'_{0,1,0}}\right|, \left|\frac{d'_{0,1,1}}{d'_{1,0,0} - d'_{0,1,0}}\right|\right). \tag{4.10}$$

The last element we need is the domain on which the x are defined. In this section, to apply the theorem we developed before, we assume that the variables are real. Therefore, it is sufficient to consider that the initial analyticity width of our domain is zero. In the present case, for $i = 1, 2$, we have $x_i = (\xi'_i + \imath\eta'_i)/\sqrt{2}$, and $x_{i+2} = (\eta'_i + \imath\xi'_i)/\sqrt{2}$. The supremum norm of x is then bounded by ρ.
On the set $\mathcal{D}_{\Lambda_0,0,\rho,0,0}$, we already calculated an upper bound of the secular Hamiltonian. Let us call $\|H^{\Lambda}_{4,D}\|_{\mathcal{D}_{\Lambda_0,0,\rho,0,0}} < \epsilon$ this bound.
Knowing that our operation needs to remove only 6 terms, we define

$$r_f = \rho - 24\rho^3\alpha.$$

Then there exists an analytic transformation $\tau : B(0, r_f) \to B(0, \rho)$, and such that $H^{\Lambda}_{4,D} \circ \tau$ is under normal form to the order 4, and $\|H^{\Lambda}_{4,D} \circ \tau\|_{\mathcal{D}_{\Lambda_0,0,r_f,0,0}} < \epsilon$. To make the computation

clearer, we did not consider here the change of variables from the Cartesian coordinates to the x_i, we will take care of this transformation in the next section.
To simplify the calculations, we can make one more assumption:

$$\rho \leq \frac{1}{\sqrt{48\alpha}}.$$

In this case, we can take $r_f = \frac{\rho}{2}$, we will have lost half of the size of our initial set after the operation.

4.3.2 Birkhoff's normal form in complex coordinates

In the previous part, we applied the BNF theorem to real variables, and assumed the coefficients of the perturbation were real. In our case, we would like to consider this operation on a set of complex coordinates, and to determine the loss of analyticity width while doing the transformation above.
First, notice that the theorem on BNF can be applied in the complex case, since the transformation we applied is real analytic. Therefore, we can take the analytic continuation of this operation, that we will call $\tilde{\tau}$. Moreover, the constant we found in the theorem will be left unchanged after this step. Indeed, the norms we were considering were based on the absolute value in $\mathbb{R}$, and changing this absolute value to a modulus does not affect the statements. The only thing we need to be worry about is the initial set for the variables x_i, for $1 \leq i \leq 4$.
Let us now call f the following transformation:

$$\begin{aligned} f : \mathbb{C}^4 &\to \mathbb{C}^4 \\ (\xi_1, \eta_1, \xi_2, \eta_2) &\mapsto (x_1, x_2, x_3, x_4), \end{aligned}$$

where the x_i are defined by the formulas:

$$\begin{cases} x_i = \dfrac{1}{\sqrt{2}}(\xi_i' + \imath\eta_i') \\ \\ x_{i+2} = \dfrac{1}{\sqrt{2}}(\eta_i' + \imath\xi_i') \end{cases}$$

The application f is clearly symplectic, and analytic on $\mathbb{C}^4$. The transformation we want to apply to our Hamiltonian $H_{4,D}^{\Lambda}$ is then $f^{-1} \circ \tilde{\tau} \circ f$. Finally, we need to apply our function $\tilde{\psi}_1$ to have a symplectic application from an initial set of action-angle coordinates, into a set on which we can calculate an upper bound of the Hamiltonian perturbation. Therefore, the final application we want to consider is the following one:

$$\Delta = \tilde{\psi}_2 \circ f^{-1} \circ \tilde{\tau} \circ f \circ \tilde{\psi}_1 \tag{4.11}$$

We now have to determine the target set of the set $\tilde{D}_{pol,r,s}$ by the symplectic application Δ.
Let $0 < m < M$ and $r, s > 0$ define the initial set $\tilde{D}_{pol,r,s}$. Using lemmas 4.3 and 4.5, we derive the following lemma:

Lemma 4.6. *Define:*

$$\begin{cases} \rho = 2\sqrt{M} \\ \rho' = \max\limits_{l \in \{m,M\}} \left(\dfrac{r}{2\sqrt{2l}} \cosh s + \sqrt{2l}(\cosh s - 1) \right) \end{cases}$$

If $\rho + 2\rho' \leq \left(2\sqrt{48\alpha}\right)^{-1}$, *where* α *is defined in* (4.10), *then we have:*

$$\Delta(\tilde{D}_{pol,r,s}) \subset \tilde{D}_{cart,4\sqrt{2}(\rho+2\rho'),2\sqrt{2}\rho+5\sqrt{2}\rho'} \tag{4.12}$$

Proof. From lemma 4.3, we know that $\tilde{\psi}_1(\tilde{D}_{pol,r,s}) \subset \tilde{D}_{cart,\rho,\rho'}$. By definition of f, we easily obtain that for $i \in [\![1,4]\!]$, $|x_i| \leq \rho + 2\rho'$. The transformation $\tilde{\tau}$ takes the set $B(0,r_f)$, where $r_f = \rho + 2\rho'$ here, and sends it into the set $B(0,r_0)$, where $r_f = r_0 - 24r_0^3\alpha$. We already noticed that if $r_0 \leq \sqrt{48\alpha}$, we have $\tilde{\tau}(B(0,r_0/2)) \subset B(0,r_0)$. Choosing r_0 such that $2(\rho+2\rho') = r_0$, and under the assumption in the lemma we have

$$\tilde{\tau}(B(0,\rho+2\rho') \subset B(0,2\rho+4\rho').$$

From the BNF theorem, we also now that $\tilde{\tau} = Id + v$, where

$$v : B(0,r_0/2) \to B(0,r_0/2).$$

Therefore, writing $(x_1',x_2',x_3',x_4') = \tilde{\tau}(x_1,x_2,x_3,x_4)$, we have, for $i \in [\![1,4]\!]$:

$$x_i' = x_i + (v(x_i))_i$$
$$\|(v(x_i))_i\| \leq \frac{r_0}{2}$$

In this case, the bound on v is so large that we will suffer a great loss of information. Surely, when we will apply f^{-1}, it is not clear what will happen to the real variables $\xi_{i,0}$ or $\eta_{i,0}$, and to the "complex part" taken in the polydisc around the real variables. Without being optimal, we have:

$$f^{-1} : B(0,2\rho+4\rho') \to \tilde{D}_{cart,4(\rho+2\rho'),2\rho+5\rho'}$$

Indeed, adding the loss of analyticity to each coordinates $(\xi_1,\eta_1,\xi_2,\eta_2)$ makes the result straightforward. For instance, if we consider $\xi_i = r_0\cos(\theta) + r_1\exp(\imath\theta_i)$ (the real part and the part in the polydisc) and $\eta_i = r_0\sin(\theta) + r_2\exp(\imath\theta_i)$, the new variables ξ_i' and η_i' verify:

$$|\xi_i' + \imath\eta_i'| = |x_i + (v(x_i))_i| \leq 2|x_i|$$

We deduce that we can take $\xi_{i,0}' + \imath\eta_{i,0}' \in B(0,4(\rho+2\rho'))$. Now for each of these variables, the part in the polydisc has been changed by at most $2\rho+4\rho'$ as well. Thus, the target set can be chosen as stated.
Finally, with lemma 4.5, we know that $\tilde{\psi}_2(\tilde{D}_{cart,4(\rho+2\rho'),2\rho+5\rho'}) \subset \tilde{D}_{cart,4\sqrt{2}(\rho+2\rho'),\sqrt{2}(2\rho+5\rho')}$. This ends to prove the lemma. ☐

When applying the transformation Δ, we lost a lot of information regarding the sets D_{cart}. Indeed, the width ρ and ρ' have been mixed up, because the norm of the transformation $\tilde{\tau}$ was large compared to the initial analyticity width. We notice here that for a choice of ρ sufficiently small, the effect on the bound of the perturbation is very small and we will therefore get along with this mixing of the coordinates.

4.4 The final Hamiltonian

4.4.1 Unperturbed Hamiltonian and perturbation

In this section, we will make explicit the new form of our Hamiltonian, that will be composed of two main parts: the first part will be the unperturbed Hamiltonian, the second one the perturbation. In the lights of our previous work, we can construct the unperturbed Hamiltonian as follows.

The initial Hamiltonian H is composed of a Kepler part and of the gravitational interaction of the two planets (in Jacobi's coordinates). Decompose the gravitational interaction into the secular part $\bar{H}_{pert}$ (not depending on the fast angles), and the non-secular part $\tilde{H}_{pert}$. Applying the function $\tilde{\psi}_2$ to the secular part, it can be expressed under the form:

$$\bar{H}_{pert} \circ \tilde{\psi}_2 = H_{2,D} + H_4^{\perp} + H_4^{\parallel} + H_{\geq 6}$$

When applying the function Δ in the whole space, one needs to be careful. Indeed, the symplectic diffeomorphism Δ is parametrized by the value Λ (and independent of the angle λ). Hence, when lifting this function to the space of the $(\Lambda, \lambda, I, \theta)$, it will create a component modifying the angles λ. In order to be able to perform this transformation, we have to verify one more hypothesis, and it will result in a loss of analyticity among the angles λ. Yet, this symplectomorphism does not affect the variables Λ, leaving unchanged the Kepler Hamiltonian:

$$\begin{aligned} H_{0,1} &= H_{Kep} + H_{2,D} \circ \tilde{\psi}_1 + H_4^{\perp} \circ \tilde{\psi}_1 \\ P_1 &= \bar{H}_{pert} \circ \Delta - H_{2,D} \circ \tilde{\psi}_1 - H_4^{\perp} \circ \tilde{\psi}_1, \end{aligned}$$

and we have:

$$H \circ \Delta = H_{0,1} + P_1 + \tilde{H}_{pert} \circ \Delta.$$

Now apply the transformation $\varphi^{\epsilon}_{X_1}$ described in the chapter 2 so as to make the term $\tilde{H}_{pert} \circ \Delta$ even smaller. Again, when lifting this symplectomorphism to the complete set of variables, since the variables (ξ, η) appeared as parameter of the Hamiltonian generating the transformation, the latter will change the variables (ξ, η). This implies a loss of analyticity over these coordinates, as well as another hypothesis. We define:

$$P_2 = (H_{0,1} + \tilde{H}_{pert} \circ \Delta) \circ \varphi^{\epsilon}_{X_1} - H_{0,1},$$

and our Hamiltonian will be:

$$H \circ \Delta \circ \varphi^{\epsilon}_{X,1} = H_{0,1} + P_1 \circ \varphi^{\epsilon}_{X_1.} + P_2$$

In fact, considering only one transformation $\varphi^{\epsilon}_{X_1}$, will not be enough, and we will need to apply 2 more times this scheme. Define $\bar{P}_2$ the integral over the fast angle of P_2, and its remainder $\tilde{P}_2$. The average Hamiltonian $\bar{P}_2$ depends on the angle g, and is not under normal form. Since we do not know its form, we want to get rid of its dependence in this angle. P_2 being a part of the perturbation, it is made of a constant part, plus even powers of the eccentricities. Calling $y = (y_1, y_2, y_3, y_4) = (\xi_1, \eta_1, \xi_2, \eta_2)$, we can therefore decompose $\bar{P}_2$ in this way:

$$\bar{P}_2(\Lambda, y) = \bar{P}_{2,0}(\Lambda) + \int_0^1 \partial_{y,y} \bar{P}_2(\Lambda, ty) \cdot y^2 (1-t) dt,$$

where $\partial_{y,y} \bar{P}_2(\Lambda, ty)$ is a 4×4 matrix representing the second derivative of $\bar{P}_2$ with respect to the y_i. Considering small enough eccentricities, one can make the term under the integral as small as wanted, hence considered as part of the perturbation. Let $R_2 = \bar{P}_2 - \bar{P}_{2,0}$, and

$$H_{0,2} = H_{0,1} + \bar{P}_{2,0}.$$

Hence, the Hamiltonian is now:

$$H \circ \Delta \circ \varphi^{\epsilon}_{X,1} = H_{0,2} + P_1 \circ \varphi^{\epsilon}_{X_1} + R_2 + \tilde{P}_2.$$

We can apply again a transformation $\varphi^{\epsilon}_{X_2}$ to the Hamiltonian such that it makes the part $\tilde{P}_2$ even smaller (by solving the cohomological equations $\{H_{0,2}, G\} = \tilde{P}_2$). After this operation, we have

$$H \circ \Delta \circ \varphi^{\epsilon}_{X_1} \circ \varphi^{\epsilon}_{X_2} = H_{0,2} + P_1 \circ \varphi^{\epsilon}_{X_1} \circ \varphi^{\epsilon}_{X_2} + R_2 \circ \varphi^{\epsilon}_{X_2} + P_3.$$

In the exact same way, we will reproduce the scheme we have just done. Hence, define $\bar{P}_3, \tilde{P}_3$, $R_3 = \bar{P}_3 - \bar{P}_{3,0}$ the unperturbed Hamiltonian

$$H_{0,3} = H_{0,2} + \bar{P}_{3,0},$$

and the transformation $\varphi^{\epsilon}_{X_3}$. Let us call

$$\varphi^{\epsilon}_{X} = \varphi^{\epsilon}_{X_1} \circ \varphi^{\epsilon}_{X_2} \circ \varphi^{\epsilon}_{X_3},$$

and the Hamiltonian is now:

$$H \circ \Delta \circ \varphi^{\epsilon}_X = H_{0,3} + P_1 \circ \varphi^{\epsilon}_X + R_2 \circ \varphi^{\epsilon}_{X_2} \circ \varphi^{\epsilon}_{X_3} + R_3 \circ \varphi^{\epsilon}_{X_3} + P_4. \tag{4.13}$$

In our version of the KAM theorem, we considered the perturbation of a linear Hamiltonian. Since $H_{0,1}$ is not already linear, we had to divide it again into two terms, a linear one and a non-linear one.
Basically, it consisted in a Taylor expansion at the order 2 in the actions, so as to express the unperturbed Hamiltonian as a sum of a linear part, and a remainder. The remainder, of order 2 in the actions will be small enough to be considered as a part of the perturbation.
Let $p = (\Lambda_1, \Lambda_2, I_3, I_4)$ be the vector of the actions, and p_0 be a specific vector in the initial set (yet to be described). We can write $p = p_0 + I'$, with I' close to 0, and expand the Hamiltonian around the vector p_0. We have:

$$H_{0,3}(p) = H_{0,3}(p_0) + H'_{0,3}(p_0) \cdot I' + \int_0^1 (1-t) H''_{0,3}(p_t) \cdot I'^2 dt,$$

where $p_t = p_0 + tI'$. Therefore, let $H_{nl}(I', p_0) = \int_0^1 (1-t) H''_{0,3}(p_t) \cdot I'^2 dt$, and

$$H_l(I', p_0) = H_{0,3}(p_0) + H'_{0,3}(p_0) \cdot I'.$$

The Hamiltonian under its final form is the following one:

$$\begin{aligned} H \circ \Delta \circ \varphi^{\epsilon}_X &= H_l + P_{tot} \\ &= H_l + P_1 \circ \varphi^{\epsilon}_X + R_2 \circ \varphi^{\epsilon}_{X_2} \circ \varphi^{\epsilon}_{X_3} + R_3 \circ \varphi^{\epsilon}_{X_3} + P_4 + H_{n,l} \end{aligned} \tag{4.14}$$

4.4.2 Bound on the norm of the perturbation

After determining precisely the different parts of the final perturbation, we can now bound it more explicitly. More precisely, we are going to bound each of its parts separately, with the work previously done.
To this end, let $(\Lambda_{0,1}, \Lambda_{0,2}, m, M) \in (\mathbb{R}^+)^4$, represent the initial condition of our system. Now let $0 < r_1, r_2, s_1, s_2$ be the analyticity width of the four action-angle variables, defined in the following domains: the two first action-angle variables $(\Lambda_1, \Lambda_2, \lambda_1, \lambda_2)$ take their values in the domain $B_{\Lambda_0}(r_1, s_1)$, where

$$B_{\Lambda_0}(r_1, s_1) = \left\{ (\Lambda, \lambda) \in \mathbb{C}^2 \times \mathbf{T}^2_{\mathbb{C}}, \max_{j \in [\![1,2]\!]} |\Lambda_j - \Lambda_{0,j}| < r_1, \max_{j \in [\![1,2]\!]} |\Im \lambda_j| < s_1 \right\};$$

the other action-angle variables will be in the set $(I_3, I_4, \theta_3, \theta_4) \in \tilde{D}_{pol,r_2,s_2}$.
We will consider thereafter that ρ and ρ' verify the assumption

$$\rho + 2\rho' < (2\sqrt{48\alpha})^{-1}.$$

Now define the constants ρ, ρ' as in lemma 4.6, and (for the sake of simplicity):

$$\left\{ \begin{array}{l} \rho_0 = 4\sqrt{2}(\rho + 2\rho') \\ \rho'_0 = 2\sqrt{2}\rho + 5\sqrt{2}\rho' \end{array} \right. \tag{4.15}$$

Finally, define the norm $\| \cdot \|_{r_1,s_1,r_2,s_2}$ as the supremum over the set $B_{\Lambda_0}(r_1, s_1) \times \tilde{D}_{pol,r_2,s_2}$.

Additional hypotheses for the lift of symplectomorphism

As said before, we consider the lifts of the four symplectomorphisms we are using: Δ, $\varphi^{\epsilon}_{X_1}$, $\varphi^{\epsilon}_{X_2}$, and $\varphi^{\epsilon}_{X_3}$. Each of these transformations induces a drift over the direction of the variables that where seen as parameters before. As an example, consider the function $\varphi^{\epsilon}_{X_i}$, which transformed the variables Λ and λ, for a fixed value of $(I_3, I_4, \theta_3, \theta_4)$. It is generated by a Hamiltonian $G_{I_3,I_4,\theta_3,\theta_4}(\Lambda, \lambda)$. When considering the lift of this Hamiltonian to our space of 8 complex variables, we consider the Hamiltonian $G(\Lambda, \lambda, I_3, I_4, \theta_3, \theta_4)$ instead. Hence, the vector field associated to it is in 8 dimensions instead of 4. To control the relation between the size of a domain and its image by the time-one map of the vector field, we require that we lose half of our analyticity widths when applying $\varphi^{\epsilon}_{X_i}$. The hypotheses of 2.16 ensures that it is true in the directions of the variables Λ and λ. It is therefore necessary to make further assumptions for the other variables, using the formula of the minimum escape time derived in lemma 2.13. Regarding the transformation Δ, the same reasoning on the polynomial P generating the transformation leads to the necessity of satisfying another hypothesis. However, this time, the polynomial P is independent of the variable λ, and hence it leaves the variable Λ unchanged. Considering that before applying Δ we have an analyticity width s_1 for the variable λ, we want to lose only half of our analyticity width for this angle.
In lemma 2.5, we wanted the minimum escape time to satisfy:

$$\bar{t} \geq \frac{\delta}{||X_P||(R+\delta)^{m-1}}.$$

The polynomial P now depends in the variable Λ, in the following way:

$$P(x) = \sum_{i \in A_{m,n}} \beta_{m,i}(\Lambda) x^i$$

It is straightforward to see that the extra hypothesis that needs to be satisfied is

$$\bar{t} \leq \frac{s_1}{2\sum_{i \in A_{m,n}} \|\beta'_{m,i}(\Lambda)\|_{B(\Lambda_0, r_1)}(R+\delta)^m}, \tag{4.16}$$

where $B(\Lambda_0, r)$ is the the usual ball. Letting $t = 1$, the assumption we consider is then:

$$2 \sum_{i \in A_{m,n}} \|\partial_{\Lambda_i} \beta_{m,i}(\Lambda)\|_{B(\Lambda_0, r_1)} (\rho_0 + \rho'_0)^4 \leq s_1$$

The computation of a bound on the derivatives of $\beta_{m,i}$ can be done in two different ways. The expression of these terms is explicit, hence one can compute it directly using their formula and then bound it. Another solution is to use Cauchy's inequality, though it is less optimal and could change artificially the validity of the previous inequation. It nevertheless eases the computation. Regarding the transformations $\varphi^{\epsilon}_{X_i}$, we divide our analyticity widths each time by 2. The additional assumption on $\varphi^{\epsilon}_{X_1}$ is:

$$\epsilon \times \|H_{pert} \circ \Delta\|_{r_1, s_1/2, r_2, s_2} \leq \gamma \frac{s_1^{\tau} r_2 s_2}{2^{12} 5^2}.$$

This hypothesis, together with all the hypotheses of 2.16 ensures that we have a function $\varphi^{\epsilon}_{X_1}$ that is well defined on the previous set and such that:

$$\varphi^{\epsilon}_{X_1}(B_{\Lambda_0}(r_1/2, s_1/4) \times \tilde{D}_{pol, r_2/2, s_2/2}) \subset B_{\Lambda_0}(r_1, s_1/2) \times \tilde{D}_{pol, r_2, s_2}.$$

For the transformations $\varphi^{\epsilon}_{X_2}$ and $\varphi^{\epsilon}_{X_3}$, we will deal with them in the next paragraph, where we gather all the definitions necessary to express them explicitly.

Bound on the non-secular part

First, we are going to bound the part P_4 of the perturbation. This bound can be determined with the help of the scheme developed in chapter 2. Instead of using directly the corollary 2.16, we use the iterated one, after 3 steps (at the end of chapter 2). Consider the Hamiltonian $H_{0,1} + \tilde{H}_{pert} \circ \Delta$, we apply the scheme to find a function φ_X^ϵ such that P_4 becomes smaller. We observe that:

$$\left\|\tilde{H}_{pert} \circ \Delta\right\|_{r_1,s_1/2,r_2,s_2} \leq \left\|\tilde{H}_{pert}\right\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}}.$$

Define the following elements iteratively:

$$K_2 \leq \left(\frac{\gamma}{2r_1 \left\|H''_{0,1}\right\|_{r_1,s_1/2,r_2,s_2}}\right)^{\frac{1}{3}},$$

$$\epsilon_2 = \frac{2^6.10^{10}}{\gamma_2^2 s_1^6}\left\|\tilde{H}_{pert}\right\|^2_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}}\left(\left\|H''_{0,1}\right\|_{r_1,s_1/2,r_2,s_2} + \frac{8}{r_1}\left\|H'_{0,1}\right\|^r_{1,r_1,s_1/2,r_2,s_2}\right) + 32K_2^2\exp\left(-\frac{K_2 s_1}{5}\right)\left\|\tilde{H}_{pert}\right\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}},$$

$$K_3^3 \leq \min\left(\frac{2\gamma}{r_1\left\|H''_{0,1}\right\|_{r_1,s_1/2,r_2,s_2}}, \frac{r_1\gamma}{16\left\|P_2\right\|_{r_1/2,s_1/4,r_2,s_2}}\right),$$

$$\epsilon_3 = \frac{2^{14}.10^{10}}{\gamma^2 s_1^6}\epsilon_2^2\left(\left\|H''_{0,1}\right\|_{r_1,s_1/2,r_2,s_2} + \frac{2^5}{r_1}\left\|H'_{0,1}\right\|^I_{1,r_1,s_1/2,r_2,s_2} + \frac{2^8}{r_1^2}\epsilon_2\right) + 32K_3^2\exp\left(-\frac{K_3 s_1}{10}\right)\epsilon_2,$$

$$K_4^3 \leq \min\left(\frac{8\gamma}{r_1\left\|H''_{0,1}\right\|_{r_1,s_1/2,r_2,s_2}}, \frac{r_1\gamma}{2^6\left\|P_2\right\|_{r_1/2,s_1/4,r_2,s_2}}, \frac{r_1\gamma}{2^7\left\|P_3\right\|_{r_1/8,s_1/8,r_2,s_2}}\right),$$

$$\epsilon_4 = \frac{2^{22}.10^{10}}{\gamma^2 s_1^6}\epsilon_3^2\left(\left\|H''_{0,1}\right\|_{r_1,s_1/2,r_2,s_2} + \frac{2^7}{r_1}\left\|H'_{0,1}\right\|^I_{1,r_1,s_1/2,r_2,s_2} + \frac{2^{10}}{r_1^2}(\epsilon_2+\epsilon_3)\right) + 32K_4^2\exp\left(-\frac{K_4 s_1}{20}\right)\epsilon_3.$$

Proposition 4.7. *Under the assumption that H_{pert} is analytic on the set $\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}$, that there exists $I_0 \in B(r_1/32, s_1/16)$, such that $H'_{0,1}(I_0) \in D(\gamma, 2)$, and that the following assumptions are verified:*

$$\left\|\tilde{H}_{pert}\right\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}} \leq \min\left(\frac{\gamma r_1 s_1^3}{2^5.10^5}, \frac{\gamma r_2 s_1^2 s_2}{2^4.10^4}\right), \qquad 2^5K_2^2\exp(-2K_2s_1/5) < 1,$$

$$\epsilon_2 \leq \min\left(\frac{\gamma r_1 s_1^3}{2^{11}.10^5}, \frac{\gamma r_2 s_1^2 s_2}{2^9.10^4}\right), \qquad 2^5K_3^2\exp(-K_3s_1/5) < 1,$$

$$\epsilon_3 \leq \min\left(\frac{\gamma r_1 s_1^3}{2^{17}.10^5}, \frac{\gamma r_2 s_1^2 s_2}{2^{14}.10^4}\right), \qquad 2^5K_4^2\exp(-K_4s_1/10) < 1,$$

there exists a symplectic map $\varphi_X^\epsilon : B(r_1/32, s_1/16) \times \tilde{D}_{pol,r_2/8,s_2/8} \to B(r_1, s_1/2) \times \tilde{D}_{pol,r_2,s_2}$*, and the following bounds hold:*

$$\begin{aligned} \|P_2\|_{r_1/2,s_1/4,r_2/2,s_2/2} &\leq \epsilon_2, \\ \|P_3\|_{r_1/8,s_1/8,r_2/4,s_2/4} &\leq \epsilon_3, \\ \|P_4\|_{r_1/32,s_1/16,r_2/8,s_2/8} &\leq \epsilon_4. \end{aligned}$$

Bound on the remainder of the BNF

We will be concerned in this part on the bound of $P_1 \circ \varphi_X^\epsilon$. First, let us bound the term P_1 of the perturbation. To perform the BNF, we needed to consider the flow of the application associated to the vector field X_P, where P is the following polynomial:

$$\begin{aligned} P : &\mathbb{C}^4 \to \mathbb{C} \\ &(x_1, x_2, x_3, x_4) \mapsto a_1(x_2^2 x_3^2 - x_1^2 x_4^2) + a_2(x_1 x_2 x_3^2 - x_1^2 x_3 x_4) + a_3(x_2^2 x_3 x_4 - x_1 x_2 x_4^2) \end{aligned}$$

with

$$a_1 = \frac{d'_{0,0,2}}{2(d'_{1,0,0} - d'_{0,1,0})}, \quad a_2 = \frac{d'_{1,0,1}}{d'_{1,0,0} - d'_{0,1,0}}, \quad a_3 = \frac{d'_{0,1,1}}{d'_{1,0,0} - d'_{0,1,0}}. \tag{4.17}$$

Define:

$$\begin{aligned} C_1 = 2^8 &\left(2\|a_1\|_{r_1,s_1} + 3\|a_2\|_{r_1,s_1} + 3\|a_3\|_{r_1,s_1}\right) \times \\ &\left(2\left\|d'_{0,0,2}\right\|_{r_1,s_1} + 3\left\|d'_{1,0,1}\right\|_{r_1,s_1} + 3\left\|d'_{0,1,1}\right\|_{r_1,s_1}\right), \end{aligned} \tag{4.18}$$

$$\begin{aligned} C_2 = 2^8 &\left(2\|a_1\|_{r_1,s_1} + 3\|a_2\|_{r_1,s_1} + 3\|a_3\|_{r_1,s_1}\right) \times \\ &\left(2\left\|d'_{2,0,0}\right\|_{r_1,s_1} + 2\left\|d'_{0,2,0}\right\|_{r_1,s_1} + \left\|d'_{1,1,0}\right\|_{r_1,s_1} + 2\left\|d'_{0,0,2}\right\|_{r_1,s_1}\right). \end{aligned} \tag{4.19}$$

We have the following result:

Proposition 4.8. *Let* $r_1, s_1, r_2, s_2 > 0$*, define* ρ *and* ρ' *as in lemma 4.6 (with* $r = r_2$ *and* $s = s_2/2$*), and* ρ_0, ρ'_0 *as in equation (4.15). If* H_{pert} *is analytic on the set* $\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho'_0+\mu,s_1}$ *with* $\mu > 0$*, and that it satisfies the hypotheses:*

$$\begin{cases} \rho + 2\rho' \leq (2\sqrt{48\alpha})^{-1}, \\ \\ 4(\|\partial_{\Lambda_i} a_1\|_{B(\Lambda_0,r_1)} + \|\partial_{\Lambda_i} a_2\|_{B(\Lambda_0,r_1)} + \|\partial_{\Lambda_i} a_3\|_{B(\Lambda_0,r_1)})(\rho_0 + \rho'_0)^4 \leq s_1, \end{cases}$$

then, on the set $B_{\Lambda_0}(r_1/32, s_1/8) \times \tilde{D}_{pol,r_2/8,s_2/8}$*, the following bound holds:*

$$\|P_1 \circ \varphi_X^\epsilon\|_{r_1/32,s_1/16,r_2/8,s_2/8} \leq \left(C_1 + C_2 + \frac{96 \times 7^7}{\mu^6}\|H_{pert}\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho'_0+\mu,s_1}}\right)(\rho_0 + 2\rho'_0)^6.$$

Proof. : We can consider the term P_1, on the set $B_{\Lambda_0}(r_1, s_1/2) \times \tilde{D}_{pol,r_2,s_2}$. The composition with the function φ_X^ϵ will make the estimate hold on the wanted set.
The transformation τ is the flow τ^t associated to the vector field X_P for $t = 1$. Notice that each a_i here depends on the variables Λ_1, Λ_2, we will denote by $\|a_i\|_{r_1,s_1}$ their supremum bound on the domain $B_{\Lambda_0}(r_1, s_1)$. Considering as well $\|x\|$ the supremum bound of (x_1, x_2, x_3, x_4) on

the domain of definition (that we will determine further), we can bound the derivative of the polynomial by:

$$|\partial_i P| \leq \left(2\left\|a_1\right\|_{r_1,s_1} + 3\left\|a_2\right\|_{r_1,s_1} + 3\left\|a_3\right\|_{r_1,s_1}\right)\|x\|^3 .$$

Now let us express explicitly the terms we are dealing with:

$$\begin{aligned}
\bar{H}_{pert} \circ \tilde{\psi}_2 \circ f^{-1} \circ \tau =& (H_{2,D,f} + H^{\perp}_{4,f} + H^{\|}_{4,f} + H_{\geq 6,f}) \circ \tau \\
=& \quad H_{2,D,f} + \{H_{2,D,f}, P\} + \int_0^1 (1-t)\left\{\{H_{2,D,f}, P\}, P\right\} \circ \tau^t \, dt \\
&+ H^{\perp}_{4,f} \quad + \int_0^1 \left\{H^{\perp}_{4,f}, P\right\} \circ \tau^t \, dt \\
&+ H^{\|}_{4,f} \quad + \int_0^1 \left\{H^{\|}_{4,f}, P\right\} \circ \tau^t \, dt \\
&+ H_{\geq 6,f} \circ \tau
\end{aligned}$$

where we defined $H_{N,f} = H_N \circ f^{-1}$ to make the formulas clearer.
The polynomial P was constructed so as to have $\{H_{2,D}, P\} = -H^{\|}_4$. Thus, we have:

$$\begin{aligned}
\bar{H}_{pert} \circ \tilde{\psi}_2 \circ f^{-1} \circ \tau - H_{2,D,f} - H^{\perp}_{4,f} =& \int_0^1 \left\{(1-t)\{H_{2,D,f}, P\} + H^{\|}_{4,f}, P\right\} \circ \tau^t \, dt \\
&+ \int_0^1 \left\{H^{\perp}_{4,f}, P\right\} \circ \tau^t \, dt + H_{\geq 6,f} \circ \tau
\end{aligned}$$

Using again the construction of P, we have:

$$\int_0^1 \left\{(1-t)\{H_{2,D,f}, P\} + H^{\|}_{4,f}, P\right\} \circ \tau^t \, dt = \int_0^1 t\left\{H^{\|}_{4,f}, P\right\} \circ \tau^t \, dt$$

Composing the secular part of the perturbation by f on both side, we obtain:

$$\begin{aligned}
\bar{H}_{pert} \circ \Delta - H_{0,1} =& \int_0^1 t\left\{H^{\|}_{4,f}, P\right\} \circ \tau^t \circ f \circ \tilde{\psi}_1 \, dt \\
&+ \int_0^1 \left\{H^{\perp}_{4,f}, P\right\} \circ \tau^t \circ f \circ \tilde{\psi}_1 \, dt + H_{\geq 6,f} \circ \tau \circ f \circ \tilde{\psi}_1
\end{aligned} \tag{4.20}$$

This explicit equation gives three terms to bound:

$$\begin{cases}
R_1 = \int_0^1 t\left\{H^{\|}_{4,f}, P\right\} \circ \tau^t \circ f \circ \tilde{\psi}_1 \, dt \\
R_2 = \int_0^1 \left\{H^{\perp}_{4,f}, P\right\} \circ \tau^t \circ f \circ \tilde{\psi}_1 \, dt \\
R_3 = H_{\geq 6,f} \circ \tau \circ f \circ \tilde{\psi}_1
\end{cases} \tag{4.21}$$

Given the definition of the terms $H^{\|}_{4,f}$ and $H^{\perp}_{4,f}$ in the coordinates (x_1, x_2, x_3, x_4) in (4.8), we can bound their derivatives by:

$$\begin{aligned}
\left|\partial_i H^{\|}_{4,f}\right| &\leq \left(2\left\|d'_{0,0,2}\right\|_{r_1,s_1} + 3\left\|d'_{1,0,1}\right\|_{r_1,s_1} + 3\left\|d'_{0,1,1}\right\|_{r_1,s_1}\right)\|x\|^3 \\
\left|\partial_i H^{\perp}_{4,f}\right| &\leq \left(2\left\|d'_{2,0,0}\right\|_{r_1,s_1} + 2\left\|d'_{0,2,0}\right\|_{r_1,s_1} + \left\|d'_{1,1,0}\right\|_{r_1,s_1} + 2\left\|d'_{0,0,2}\right\|_{r_1,s_1}\right)\|x\|^3
\end{aligned}$$

With these bounds, we can calculate the bounds on R_1 and R_2 on the domain of definition:

$$\begin{aligned}
\|R_1\|_{r_1,s_1/2,r_2,s_2} &\leq \int_0^1 t \left\| \left\{ H^{\|}_{4,f}, P \right\} \circ \tau^t \circ f \circ \tilde{\psi}_1 \right\|_{r_1,s_1/2,r_2,s_2} dt \\
&\leq \int_0^1 \left\| \left\{ H^{\|}_{4,f}, P \right\} \circ \tau^t \circ f \right\|_{r_1,s_1/2,\tilde{D}_{cart,\rho_0,\rho_0'}} dt \\
&\leq \int_0^1 \left\| \left\{ H^{\|}_{4,f}, P \right\} \circ \tau^t \right\|_{r_1,s_1/2,B(0,\rho_0+2\rho_0')} dt \\
&\leq \int_0^1 \left\| \left\{ H^{\|}_{4,f}, P \right\} \right\|_{r_1,s_1,B(0,2\rho_0+4\rho_0')} dt \\
&\leq \left\| \left\{ H^{\|}_{4,f}, P \right\} \right\|_{r_1,s_1,B(0,2\rho_0+4\rho_0')} \\
&\leq C_1 \left(\rho_0 + 2\rho_0'\right)^6, \qquad (4.22)
\end{aligned}$$

The constant 2^8 in the formula of C_1 comes from the fact that $(2\rho_0 + 4\rho_0')^6 = 2^6(\rho_0 + 2\rho_0')^6$ and from the fact that we have 4 term in the Poisson bracket. For the same reason, we have:

$$\|R_2\|_{r_1,s_1/2,r_2,s_2} \leq C_2 \left(\rho_0 + 2\rho_0'\right)^6 \qquad (4.23)$$

The last term cannot be bounded in the same way. Indeed, we do not know explicitly the different terms appearing in this part of the Hamiltonian. Though, we know that they are at least of order 6 in eccentricity so we can derive a bound on its norm using Taylor's theorem and Cauchy's inequalities for an analytic function. Since the transformation $\tilde{\psi}_2$ is linear, we can write:

$$\begin{aligned}
\|R_3\|_{r_1,s_1/2,r_2,s_2} &= \left\| H_{\geq 6,f} \circ \tau \circ f \circ \tilde{\psi}_1 \right\|_{r_1,s_1/2,r_2,s_2} \\
&= \|H_{\geq 6} \circ \Delta\|_{r_1,s_1/2,r_2,s_2} \\
&= \|H_{\geq 6}\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}}
\end{aligned}$$

We will therefore use a Taylor theorem, for a function of 4 variables $(\xi_1, \xi_2, \eta_1, \eta_2)$, and evaluate its remainder at the order 6. The variables ξ_i and η_i belong to the set $\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}$, and they can be bounded by $\rho_0 + \rho_0'$. We have

$$\|R_3\|_{r_1,s_1/2,r_2,s_2} \leq \sum_{|\beta|=6} \left(\sup_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}} \left| \frac{1}{\beta!} \partial_\beta \bar{H}_{pert} \right| (\rho_0 + \rho_0')^6 \right),$$

where $\beta = (\beta_1, \beta_2, \beta_3, \beta_4) \in \mathbb{N}^4$, $|\beta| = \beta_1 + \beta_2 + \beta_3 + \beta_4 = 6$, and $\beta! = \beta_1!\beta_2!\beta_3!\beta_4!$. We can bound the derivative using Cauchy's inequality. The terms $\beta!$ then cancel, we are left with:

$$\begin{aligned}
\|R_3\|_{r_1,s_1/2,r_2,s_2} &\leq \sum_{|\beta|=6} \left\| \partial_\beta \bar{H}_{pert} \right\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}} (\rho_0 + \rho_0')^6 \\
&\leq \sum_{|\beta|=6} \|\partial_\beta H_{pert}\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0',s_1}} (\rho_0 + \rho_0')^6 \\
&\leq \sum_{|\beta|=6} \frac{1}{\mu^6} \|\partial_\beta H_{pert}\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0'+\mu,s_1}} (\rho_0 + \rho_0')^6 \\
&\leq \sum_{|\beta|=6} \|H_{pert}\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0'+\mu,s_1}} \left(\frac{\rho_0 + \rho_0'}{\mu} \right)^6 \\
&\leq 84 \|H_{pert}\|_{\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0'+\mu,s_1}} \left(\frac{\rho_0 + \rho_0'}{\mu} \right)^6
\end{aligned}$$

This estimate, with the definition of ρ_0 and ρ_0', completes the proof. □

Only the estimate on the part of order at least 6 in eccentricities suffers a loss of analyticity, the other part being defined explicitly.

Bounds on the remainders of the transitional Hamiltonian P_i

We are interested in bounding the terms $R_2 \circ \varphi^{\epsilon}_{X_2} \circ \varphi^{\epsilon}_{X_3}$ and $R_3 \circ \varphi^{\epsilon}_{X_3}$.

Proposition 4.9. *Under the assumptions of proposition 4.7 and that the Hamiltonian H_{pert} is analytic on the set $\mathcal{D}_{\Lambda_0,r_1,\rho_0,\rho_0'+\mu,s_1}$, then we have the estimates:*

$$\left\|R_2 \circ \varphi^{\epsilon}_{X_2} \circ \varphi^{\epsilon}_{X_3}\right\|_{r_1/32,s_1/16,r_2/8,s_2/8} \leq 4\epsilon_2 \left(\frac{\rho_0 + 2\rho_0'}{\mu}\right)^2,$$
$$\left\|R_3 \circ \varphi^{\epsilon}_{X_3}\right\|_{r_1/32,s_1/16,r_2/8,s_2/8} \leq 4\epsilon_3 \left(\frac{\rho_0 + 2\rho_0'}{\mu}\right)^2.$$

Proof. Using the definition of these variables, and the Cauchy's estimate on the second derivative of P_2 with respect to the variables y_i, the proof is straightforward. □

Bound on the remainder of the Taylor expansion

We are interested in bounding the norm of H_{nl}. In this aim, we will simply do a classical estimate on the remainder of a Taylor expansion, though taking into account the special form of the Hamiltonian $H_{0,3}$.
We have

$$H_{0,3} = H_{Kep} + H_{2,D} \circ \tilde{\psi}_1 + H_4^{\perp} \circ \tilde{\psi}_1 + \bar{P}_{2,0} + \bar{P}_{3,0}.$$

Moreover, the application $\tilde{\psi}_1$ takes the four actions in the set $B_{\Lambda_0}(r_1, s_1) \times \tilde{D}_{pol,r_2,s_2}$ into the set $\tilde{D}_{cart,\Lambda_0,r_1,\rho,\rho',s_1}$. At the moment, we still use different r_1 and r_2 as analyticity widths when trying to estimate the rests. Though, since H_{Kep} is much larger than the perturbation, and that we need to have $H_{Kep} r_1^2 \sim \epsilon^2$, we will need to fix a smaller value of r_1. Define a new analyticity width $r' < \min(r_1/32, r_2/8)$. We have

$$B_{\Lambda_0}(r', s_1) \times \tilde{D}_{pol,r',s_2} \subset B_{\Lambda_0}(r_1, s_1) \times \tilde{D}_{pol,r_2,s_2}$$
$$\tilde{\psi}_1\left(B_{\Lambda_0}(r', s_1) \times \tilde{D}_{pol,r',s_2}\right) \subset \tilde{D}_{cart,\Lambda_0,r_1,\rho,\rho',s_1}.$$

We have the proposition:

Proposition 4.10. *Assume $H_{0,3}$ is defined as before, and analytic on the domain $\tilde{D}_{cart,\Lambda_0,r_1,\rho,\rho',s_1}$. Assume that it verifies the assumptions of proposition 4.7. Call*

$$D_1 = \max_{i=1,2}\left(\sup_{p \in B_{\Lambda_0}(r_1,s_1)} \left|\left(H''_{Kep}(p)\right)_{i,i}\right|\right),$$
$$D_2 = \max_{1 \leq i,j \leq 4} \left\|\left(\left(H_{2,D} + H_4^{\perp}\right) \circ \tilde{\psi}_1\right)''_{i,j}\right\|_{r_1,s_1,r_2,s_2}.$$

We have the inequality:

$$\|H_{nl}\|_{r',s_1,r',s_2} \leq \left(D_1 + 8D_2 + 16\frac{\epsilon_2 + \epsilon_3}{(\min(r_1/32, r_2/8) - r')^2}\right) r'^2. \tag{4.24}$$

Proof. The proof is straightforward: one has to compute every term of the Hamiltonian H_{nl} by considering the Taylor expansion of order 2 of each terms, and eventually using Cauchy's

inequality.
First:

$$H_{nl}(I', p_0) = \int_0^1 (1-t) H''_{0,3}(p_t) \cdot I'^2 dt$$

The Kepler Hamiltonian not depending on the two last actions, and its Hessian being diagonal, a bound for the first term is:

$$\left\| \int_0^1 (1-t) H''_{Kep}(p_t) \cdot I'^2 dt \right\|_{r', s_1, r', s_2} \leq \max_{i=1,2} \left(\sup_{p \in B_{\Lambda_0}(r_1, s_1)} \left| \left(H''_{Kep}(p) \right)_{i,i} \right| \right) r'^2 = D_1 r'^2$$

We will just recall the formulas for the two Hamiltonian $H_{2,D}$ and $H_4^\perp$, their second derivative being too long to express explicitly here. Recall that the constants $d_{i,j,k}$ depend on the variables Λ_i. We have:

$$\begin{aligned} \left(H_{2,D} + H_4^\perp \right) \circ \tilde{\psi}_1(\Lambda_1, \Lambda_2, I_3, I_4) = & d'_{0,0,0} - d'_{1,0,0} I_3 - d'_{0,1,0} I_4 \\ & + d'_{2,0,0} I_3^2 + d'_{0,2,0} I_4^2 + (d'_{1,1,0} + 2d'_{0,0,2}) I_3 I_4 \end{aligned}$$

The terms implied in the Hessian of the latter function will roughly be of the same order, and therefore we can bound them as follows:

$$\left\| \int_0^1 (1-t) \left((H_{2,D} + H_4^\perp) \circ \tilde{\psi}_1 \right)'' (p_t) \cdot I'^2 dt \right\|_{r_1, s_1, r_2, s_2} \leq 8 D_2 r'^2,$$

where the factor 8 comes from the number of terms coming from the matrix times the integration factor $\frac{1}{2}$. As for the terms $\bar{P}_{2,0}$ and $\bar{P}_{3,0}$, using Cauchy's inequality gives the result directly. □

Final bound on an isotropic domain

After bounding the different terms of the perturbation, we can now fix our analyticity width, while setting the analyticity width necessary on the perturbation to evaluate those terms.
Since our KAM theorem is isotropic in the actions, and in the angles, we wish to fix $r' = 4r_2$ and $s_1 = 2s_2$, so as to have $r'/32 = r_2/8$ and $s_1/16 = s_2/8$. This assumption is required to make the domain isotropic on the actions and the angles when applying the KAM theorem. Now, the requirements of propositions 4.7, 4.8, 4.9 and 4.10 are satisfied if we suppose that for some $\mu > 0$, H_{pert} is analytic on $\mathcal{D}_{\Lambda_0, r_1, \rho_0, \rho'_0 + \mu, s_1}$. The assumptions then need to be checked one by one, using the construction of $K_2, \epsilon_2, K_3, \epsilon_3, K_4, \epsilon_4$.
The perturbation we will consider will, under all the assumptions made, be of size

$$\epsilon = \epsilon_4 + S_1 \left(\frac{\rho_0 + 2\rho'_0}{\mu} \right)^6 + 4(\epsilon_2 + \epsilon_3) \left(\left(\frac{\rho_0 + 2\rho'_0}{\mu} \right)^2 + 4 \left(\frac{r'}{r_1 - r'} \right)^2 \right) + (D_1 + 8D_2) \left(\frac{r'}{r_1 - r'} \right)^2$$

on the set $\mathcal{D}_{\Lambda_0, r', \rho_0, \rho'_0, s_1}$, with

$$S_1 = C_1 + C_2 + \frac{96 \times 7^7}{\mu^6} \left\| H_{pert} \right\|_{\mathcal{D}_{\Lambda_0, r_1, \rho_0, \rho'_0 + \mu, s_1}}.$$

The transformation is then defined on the set $B_{\Lambda_0}(r_1/32, s_1/16) \times \tilde{D}_{pol, r_1/32, s_1/16}$. After computing this sum, if it satisfies the KAM assumptions, then one can apply KAM theorem to the Hamiltonian $H_{0,3} + P$.
There remains still one value to compute, that is the analyticity width in the frequencies.

4.5 Analyticity width in the frequencies

4.5.1 Upper bound of the analyticity width

In this section, we wish to find the analyticity width in frequencies, *ie* the value of h in our KAM theorem. Assume that the perturbation is analytic on the domain $B_{\Lambda_0}(r_f, s_f) \times \tilde{D}_{pol,r_f,s_f}$ with $r_f, s_f > 0$ (for instance $r_f = r_1/32$ and $s_f = s_1/16$). Recall that the frequency vector is defined in the following way:

$$\omega = H'_{0,3}(p),$$

where $(p, 0)$ belongs to the set $B_{\Lambda_0}(r_f, s_f) \times \tilde{D}_{pol,r_f,s_f}$, in order to be defined. Let us cut the analyticity width r_f in three parts: first consider the set

$$\Omega = \left\{\omega \in \mathbb{R}^n,\ \exists (p,0) \in B_{\Lambda_0}\left(\frac{r_f}{4}, s_f\right) \times \tilde{D}_{pol,r_f/4,s_f} \text{ such that } \omega = H'_{0,3}(p)\right\} \tag{4.25}$$

This set will be the initial set of frequencies. More precisely, we will consider the set of frequencies $\Gamma_\gamma = \Omega \cap D(\gamma, \tau)$, the frequencies in the set Ω that verify the Diophantine condition with constants γ and τ. Now, recall the definition of the set of frequencies with analyticity width h:

$$\Omega_\gamma^h = \left\{\omega \in \mathbb{C}^n,\ \exists \omega' \in \Gamma_\gamma \text{ s.t. } \sup_{i \in [\![1,4]\!]} |\omega_i - \omega'_i| < h\right\}.$$

This set consists in the neighborhood of the tori of frequency in Γ_γ. We will fix the value of h so that the image of this set by the function $H'^{-1}_{0,3}$ is contained in the set $B_{\Lambda_0}(r_f/2, s_f) \times \tilde{D}_{pol,r_f/2,s_f}$. Finally, we will define a new set for the action-angle variables, close to zero, as follows:

$$D_{r_f/2,s_f} = \left\{(I', \theta) \in \mathbb{C}^4 \times \mathbb{T}^n_{\mathbb{C}},\ \sup_{i \in [\![1,4]\!]} |I'_i| < \frac{r_f}{2},\ \sup_{i \in [\![1,4]\!]} |\Im \theta_i| < s_f\right\}$$

Now, we can take $(\omega, I', \theta) \in \Omega_\gamma^h \times D_{r_f/2,s_f}$, it guarantees that:

$$(H'^{-1}_{0,3}(\omega) + I', \theta) \in B_{\Lambda_0}(r_f, s_f) \times \tilde{D}_{pol,r_f,s_f},$$

with a slight abuse of notation coming from the fact that the action-angles variables are not in the right order. Considering an analyticity width $r' < r_f$ for the actions I', we will be able to apply the KAM theorems using these sets, the only necessary step being to find a precise bound for h.

Determining a suitable value of h will be done in several steps. First, let us expand the definition of Ω_h:

$$\omega \in \Omega_\gamma^h \Leftrightarrow \exists\ \omega_0 \in \Gamma_\gamma,\ \vartheta \in B(0,h) \subset \mathbb{C}^4,\ \omega = \omega_0 + \vartheta$$

Now we ask that $H'^{-1}_{0,3}(\omega) \in B_{\Lambda_0}(r_f/2, s_f) \times \tilde{D}_{pol,r_f/2,s_f}$, therefore, we want the following condition to be verified:

$$H'^{-1}_{0,3}(\omega_0 + \vartheta) \in B_{\Lambda_0}(r_f/2, s_f) \times \tilde{D}_{pol,r_f/2,s_f}$$

With the definition of the set Ω, and using a Taylor expansion to the first order, we require:

$$\left\| \int_0^1 (H'^{-1}_{0,3})'(\omega_0 + t\vartheta) \cdot \vartheta dt \right\| < \frac{r_f}{4}, \tag{4.26}$$

where $\|\cdot\|$ represents the sup-norm of our complex vector. Let us single out the coefficients of our 4-vector: let $\vartheta = (\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4) \in B(0, h)$, we have for $1 \le i \le 4$:

$$\begin{aligned}
\left|\left(\int_0^1 (H_{0,3}'^{-1})'(\omega_0 + t\vartheta)\cdot\vartheta dt\right)_i\right| &\le \int_0^1 \sum_{1\le j\le 4}\left|\left((H_{0,3}'^{-1})'(\omega_0+t\vartheta)\right)_{i,j}\vartheta_j\right| dt \\
&\le 4h \times \int_0^1 \max_{1\le j\le 4}\left|\left((H_{0,3}'^{-1})'(\omega_0+t\vartheta)\right)_{i,j}\right| dt \\
&\le 4h \times \max_{1\le j\le 4}\left(\sup_{\Omega_h}\left|\left((H_{0,3}'^{-1})'\right)_{i,j}\right|\right) \\
&\le 4h \times \max_{1\le j\le 4}\left(\sup_{\Omega_h}\left|\left(H_{0,3}''^{-1}\circ H_{0,3}'^{-1}\right)_{i,j}\right|\right) \\
&\le 4h \times \max_{1\le j\le 4}\left(\sup_{B_{\Lambda_0}(r_f/2,s_f)\times\tilde{D}_{pol,r_f/2,s_f}}\left|\left(H_{0,3}''^{-1}\right)_{i,j}\right|\right)
\end{aligned}$$

Therefore we deduce:

$$\left\|\int_0^1 (H_{0,3}'^{-1})'(\omega_0+t\vartheta)\cdot\vartheta dt\right\| \le 4h \times \max_{1\le i,j\le 4}\left\|\left(H_{0,3}''^{-1}\right)_{i,j}\right\|_{r_f/2,s_f,r_f/2,s_f}$$

Thus, a condition on h so that our initial requirement holds is:

$$h = \frac{r_f}{16}\left(\max_{1\le i,j\le 4}\left\|\left(H_{0,3}''^{-1}\right)_{i,j}\right\|_{r_f/2,s_f,r_f/2,s_f}\right)^{-1} \tag{4.27}$$

In the KAM theorem, we require furthermore that the set of initial frequencies is at least at distance h from the boundary. In the ball we just found, this is therefore not possible. Hence we have to define $h' = h/2$ as the analyticity width in frequencies to be sure this set is non-empty. With the latter value of h', it will be possible to apply the KAM theorem. The last calculation left is the determination of the value of the maximal coefficient of the inverse of the Hessian of $H_{0,3}$, that we will discuss in the next part.

Before, let us discuss quickly the previous result, and another possibility to find a value of h. In the previous calculation we used the supremum norm on the vectors of $\mathbb{C}^4$. By equivalence of all the norms in our space, we could have used the euclidean norm. By this mean, instead of the sum over all the modulus of the coefficients of $\left(\left((H_{0,3}'^{-1})'(\omega_0+t\vartheta)\right)_{i,j}\vartheta_j\right)_{1\le i\le 4}$, we would have found an upper bound of the euclidean norm of the vector ϑ, multiplied by the square root of the spectral radius of the matrix $\left(H_{0,3}''^{-1}\right)^* H_{0,3}''^{-1}$. Since determining $H_{0,3}''^{-1}$ is not straightforward, we could instead use the minimum of the modulus of the eigenvalues of the matrix $H_{0,3}''$. Now, notice that for small analyticity width, and in the case of the ratio between the semi-major axis a_1/a_2 is small, we can see the matrix $H_{0,3}''$ as a perturbation of the following matrix:

$$\begin{pmatrix}
-\frac{\alpha_1}{\Lambda_{1,0}^4} & 0 & 0 & 0 \\
0 & -\frac{\alpha_2}{\Lambda_{2,0}^4} & 0 & 0 \\
0 & 0 & -\alpha_3\frac{\Lambda_{1,0}^2}{\Lambda_{2,0}^6} & \alpha_4\frac{\Lambda_{1,0}^3}{\Lambda_{2,0}^7} \\
0 & 0 & \alpha_4\frac{\Lambda_{1,0}^3}{\Lambda_{2,0}^7} & \alpha_5\frac{\Lambda_{1,0}^4}{\Lambda_{2,0}^8}
\end{pmatrix},$$

where the α_i depend only on the masses of the bodies. This matrix has a determinant strictly negative and is symmetric, therefore has real eigenvalues that can be easily computed. In order to determine the smallest eigenvalue of the complete matrix $H''_{0,3}$, one can therefore use the Bauer-Fike theorem on the perturbation of the eigenvalues of a matrix, which gives an upper bound for the deviation of the eigenvalues from the initial matrix given the size of the perturbation. With this result, one can possibly derive a value of h.
However, the calculation required to obtain a value of h by this mean does not seem significantly less important than with the method we are going to describe, that is why we will not follow this path.

4.5.2 Calculation of an explicit upper bound

The aim of this part is to describe a way to derive a value of h, or more precisely of

$$\max_{1\leq i,j\leq 4}\left\|\left(H''^{-1}_{0,3}\right)_{i,j}\right\|_{r_f/2,s_f,r_f/2,s_f}.$$

To find the value of the coefficient of this (inverse) matrix, we will use the equality:

$$H''^{-1}_{0,3}=\frac{1}{\det(H''_{0,3})}adj(H''_{0,3}).$$

The adjugate adj of a matrix is by definition the transpose of the cofactor matrix. To simplify the problem of determining a bound on the maximal coefficient, we will make a further simplification:

$$\max_{1\leq i,j\leq 4}\left\|\left(H''^{-1}_{0,3}\right)_{i,j}\right\|_{r_f/2,s_f,r_f/2,s_f}\leq \left\|\left(\det(H''_{0,1})\right)^{-1}\right\|_{r_f/2,s_f,r_f/2,s_f}\times\max_{1\leq i,j\leq 4}\left\|\left(adj(H''_{0,3})\right)_{i,j}\right\|_{r_f/2,s_f,r_f/2,s_f}.$$

Now let us discuss these two values, and how to bound them.
First, the determinant of this 4×4 matrix is composed of 24 terms. However, since the Hamiltonian is composed of the Kepler part, that is composed of two terms, depending only on one variable, and a small part coming from the perturbation, we can deduce that the largest term of the determinant is the following one:

$$f=\left(H''_{0,3}\right)_{1,1}\left(H''_{0,3}\right)_{2,2}\left(\left(H''_{0,3}\right)_{3,3}\left(H''_{0,3}\right)_{4,4}-\left(H''_{0,3}\right)^2_{3,4}\right)$$

The other terms will be at most the product of one of the large terms $\left(H''_{0,3}\right)_{1,1}$ or $\left(H''_{0,3}\right)_{2,2}$, and of three other small terms of the size of the perturbation. Therefore, we can bound the determinant by the supremum of the largest term minus 22 times the $\left(H''_{0,3}\right)_{1,1}$ multiplied by the highest remaining term to the power 3. In case the perturbation and r_f are small enough, we will have the following upper bound on the inverse of the determinant:

$$\max_{1\leq i,j\leq 4}\left\|\left(H''^{-1}_{0,3}\right)_{i,j}\right\|_{r_f/2,s_f,r_f/2,s_f}\leq\frac{1}{2}\inf_{B_{\Lambda_0}(r_f/2,s_f)\times\tilde{D}_{pol,r_f/2,s_f}}|f|$$

We will not give here a condition on the size of the perturbation nor the value of r_f so that this inequality holds, we will just insure that it is verified when computing it.
Let us now take a closer look at the adjugate matrix. By the same reasoning as previously, we can see that the largest cofactors will be the one involving the two terms $\left(H''_{0,3}\right)_{1,1}$ or $\left(H''_{0,3}\right)_{2,2}$. Thus, we are looking for an upper bound on the terms on the lower square of the adjugate matrix. Looking at the coefficient $(3,3),(4,4),(3,4)$ of the matrix $H''_{0,1}$, we can see by the power

on the term Λ_1/Λ_2 that the term with highest modulus will be the term $\left(H''_{0,3}\right)_{3,3}$. It implies that the largest term of the adjugate matrix will be in position $(4,4)$. As well, for small enough perturbation and analyticity width, we can derive:

$$\max_{1\leq i,j\leq 4}\left\|\left(adj(H''_{0,3})\right)_{i,j}\right\|_{r_f/2,s_f,r_f/2,s_f} \leq \frac{1}{2}\left\|\left(H''_{0,3}\right)_{1,1}\right\|_{r_f/2,s_f,r_f/2,s_f}\left\|\left(H''_{0,3}\right)_{2,2}\right\|_{r_f/2,s_f,r_f/2,s_f}\left\|\left(H''_{0,3}\right)_{3,3}\right\|_{r_f/2,s_f,r_f/2,s_f}.$$

In practice, giving an upper bound on all this terms will be feasible, so we will not go any further in the calculations.
By this mean we can deduce the value of the missing analyticity width h, and we will apply our theorem on the set $\Omega_h \times D_{r_f/2,s_f}$.

Chapter 5

On the Diophantine Condition

When stating the KAM theorem, we saw that it was relying on an arithmetic condition, called the Diophantine condition. This Diophantine condition appears when solving the cohomological equation, to ensure that the solution is still analytic on some smaller set. We will call Diophantine vectors the vectors of $\mathbb{R}^n$ verifying a Diophantine condition. The fact that these Diophantine vectors have full Lebesgue measure in $\mathbb{R}^n$ is substantial in KAM theory. In different classes of differentiability, there is a strong link between the arithmetic condition, and the possibility of application of such a theorem to a Hamiltonian that is the perturbation of an integrable one. To understand better these relations and see examples or counter-examples, the reader can peruse [72]. The 2-dimensional Diophantine condition is a significant particular case, since it allows one to use of a lot of mathematical tools. Indeed, dividing the coefficients of the vector by the largest coefficient, one is reduced to the classical problem of the approximation of a number by rationals. In this case, the Diophantine vector (or number) corresponds to an irrational number that is badly approximated by rationals. One can consult the pedagogical paper [25] so as to understand physically the problem emerging from a non-Diophantine vector in a celestial mechanics problem. A tool for the study of such numbers is for example the decomposition in continued fraction of a real number, one can refer to [28] to understand this operation better. However, when increasing the dimension (generally for $n \geq 3$), the tools used in the study of the approximation of an irrational by rationals do not generalize easily. In our application of the KAM theorem, we use two Diophantine conditions, one in dimension 2 and one in dimension 4. We will therefore, in this chapter, present and justify the existence of Diophantine vectors in the regions we consider.

5.1 Abundance and basic properties

5.1.1 Abundance of Diophantine vectors

For $n \in \mathbb{N} \setminus \{0\}$, and $\gamma, \tau > 0$, define:

Definition 5.1.

$$\begin{aligned} D(\gamma,\tau) &= \left\{ \omega \in \mathbb{R}^n,\ |k \cdot \omega| > \frac{\gamma}{|k|_1^\tau}, k \in \mathbb{Z}^n \setminus \{0\} \right\}, \\ D(\tau) &= \bigcup_{\gamma > 0} D(\gamma,\tau), \\ Dioph &= \bigcup_{\tau > 0} D(\tau) \end{aligned} \tag{5.1}$$

where $|k|_1 = \sum_{i=1}^n |k_i|$ for $k = (k_1, k_2, ..., k_n) \in \mathbb{Z}^n$ and "$\cdot$" is the canonical scalar product of two vectors.

The question of the abundance of the Diophantine vectors (vectors belonging to the set *Dioph*) has been studied in the field of number theory. We will briefly give the most important results here, as a starter (the reader can find a deeper study of these vectors in [28, 48]).
We will consider our space to be $E = \mathbb{R}^n$ with $n \geq 1$ in the whole chapter. Let us remind some definitions.

Definition 5.2. Let A be a subset of E. A is said to have null Lebesgue measure if and only if for every $\epsilon > 0$, there is a countable collection of n-balls $I_n \subset \mathbb{R}^n$, for which the sum of their volume is less than ϵ, and such that their union covers A.
A set B is of full Lebesgue measure if and only if its complementary is of null Lebesgue measure.

Definition 5.3. Let A be a subset of E. A is said to be meagre if it can be expressed as the countable union of subsets that are nowhere dense in E.

Those definitions give two ways of thinking the abundance of a type of numbers in $\mathbb{R}^n$. The scarcity in term of Lebesgue measure is expressed by the fact of being of null measure, whereas in terms of Baire spaces, it corresponds to being meagre. We have the following results:

Proposition 5.4. *The set of Diophantine numbers is meagre and of full Lebesgue measure.*

Proof. Let us verify first that *Dioph* is meagre. For this purpose, let us write it as follows:

$$Dioph = \left\{ \omega \in E,\ s.t.\ \exists \tau \in \mathbb{N},\ \exists n \in \mathbb{N},\ \forall k \in \mathbb{Z}^n \setminus \{0\} :\ |k \cdot \omega| \geq \frac{1}{n|k|_1^\tau} \right\}.$$

Under this form, it is clear that the set of Diophantine numbers is the union of closed sets indexed by τ and n which interior is the empty set. Therefore, *Dioph* is meagre.
To show that *Dioph* is of full Lebesgue measure, it is sufficient to show that for some $\tau > 0$, $D(\tau)$ is of full Lebesgue measure, or equivalently, that $E \setminus D(\tau)$ is of null Lebesgue measure. To this end, consider the Diophantine numbers of dimension n in the set $[0,1]^n$. We can define:

$$\begin{aligned} C(\gamma,\tau,k) &= \left\{ \omega \in [0,1]^n,\ |k \cdot \omega| \leq \frac{\gamma}{|k|_1^\tau} \right\} \qquad (5.2) \\ C(\gamma,\tau) &= \bigcup_{k \in \mathbb{Z}^n \setminus \{0\}} C(\gamma,\tau,k) = [0,1]^n \setminus (D(\gamma,\tau) \cap [0,1]^n) \end{aligned}$$

For $k \in \mathbb{Z}^n \setminus \{0\}$, the set $C(\gamma,\tau,k)$ is contained between two hyperplanes separated by a distance $\frac{2\gamma}{|k|_1 \times |k|_1^\tau}$, its Lebesgue measure in the set $[0,1]^n$ is therefore $\lambda(C(\gamma,\tau,k)) = \frac{2\gamma}{|k|_1^{\tau+1}}$. Let us now write:

$$\begin{aligned} C(\gamma,\tau) &= \bigcup_{k \in \mathbb{Z}^n \setminus \{0\}} C(\gamma,\tau,k) \\ &= \bigcup_{l \in \mathbb{N}^\star} \bigcup_{k, |k|_1 = l} C(\gamma,\tau,k) \end{aligned}$$

The number of vectors $k \in \mathbb{Z}^n \setminus \{0\}$ verifying the condition $|k|_1 = l$ can be bounded quite easily. One has to fix one of the components of k to the value $\pm l$, and then to find how many vectors of dimension $n-1$ one can write such that their components take their values between $-l$ and

l. Therefore, the cardinal of this set is bounded by $4^n l^{n-1}$. We then have:

$$\begin{aligned}\lambda(C(\gamma,\tau)) &\leq \sum_{l\in\mathbb{N}^\star}\sum_{k,|k|_1=l}\lambda(C(\gamma,\tau,k))\\ &\leq \sum_{l\in\mathbb{N}^\star}\sum_{k,|k|_1=l}\frac{2\gamma}{|k|_1^{\tau+1}}\\ &\leq \sum_{l\in\mathbb{N}^\star}4^n\times l^{n-1}\frac{2\gamma}{l^{\tau+1}}\\ &< 2^{2n+1}\gamma\sum_{l\in\mathbb{N}^\star}\frac{l^{n-1}}{l^{\tau+1}}\\ &< 2^{2n+1}\gamma\sum_{l\in\mathbb{N}^\star}l^{n-\tau+2}\end{aligned}$$

Choosing $\tau = n$, we conclude:

$$\lambda(C(\gamma,\tau)) < \frac{4^n}{3}\pi^2\gamma \tag{5.3}$$

Therefore, $\lambda(D(\gamma,\tau)\cap[0,1]^n) = 1 - 4^n\pi^2\gamma/3$ for γ sufficiently small. Taking the union for all $\gamma > 0$, we deduce that $D(\gamma,\tau)\cap[0,1]^n$ has measure 1 in $[0,1]^n$. This result being independent of the subset considered ($[0,1]^n$ here), we showed that $D(\tau)$ and therefore *Dioph* have full Lebesgue measure. □

In our case, the application of the KAM theorem relies on the knowledge of the explicit values of the constants appearing in the problem, such as the constants γ and τ. In the proof of the last proposition, we can see that for $\tau > n-1$, almost every real belongs to the set $D(\tau)$. It means that for γ small enough, the set of Diophantine numbers in $D(\gamma,\tau)$ is of measure strictly positive.
When considering a random non-empty set E in $\mathbb{R}^n$ and some $\tau > n-1$, we would like to know for what value of γ it is true that $D(\gamma,\tau)\cap A \neq \emptyset$. To simplify the result, we will suppose that $E \subset [0,1]^n$, as it does not change fundamentally the result (it is only a matter of rescaling). We have the corollary:

Corollary 5.5. *Let $n \geq 1$, $\tau = n$, and $E \subset [0,1]^n$ such that its Lebesgue measure verifies $\lambda(E) > 0$. If*

$$\gamma \leq \frac{3\lambda(E)}{4^n\pi^2}, \tag{5.4}$$

then we have $E\cap D(\gamma,\tau) \neq \emptyset$.

Proof. This lemma follows from the proof of the previous theorem. Indeed, in the last section, we proved the following result:

$$\lambda(C(\gamma,\tau)) < \frac{4^n}{3}\pi^2\gamma$$

Therefore, if $\lambda(C(\gamma,\tau)) < \lambda(E)$, it follows that $\lambda(E\cap D(\gamma,n)) > 0$. □

As an example, in the case $n = 4$, for $E \subset [0,1]^n$ and $\gamma = \lambda(E)/850$, there exists a Diophantine vector in the set E.

5.1.2 From the 2-dimensional case to an equivalent 1-dimensional case

The particularity of the 2 dimensional case for the Diophantine condition (the 1 dimensional case being trivial) comes from the fact that we can find an equivalent expression of the Diophantine condition, related to the approximation of a number by rationals. This last problem has been studied a lot, and therefore, it allows one to use a lot of different tools when working on the Diophantine vectors of dimension 2.

Definition 5.6. Define:

$$\begin{aligned} D_2(\gamma,\tau) &= \left\{\alpha \in \mathbb{R},\ s.t.\ \forall (p,q) \in \mathbb{Z}\times\mathbb{N}^\star, \left|\alpha - \frac{p}{q}\right| \geq \frac{\gamma}{q^{\tau+1}}\right\} \\ D_2(\tau) &= \bigcup_{\gamma>0} D_2(\gamma,\tau), \\ Dioph_2 &= \bigcup_{\tau>0} D_2(\tau) \end{aligned} \tag{5.5}$$

We will prove the following lemma:

Lemma 5.7. $\forall(\omega_1,\omega_2) \in (\mathbb{R}^+)^2$ *such that* $\omega_1 < \omega_2$*:*

$$(\omega_1,\omega_2) \in D(\gamma,\tau) \Leftrightarrow \frac{\omega_1}{\omega_2} \in D_2\left(\frac{\gamma}{\omega_2},\tau\right)$$

Proof. Assume we have a vector $\omega = (\omega_1,\omega_2) \in (\mathbb{R}^+)^2 \cap D(\gamma,\tau)$, the Diophantine condition is the following:

$$|k_1\omega_1 + k_2\omega_2| \geq \frac{\gamma}{(|k_1|+|k_2|)^\tau}, \qquad \forall (k_1,k_2) \in \mathbb{Z}^2 \setminus \{0\}$$

It is clear that if ω is Diophantine, then neither ω_1 nor ω_2 is null, and they cannot be equal as well. Suppose now that $|\omega_1| < |\omega_2|$. Then, defining $\alpha = \omega_1/\omega_2$, we can divide the preceding condition by ω_2:

$$|k_1\alpha + k_2| \geq \frac{\gamma'}{(|k_1|+|k_2|)^\tau}, \qquad \forall (k_1,k_2) \in \mathbb{Z}^2 \setminus \{0\},$$

with $\gamma' = \gamma/\omega_2$. In the case $k_1 = 0$, the Diophantine condition requires $|k_2| \geq \gamma'/|k_2|^\tau$, or equivalently $|k_2|^{\tau+1} \geq \gamma'$. We obtain $\gamma' \leq 1$. In fact, this upper bound holds for any k_1, and is called the pigeonhole principle of Dirichlet (again, see [28]). Now if $k_1 \neq 0$, for all $k_2 \in \mathbb{Z}$ we have:

$$\left|\alpha + \frac{k_2}{k_1}\right| \geq \frac{\gamma'}{|k_1|(|k_1|+|k_2|)^\tau} \geq \frac{\gamma'}{(|k_1|+|k_2|)^{\tau+1}}$$

Therefore the condition to belong to the set $D_2(\gamma',\tau)$ holds.
For equivalents reason, if ω_1/ω_2 belongs to the set $D_2(\gamma,\tau)$, then the vector (ω_1,ω_2) is in the set $D(\omega_2\gamma,\tau)$. □

The second definition of a Diophantine number is a one-dimensional condition (we got rid of the second one, since only the ratio of the two components of the Diophantine vector matters). The condition given by $D_2(\gamma,\tau)$ can be thought as if we require our Diophantine number to be at least at a distance $1/q^{\tau+1}$ to any rational p/q.
In the three-body problem, it is not always the case that the frequency vector is far away from the rationals (or from the resonant cases). For instance, the system Sun-Jupiter-Saturn is close to the resonance 2:5, which means that the ratio of the periods of rotation of the planets around the sun is close to 2/5.

5.1.3 Continued fraction

When considering a Diophantine number, the main problem lies in the approximation of this number by rationals. This approximation is relative to the denominator we are allowed to work with: the larger the denominator, the closer must be the approximation, as we saw in the different definitions of Diophantine vectors ($\in Dioph$) or numbers ($\in Dioph_2$). Obviously, rationals are well-approximated by a rational: when we authorize the denominator to be large enough, the difference between the initial rational and the rational approximation is zero. The Diophantine vectors are badly approximated by rationals, though it is not the case for every irrational numbers, the latter are called Liouville numbers.
A representation convenient to work with is the continued fraction theory, of which we are going to give the basics here, as well as some results concerning Diophantine numbers. For a more exhaustive reference on this subject, one can consult [33].
The continued fractions of a real number is obtained by an algorithm, that has a finite number of steps if the initial number is rational, and an infinite number otherwise.

Definition 5.8. Continued fraction algorithm: Let $\alpha \in \mathbb{R}$, call $a = \lfloor \alpha \rfloor$, and $b = \alpha - a$. If $b = 0$, then the continued fraction of α is $[a]$, if $b \neq 0$, then the continued fraction of α is $[a, a_1, a_2, ...]$ where $[a_1, a_2, ...]$ is the continued fraction of $1/b$.

Hence, any real number α can be written in the form $\alpha = [a_0, a_1, ..., a_n, ...]$, with $a_i \in \mathbb{N}$ for $i \in \mathbb{N}$. Now, define the following sequences for $n \in \mathbb{N}$:

$$\begin{aligned} p_{-2} &= 0, & p_{-1} &= 1, & p_n &= a_n p_{n-1} + p_{n-2}, \\ q_{-2} &= 1, & q_{-1} &= 0, & q_n &= a_n q_{n-1} + q_{n-2}. \end{aligned}$$

The fractions p_n/q_n for $n \in \mathbb{N}$ are called the convergents of of α. They approach very well the real number α, in fact:

Proposition 5.9. *For $\alpha \in \mathbb{R}$, $n \in \mathbb{N}$, we have:*

$$\left|\alpha - \frac{p_n}{q_n}\right| < \left|\frac{p}{q}\right| \qquad \forall p \in \mathbb{Z}, q \in \mathbb{N} \setminus \{0\}, \text{ such that } q < q_n.$$

This proposition is a corollary of the fact that the following inequality holds:

$$\frac{1}{q_n(q_n + q_{n+1})} < \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}}. \tag{5.6}$$

We will prove this inequality quite easily in the next section. Observe besides that the even convergents are approaching the real number by lower values, whereas the odd convergents are approaching it by upper values.
A useful result in the theory of Diophantine approximation is the following (see [72]):

$$\begin{aligned} \alpha \in D_2(\tau) &\Leftrightarrow q_{n+1} = O(q_n^{\tau+1}), \\ \alpha \in Dioph_2 &\Leftrightarrow \log(q_{n+1}) = O(\log(q_n)). \end{aligned}$$

Loosely speaking, this result means that the sequence $(a_n)_{n\in\mathbb{N}}$ cannot grow too fast. In other words, if we fix the value

$$\gamma_n = q_n^{\tau+1} \times \left|\alpha - \frac{p_n}{q_n}\right|,$$

for the γ_n to decrease at some step n ($\gamma_n = \min_{0\leq i\leq n}(\gamma_i)$), one has to have a_n very large compared to q_n. A number that is not Diophantine is a number for which $\liminf_{n\to\infty} \gamma_n = 0$, and therefore for which a subsequence of the $(a_n)_{n\in\mathbb{N}}$ grows very fast. This discussion enlightens the fact that the complement of the set of Diophantine numbers is of null-measure.

5.2 Stern-Brocot tree and approximations of irrational numbers

The continued fraction theory is a convenient way to understand the real numbers, and how they can be approximated. Another way to see the numbers and their distribution is to work with the Stern-Brocot tree, that is a visual tool allowing to understand better the approximations of real numbers. It will moreover allow us to derive a simple result that will be important when considering the plane planetary three-body problem.

5.2.1 Building the Stern-Brocot tree

The Stern-Brocot tree is a binary tree, introduced independently by Stern, in 1858, and Brocot, in 1861, to compute all the irreducible fractions, in a simple and organized way. We will explain here how to build this tree, and show its basic properties (see [55], [8] and [9] for a deeper study of the subject).

The Stern-Brocot tree is built in an iterative way. First define the following addition between two positive rationals:

Definition 5.10. Let $\frac{p_1}{q_1}$ and $\frac{p_2}{q_2}$ be two positive rationals:

$$\frac{p_1}{q_1} \oplus \frac{p_2}{q_2} = \frac{p_1 + p_2}{q_1 + q_2}.$$

This simple addition rule is the key of the tree we will build (though it would pull out hair of any school teacher). Observe that the result belongs to the interval:

$$\left]\min\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right), \max\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right[.$$

Consider the following fractions that will be the roots of the tree, $\frac{0}{1}$ and $\frac{1}{0}$, and place them on the top left and top right. This choice of roots will allow us to construct every positive real numbers. Choosing $\frac{0}{1}$ and $\frac{1}{1}$ as the roots would give all the rationals in the set $[0, 1]$. One can now add the two roots with the previously defined addition, and put the result in the middle, that is $\frac{1}{1}$. The fraction obtain will be called the child; his parents are the two fractions from which the child is obtained by the previous addition rule (here the parents are the two roots of the tree). The parents and the child are called adjacent. For the next step, take every pair of adjacent fractions ($\left(\frac{0}{1}, \frac{1}{1}\right)$ and $\left(\frac{1}{1}, \frac{1}{0}\right)$), and add them in the same way as before, linking the "youngest child" to the new fractions and each time placing the child in the middle of the two parents. By iterating this process, one can construct a infinite tree, starting from $\frac{1}{1}$, resembling the one in figure 5.1 after 5 steps. We will show the following results:

Proposition 5.11. *Two adjacent fractions* $\frac{p_1}{q_1}$ *and* $\frac{p_2}{q_2}$ *in the Stern-Brocot tree verifies* $p_1q_2 - p_2q_1 = 1$.

Corollary 5.12. *The difference between a parent fraction* $\frac{p_1}{q_1}$ *and a child* $\frac{p_2}{q_2}$ *is equal to* $\frac{1}{q_1q_2}$.

Proposition 5.13. *Every strictly positive irreducible rational is represented in the Stern-Brocot tree.*

Proof. **Proof of proposition 5.11:** Let us show this result by recurrence on the depth of the tree:

- The roots of the tree are $\frac{0}{1}$ and $\frac{1}{0}$, hence it is true at depth 0.
- Assume the proposition to be true at step n of the construction of the Stern-Brocot tree (hence

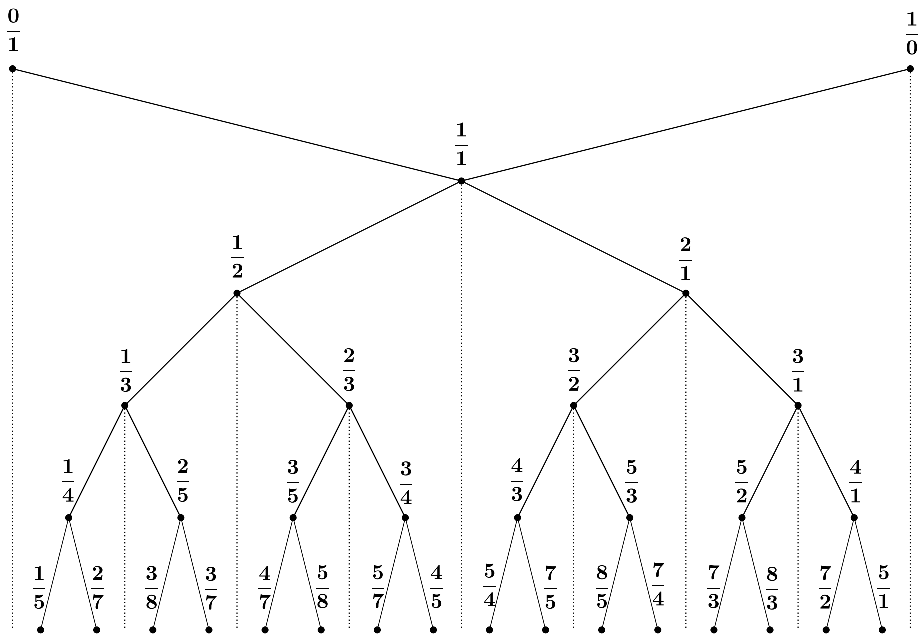

Figure 5.1: Stern-Brocot tree, starting from the two initial roots $\frac{0}{1}$ and $\frac{1}{0}$. The tree shows 5 steps of the main procedure.

for a depth n). Consider two adjacent fractions $\frac{p_1}{q_1}$ and $\frac{p_2}{q_2}$ at the step $n+1$. One is the parent of the other one, for instance let $\frac{p_1}{q_1}$ be the parent. We have

$$\frac{p_2}{q_2} = \frac{p_1 + p}{q_1 + q},$$

where $\frac{p}{q}$ is the other parent. We have:

$$p_1(q_1 + q) - q_1(p_1 + p) = p_1 q - p q_1 = 1,$$

by hypothesis of recurrence, hence the proposition.
The corollary is immediate.
Proof of proposition 5.13: Let $\frac{a}{b}$ be a strictly positive irreducible rational, not represented at step n of the construction of the Stern-Brocot tree. Its value belongs to an interval of adjacent irreducible rationals of the Stern-Brocot tree at step n: $\left]\frac{p_1}{q_1}, \frac{p_2}{q_2}\right[$. At step $n+1$, there exists 3 possible cases:

1. $\dfrac{a}{b} = \dfrac{p_1 + p_2}{q_1 + q_2}$. In this case, the initial rational belongs to the Stern-Brocot tree at the next step.

2. $\dfrac{a}{b} < \dfrac{p_1 + p_2}{q_1 + q_2}$. Then replace $\dfrac{p_2}{q_2}$ by $\dfrac{p_1 + p_2}{q_1 + q_2}$.

3. $\dfrac{a}{b} > \dfrac{p_1 + p_2}{q_1 + q_2}$. Then replace $\dfrac{p_1}{q_1}$ by $\dfrac{p_1 + p_2}{q_1 + q_2}$. The algorithm consists in iterating this process until we are in the first case. This algorithm stops after a finite number of steps. Indeed, we have after some step k:

$$\frac{p}{q} < \frac{a}{b} < \frac{p'}{q'}$$

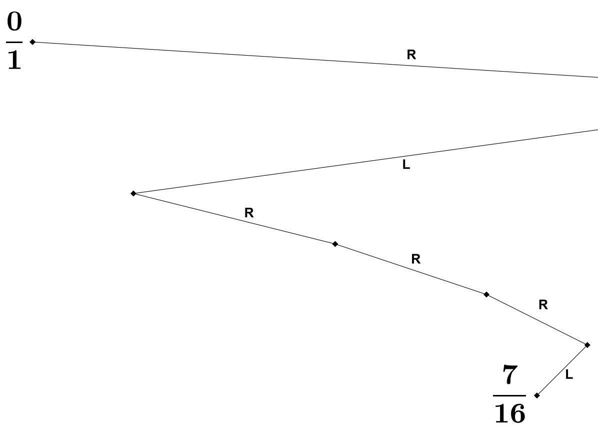

Figure 5.2: Path followed for the word RL^2R^3L: starting from the rational $\frac{0}{1}$, the image of this word is the rational $\frac{7}{16}$.

where the two new fractions are adjacent. Hence:

$$aq - bp \geq 1,$$
$$bp' - aq' \geq 1,$$

from which we deduce:

$$\begin{aligned} a + b &= (p' + q')(aq - bp) + (p + q)(bp' - aq') \\ &\geq p' + q' + p + q. \end{aligned}$$

The two values $p + p'$ and $q + q'$ strictly increasing at each steps, and $a + b$ being fixed, the algorithm stops, hence the proposition. □

Observe that if we remove some branches of the tree we just built, we obtain the Farey sequence at some order, which is the the sequence of irreducible fractions between 0 and 1 which have denominators less than n and is arranged in increasing order.

5.2.2 Binary paths and continued fractions

In the binary Stern-Brocot tree, one can go down the tree by choosing at each level a direction, left or right, which will lead to a rational. Since every irreducible positive rational is represented (by adding the root 0/1 in the tree), one can define a one-to-one correspondence between them and a specific path. Define the following set:

$$\mathcal{M}_{\{L,R\}} = \Big\{ R^{n_1} L^{n_2} \ldots R^{n_i} L^{n_{i+1}} \ldots, \exists l \in \mathbb{N}, (n_i)_{i \in \mathbb{N}} \in (\mathbb{N})^{\mathbb{N}}, \text{ such that} \\ \forall 1 \leq i \leq l,\ n_i > 0, \text{ and } \forall i > l,\ n_i = 0 \Big\}.$$

This set is a set of all the finite words composed of the two letters L and R, starting with letter R (the empty string belongs to this set). Now, for any word in $\mathcal{M}_{\{L,R\}}$, consider the following algorithm starting at the fraction 0/1 of the Stern-Brocot tree: if the first letter of the word is R, then go down to the right child, if it is L, go down to the left child; remove the first letter of the word, and iterate until the word is the empty string. At the end of the process, we associate the word and the rational we are on (see figure 5.2).

Call the function that associates the word to the rational obtained f. Using the proposition 5.13, it is straightforward that:

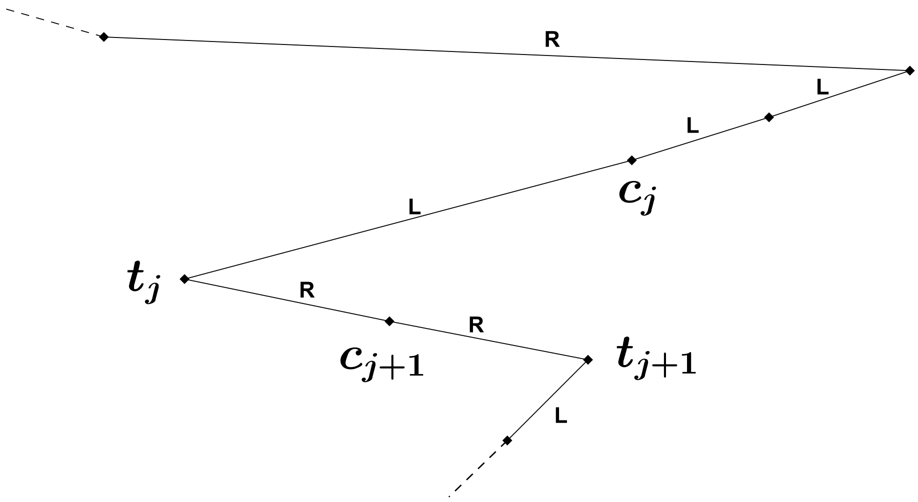

Figure 5.3: Example of convergents and truncations of some order for a specific path in the Stern-Brocot tree. The truncations are the words that finish at some turn, the convergents the word cut right before a turn happens.

Proposition 5.14. *The following function is a one-to-one correspondence:*

$$f: \mathcal{M}_{\{L,R\}} \to \mathbb{Q} \cap \mathbb{R}^+. \tag{5.7}$$

Observe that by taking the symmetric of the tree we have built with respect to the root 0/1, we could have built a left tree with negative numbers, then allowing the words to start by the letter L, and hence the bijection would be a bijection between all the words and $\mathbb{Q}$.

When looking at the set $\mathcal{M}_{\{L,R\}}$, we can define different interesting tools:

Definition 5.15. The length $l(a)$ of a word $a \in \mathcal{M}_{\{L,R\}}$ is defined by $l(a) = \max_{i\in\mathbb{N}}(n_i \neq 0)$ where $a = R^{n_1}L^{n_2}....$ The truncation of order j of a word a is $t_j(a) = a = R^{n_1}L^{n_2}...x^{n_j}$ where $x \in \{L, R\}$. The convergent of order j of a word a is $c_j(a) = a = R^{n_1}L^{n_2}...x^{n_j-1}$ where $x \in \{L, R\}$.

The length of the word corresponds to the number of times we did a turn on the way down, counting from one. The truncation of order j hence corresponds to the fraction located at the j^th turn on the way down. The convergent of order j is the fraction reached just before reaching this turn (see figure 5.3).

Now we can link this construction to the continued fraction.

Proposition 5.16. *Let $a \in \mathcal{M}_{\{L,R\}}$ be defined by $a = R^{n_1}L^{n_2}...R^{n_i}L^{n_{i+1}}...$, then*

$$f(a) = \left[n_1 - 1, n_2, ..., n_{l(a)-1}, n_{l(a)} + 1\right] \tag{5.8}$$

$$f(c_j(a)) = \frac{p_j}{q_j} \tag{5.9}$$

where the last fraction is the convergent of the fraction $f(a)$.

The Stern-Brocot tree is therefore a visualization of the continued fractions, that allows us to see more clearly the different implications of the continued fractions. Note that if one wants to consider only the Diophantine numbers between 0 and 1, it is possible to start from the fraction 1/1, consider only the words that starts by L, and to let the first term of the continued fraction be zero.

Proof. Observe on the Stern-Brocot tree that for a word $a = R^{n_1}...x^{n_{l(a)-1}}y^{n_{l(a)}}$ of length $l(a) > 1$, with $(x, y) = (L, R)$ or (R, L), we have the relation:

$$f(a) = f(R^{n_1}...x^{n_{l(a)-1}}y^{n_{l(a)}-1}) \oplus f(R^{n_1}...x^{n_{l(a)-1}-1}). \tag{5.10}$$

Using this relation, we can proceed by recurrence on the length of the word.
Consider a word $a = R^n$, the previous rule is not direct since $l = 1$. Though observe that we have:

$$f(R^n) = \frac{n \times 1 + 0}{n \times 0 + 1} = \frac{n}{1} = n = (n-1) + 1.$$

The words of length $l = 2$ can be written $R^n L^m$. Using relation (5.10), we see that:

$$\begin{aligned}
f(R^n L^m) &= f(R^{n-1}) \oplus f(R^n L^{m-1}) \\
&= f(R^{n-1}) \oplus f(R^{n-1}) \oplus f(R^n L^{m-2}) \\
&= f(R^{n-1}) \oplus f(R^{n-1}) \oplus ... \oplus f(R^n) \\
&= \frac{n-1}{1} \oplus \frac{n-1}{1} \oplus ... \oplus \frac{n}{1} \\
&= \frac{m \times (n-1) + n}{m+1} \\
&= \frac{(m+1)(n-1)+1}{m+1} = (n-1) + \frac{1}{m+1} \\
f(R^n L^m) &= [n-1, m+1]
\end{aligned}$$

The proposition is proved for words of length $l \leq 2$.
Consider the proposition to be true for a word of length $l > 1$. Let $a = R^{n_1}...x^{n_l}y^{n_{l+1}}$ be a word of length $l + 1$, with $(x, y) = (L, R)$ or (R, L).

$$f(a) = f(R^{n_1}...x^{n_l - 1}) = [n_1 - 1, n_2, ..., n_{l-1}, n_l + 1] = \frac{p_l}{q_l}$$

Using the relation (5.10) n_l times, we have in fact:

$$\begin{aligned}
f(a) &= f(R^{n_1}...x^{n_l}) \oplus \frac{n_l p_l}{n_l q_l} \\
&= f(R^{n_1}...y^{n_{l-1}}x^{n_l - 1}) \oplus f(R^{n_1}...y^{n_{l-1}-1}) \oplus \frac{n_l p_l}{n_l q_l} \\
&= \frac{p_l}{q_l} \oplus \frac{p_{l-1}}{q_{l-1}} \oplus \frac{n_l p_l}{n_l q_l} \\
&= \frac{(n_l+1)p_l + p_{l-1}}{(n_l+1)q_l + q_{l-1}} \\
&= [n_1 - 1, n_2, ..., n_l, n_{l+1} + 1] .
\end{aligned}$$

□

Observe that the fact that the roots of the tree are $\frac{0}{1} = \frac{p_{-2}}{q_{-2}}$ and $\frac{1}{0} = \frac{p_{-1}}{q_{-1}}$ is responsible for this equivalence between the this approach and the continued fraction one. The visualization that offers the Stern-Brocot tree allows to think only with vertices, which represent the words.

5.2.3 Infinite length words and Diophantine condition

Considering the closure of the set $\mathcal{M}_{\{L,R\}}$, we can now work on the irrational numbers with the Stern-Brocot tree. Call

$$\mathcal{M} = Cl\left(\mathcal{M}_{\{L,R\}}\right),$$

we have:

$$f : \mathcal{M} \to \mathbb{R}^+$$

The words we are now looking at are the words $a = R^{n_1}L^{n_2}...R^{n_i}L^{n_{i+1}}...$ with $n_i \in \mathbb{N}$. These words correspond to the (infinite) continued fraction $[n_1 - 1, n_2, ..., n_i, ...]$.
Observe that we can easily prove the inequality (5.6) with the construction of the Stern-Brocot tree. Indeed, any Diophantine vector α with corresponding word a belongs to the interval determined by the image by f of the words $c_j(a)$ and $t_j(a)$, but does not belong to the interval determined by the image by f of the words $c_j(a)$ and $t_{j+1}(a)$ (recall figure 5.3). These two fractions are given by the formulas:

$$f(c_j(a)) = \frac{p_j}{q_j} \qquad f(t_{j+1}(a)) = \frac{p_j + p_{j+1}}{q_j + q_{j+1}},$$

since the parent of $t_{j+1}(a)$ are $c_j(a)$ and $c_{j+1}(a)$. The inequality (5.6) is then straightforward, using corollary 5.12 for the adjacent fractions $(c_j(a), c_{j+1}(a))$ and $(c_j(a), t_{j+1}(a))$.

With the infinite words, we can consider Diophantine numbers. The properties of Diophantine numbers obtained with the help of continued fractions can be demonstrated using the Stern-Brocot tree and the relation between adjacent fractions.
In the Stern-Brocot tree, the Diophantine condition can be visualized by considering the words in which the number of times we go left or right in a row is not too large compared to the depth we are at. Observe for instance that the golden ratio is given by the word $[R^2LRLR...]$.
The frequencies we will be interested in while applying theorem 2.15 are of different order of magnitudes. In this case, one can consider an integer K such that the first frequency is 1, and the second one is close to $1/K$. Using the Stern-Brocot tree construction we will prove two propositions for this kind of frequency distribution.

Proposition 5.17. *Let $K \geq 2$, then the Lebesgue measure of the set:*

$$\left]\frac{1}{K+1}, \frac{1}{K}\right[\cap D_2\left(\frac{1}{K+1}, 2\right)$$

is strictly positive.

We chose to let $\tau = 2$ in this proposition, though any other choice $\tau > 2$ would have been possible (possibly changing the minimum value of K). This proposition means that if we can let the frequency ω change by a small value, then we can find a Diophantine vector close to it, and with a value of γ close to the maximal one (the value of γ is bounded by ω). By symmetry of the denominators in the Stern-Brocot tree, we have as well:

Corollary 5.18. *Let $K \geq 2$, then the Lebesgue measure of the set:*

$$\left]1 - \frac{1}{K}, 1 - \frac{1}{K+1}\right[\cap D_2\left(\frac{1}{K+1}, 2\right)$$

is strictly positive.

To prove the proposition, we will first prove the following lemma.

Lemma 5.19. *Let $\tau > 1$ and $\frac{p}{q} < \frac{r}{s}$ be two adjacent fractions in the Stern-Brocot tree, then the Lebesgue measure of the set*

$$\left\{\omega \in \left]\frac{p}{q}, \frac{r}{s}\right[\text{ such that } \forall a \in \mathbb{Z},\ b \in \mathbb{N} \setminus [\![0, \min(q,s) - 1]\!],\ \left|\omega - \frac{a}{b}\right| \geq \frac{\gamma}{b^\tau}\right\}$$

with γ verifying the following assumption

$$\gamma\left(\frac{2qs}{(q+s)^\tau}+\frac{2}{\tau-2}\frac{1}{(q+s)^{\tau-2}}+\frac{q}{s^{\tau-1}}+\frac{s}{q^{\tau-1}}\right)<1 \tag{5.11}$$

is strictly positive.

Proof. The idea of the proof relies on the study of the complement of this set in the interval $\left]\frac{p}{q},\frac{r}{s}\right[$. The other important idea is that, given the construction of the Stern-Brocot tree, the only intervals we have to consider in the complement are the ones around the fractions of the form $\frac{a}{kq+ls}$ with $k,l\geq 1$. Indeed, since the two initial fractions are adjacent in the Stern-Brocot tree, all children of these rationals are of this form. There still exists some repetitions when considering all these fractions (for instance $\frac{2a}{2(kq+ls)}=\frac{a}{kq+ls}$), but this will be sufficient here. In the initial interval, we remove an interval of size $\frac{\gamma}{q^\tau}$ on the left, and $\frac{\gamma}{s^\tau}$ on the right. The measure of the other intervals we are considering is bounded by:

$$S=\sum_{k,l\geq 1}\frac{2\gamma}{(kq+ls)^\tau}$$

Bounding this term by an integral, we have:

$$\begin{aligned}
S&=2\gamma\sum_{k\geq 1}\left(\frac{1}{(kq+s)^\tau}+\frac{1}{\tau-1}\frac{1}{s(kq+s)^{\tau-1}}\right)\\
&=2\gamma\left(\frac{1}{(q+s)^\tau}+\frac{1}{\tau-1}\left(\frac{1}{q(q+s)^{\tau-1}}+\frac{1}{s(q+s)^{\tau-1}}\right)+\frac{1}{(\tau-1)(\tau-2)}\frac{1}{qs(q+s)^{\tau-2}}\right)\\
&=2\gamma\left(\frac{1}{(q+s)^\tau}+\frac{1}{\tau-2}\frac{1}{qs(q+s)^{\tau-2}}\right).
\end{aligned}$$

The condition on the measure of the set we are interested in to be strictly positive is therefore:

$$\gamma\left(\frac{2}{(q+s)^\tau}+\frac{2}{\tau-2}\frac{1}{qs(q+s)^{\tau-2}}+\frac{1}{q^\tau}+\frac{1}{s^\tau}\right)<\frac{1}{qs}.$$

Hence the lemma. □

One can improve this limit for γ by considering the first intervals of the complement, in order to remove some repetition.
Regarding the Diophantine condition, it is necessary to check if the intervals $\left]\frac{a}{b}-\frac{\gamma}{b^\tau},\frac{a}{b}+\frac{\gamma}{b^\tau}\right[$ for some $b\leq\min(q,s)$ intersect or not the interval $\left]\frac{p}{q},\frac{r}{s}\right[$. If it is not, then a constant γ such that the measure of the set in the lemma is strictly positive implies that the measure of Diophantine vectors in this interval is strictly positive too.
We can now prove the proposition.

Proof. First, observe that for $\gamma=1/(K+1)$:

$$\left]\frac{1}{K+1},\frac{1}{K}\right[\cap\,]-\gamma,\gamma[\,=\emptyset.$$

Hence, for the denominators equal to one, there is no problem. Using lemma 5.19 in the case $\tau=3$, $\frac{p}{q}=\frac{1}{K+1}$ and $\frac{r}{s}=\frac{1}{K}$, the condition (5.11) of the lemma is:

$$\frac{2}{(2K+1)^3}+\frac{2}{K(K+1)(2K+1)}+\frac{1}{K^3}+\frac{1}{(K+1)^3}<\frac{1}{K}.$$

Bounding $K+1$ from below by K, we obtain:

$$\frac{2}{(2K+1)^3} + \frac{2}{K(K+1)(2K+1)} + \frac{1}{K^3} + \frac{1}{(K+1)^3} < \frac{1}{4K^2} + \frac{1}{K^2} + \frac{1}{K^2} + \frac{1}{K^2} < \frac{13}{4K^2}.$$

Since $K \geq 2$, $\gamma = 1/(K+1)$ satisfies indeed the inequality of lemma 5.19.
In all the work we have done, we only considered the denominators of the fractions. By symmetry of the denominators in the Stern-Brocot tree between the fractions $0/1$ and $1/1$, the corollary is immediate. □

Observe on the Stern-Brocot tree the particularity of the golden ratio. At each step, when going down, we change side, "bouncing" from left to right or right to left, therefore avoiding being close to a rational. In this particular case, the optimal constant γ such that the golden ratio $\phi = \frac{1+\sqrt{5}}{2}$ belongs to a set $D(\gamma, \tau)$ for $\tau = 1$ is equal to:

$$\gamma = \frac{3}{2} - \frac{\sqrt{5}}{2}.$$

This limit is reached when considering $\phi - \frac{2}{1}$. The maximal value of γ is therefore reached for a denominator equal to one. We are interested in finding other irrationals of this kind, such that the condition on γ is fixed by the fraction of denominator equal to 1.

Lemma 5.20 (Optimal Diophantine condition in γ and τ)**.** *Let $K \geq 2$, define*

$$\omega = \frac{2K - 1 - \sqrt{5}}{2K^2 - 2K - 2} = \frac{1}{K + (\phi - 1)}, \tag{5.12}$$

then we have $\omega \in D_2(\omega, 2)$.

The irrational ω we have just built corresponds to the word $RL^K RLRLR$..., *i.e.* when reaching the fraction $1/(K+1)$ in the Stern-Brocot tree, then one has to alternate right and left to reach ω.

Proof. For $K \geq 2$, we have obviously:

$$\left|\omega - \frac{0}{1}\right| \geq \frac{\omega}{1^2}$$

Now we need to consider the other convergents of this irrational number. The second one is $\frac{1}{1}$, it is clear from the definition of ω that $1 - \omega > \omega$. The next convergents will be:

$$\frac{1}{K}; \ \frac{1}{K+1}; \ \frac{2}{2K+1}; \ \frac{3}{3K+2}; \ \frac{5}{5K+3} \ldots$$

or by definition, they will be of the form $\frac{p_n}{q_n} = \frac{u_{n+1}}{u_{n+1}K + u_n}$, where u_n is the Fibonacci sequence. We want to check that:

$$q_n^2 \left|\omega - \frac{p_n}{q_n}\right| \geq \omega,$$

or equivalently

$$\left|(u_{n+1}K + u_n)^2 \omega - K u_{n+1}^2 - u_n u_{n+1}\right| \geq \omega.$$

To simplify the discussion, we will check this inequality only for the convergents approaching ω by lower values (the other ones not being important in our case). We will prove that the left term of the previous inequation actually converges to a constant value, not depending on K, and higher than ω. Indeed:

$$\begin{aligned}(u_{n+1}K + u_n)^2\omega - Ku_{n+1}^2 - u_n u_{n+1} &= u_{n+1}^2 \left(\left(K + \frac{u_{n+1}}{u_n}\right)\omega\right)\left(K + \frac{u_{n+1}}{u_n} - \frac{1}{\omega}\right) \\ &= u_{n+1}^2 \left(\left(K + \frac{u_{n+1}}{u_n}\right)\omega\right)\left(\frac{u_n}{u_{n+1}} - (\phi - 1)\right).\end{aligned}$$

The term $\left(K + \frac{u_{n+1}}{u_n}\right)\omega$ converges to one. Now we have to study the following term (independent of K):

$$u_{n+1}^2 \left(\frac{u_n}{u_{n+1}} - (\phi - 1)\right).$$

Using the properties of the Fibonacci sequence, and the definition of ϕ, we obtain:

$$u_{n+1}^2 \left(\frac{u_n}{u_{n+1}} - (\phi - 1)\right) \to \frac{1}{\sqrt{5}} > \omega.$$

Hence, the worst case is the case with denominator 1. □

This lemma ensures the existence of a Diophantine number ω that is optimal in γ and τ in some intervals, *i.e.* $\omega \in D_2(\omega, 1)$. It is therefore possible by changing slightly the initial frequencies to consider that they have a ratio of the previous form. We will use the following definitions when considering Diophantine vectors:

Definition 5.21. The Diophantine vector $\omega = (\omega_1, ..., \omega_n) \in \mathbb{R}^n$ is called optimal in γ and τ if $\omega \in D(\min_i(\omega_i), n)$. It is called optimal in γ for some fixed τ if $\omega \in D(\min_i(\omega_i), \tau)$.

With the previous work, we have the proposition:

Proposition 5.22. *Let $\omega_1 \in \mathbb{R}$ and $\omega_2 = \omega_1 \times (K + \phi - 1)$, with $K \geq 2$, then the vector (ω_1, ω_2) is optimal in γ and τ.*

5.3 Resonances and gaps of Diophantine conditions

In the KAM theorem, for a fixed value τ in the Diophantine condition, some of the hypotheses depend on γ, and if they do, they depend linearly on it. We can therefore wonder how does the distribution of the Diophantine vectors at a fixed γ and τ look like, so as to know what mass would be KAM stable given an initial position. We therefore want to give a physical criterion representing a power of stability in a fixed interval.

Physically, we will consider the following question. Consider a star and a planet, for instance the system Sun-Jupiter. We consider that the conditions to apply the KAM theorem [62] to a third mass are satisfied, and that they are only related to the factor γ (τ is fixed). Under these assumptions, one can model the total mass KAM stable in an interval I by considering the value of:

$$m \sim \gamma \times \lambda(D(\gamma, \tau) \cap I),$$

where λ is the Lebesgue measure.

This modeling suffers a lot of flaws, though it allows to consider a stable mass on a specific interval. On the same interval, for some $\gamma' < \gamma$, we have:

$$(D(\gamma', \tau) \setminus D(\gamma, \tau)) \cap I \neq \emptyset.$$

Hence, we can consider that the mass that would be stable in this interval is closer to the value

$$\begin{aligned} m &\sim \gamma \times \lambda(D(\gamma,\tau)\cap I) + \gamma' \times \lambda\left(D(\gamma',\tau)\setminus D(\gamma,\tau)\right)\cap I) \\ &\sim \gamma' \times \lambda(D(\gamma',\tau)\cap I) + (\gamma-\gamma')\times\lambda(D(\gamma,\tau)\cap I). \end{aligned}$$

Step by step, by decreasing γ, we can obtain the desired value.

Definition 5.23. The power of stability in an interval $I \subset [0,1]$ is the following value:

$$PoS(I) = \frac{1}{\lambda(I)}\int_0^1 \lambda(D(\gamma,\tau)\cap I)d\gamma.$$

We stopped the integral at the value 1, though the maximal value depends on τ.
Given the difficulty of obtaining a precise expression of the measure of the set $D(\gamma,\tau)$, we do an approximation using a Riemann sum, and try to compute this value numerically. Consider a variable $N > 0$, we divide the integral in a sum for different intervals of the value γ: $\left]\frac{k}{N},\frac{(k+1)}{N}\right[$ for $k = 0, ..., N-1$ (we can stop before $N-1$ in the actual computation). We consider as well the sets $D_2(\gamma,\tau)$ instead of $D(\gamma,\tau)$ so as to simplify the computation. Therefore, the planet will have a frequency equal to one around the star, and we put our interest only on the frequency of the other body. The numerical scheme we will apply is the following:
• We first compute all the rational fractions with denominator less or equal to m in the interval $[0,1]$.
• We consider the complementary set of $D_2(k/N,\tau)$, and compute the union of every interval of the form:

$$C_{m,N} = \bigcup_{(p,q):\frac{p}{q}\in\mathbb{Q}\cap[0,1],0<q\leq m}\left]\frac{p}{q}-\frac{k}{Nq^\tau},\frac{p}{q}+\frac{k}{Nq^\tau}\right[.$$

In figure 5.4, we represented the first intervals to eliminate as a function of γ.

• We intersect this set with an interval I, that we will take if the form $I = I_{k',M} = \left]\frac{k'}{M},\frac{k'+1}{M}\right[$ where $M \geq 1$ and $k \in [\![0, M-1]\!]$, and calculate its Lebesgue measure. We then add the results using:

$$PoS(I_{k',M}) \sim M\sum_{k=0}^{N-1}\frac{(k+1/2)}{N}\times\lambda\left(D_2\left(\frac{k}{N},\tau\right)\setminus D_2\left(\frac{(k+1)}{N},\tau\right)\right). \tag{5.13}$$

When computing the formula (5.13) by this mean, we can see that two approximations were made. First, we considered a Riemann sum, which relies on choosing a large value of N to approach the real value of $PoS(I)$. Secondly, we omitted to remove the size of all the intervals around rationals with denominators larger than m. We will justify the computation here by estimating the difference between the truncated sum we just did and the sum over all rationals. Observe that for some $0 < k' < M-1$, the size of the interval $I_{k',M}$ is $1/M$. Hence, there exists only one rational of denominator $M+1$ in this interval. This stays true until we reach the rationals with denominator $2M+1$: then there exists two rationals in this interval. With the same reasoning, there is in fact $\lfloor (l-1)/M \rfloor$ rationals in the interval $I_{k',M}$ with denominator $l > M$. It is not the case for rational with denominators less than M, hence we will fix $m \geq M$. We can now bound the error on the Lebesgue measure of the considered set by adding the size of all the intervals we did not considered, even if some of them are not disjoint. It gives:

$$\begin{aligned} \lambda\left(D_2\left(\frac{k}{N,\tau}\right)\right) - \lambda\left(C_{m,n}\cap\left[\frac{k'}{M},\frac{k'+1}{M}\right[\right) &\leq \sum_{q>m}\frac{2kM}{Nq^{\tau+1}}\times\left\lfloor\frac{q-1}{M}\right\rfloor \\ &\leq \sum_{q>m}\frac{2kM}{NMq^\tau} \\ &\leq \frac{2k}{Nm^{\tau-1}}. \end{aligned}$$

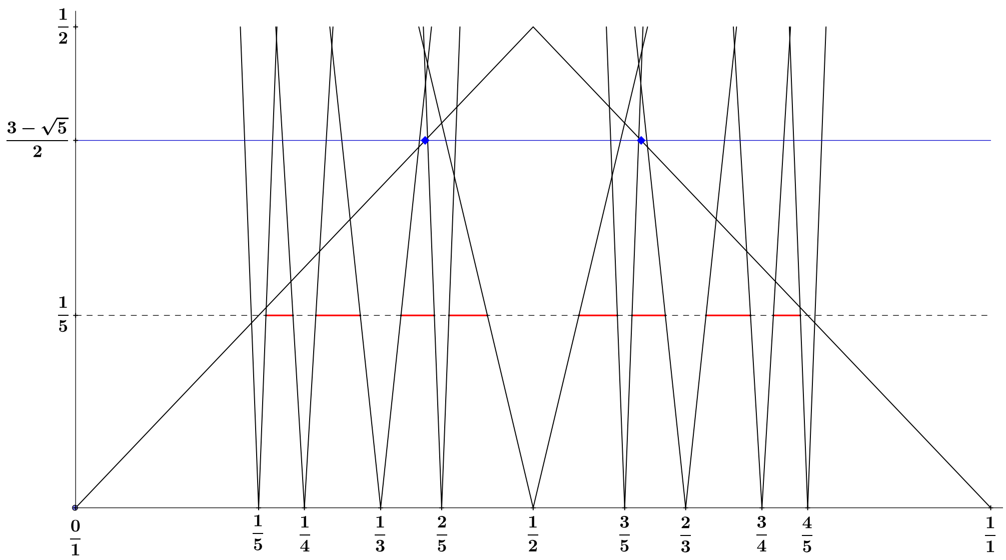

Figure 5.4: In this figure, $\tau = 1$. In abscissa, the interval $[0, 1]$, and in ordinates, the value of γ. The interval to eliminate around a rational is increasing with the value of γ. In red is represented an example for $\gamma = 1/5$, for which we removed the intervals around the rationals of denominator less than 5. In blue is represented the maximum γ can reach for $\tau = 2$, there are two reals in $[0, 1]$ (related to the golden ratio) verifying this condition.

Summing this error over all the $k \leq 4/10N$ (the maximal value for γ is $(3 - \sqrt{5})/2 \sim 0.4$), we obtain the following error on the evaluation of the power of stability:

$$\delta_{m,N}\left(PoS(I_{k',M})\right) \leq \frac{1}{m^{\tau-1}}\left(\frac{3}{15N} + \frac{16}{25} + \frac{16N}{375}\right).$$

Hence, if we choose m sufficiently large compared to N, the error gets small. Observe that the estimate on the error is not optimal. With this error, we need τ to be large enough to get an idea of the real power of stability. However, it is a large upper bound of the true error.

The results are shown in figure 5.5 in the case $\tau = 2$ and for the following values: $N = 400$, $M = 10000$ $m = 10000$. For these values, the error on the power of stability is less than $\delta \leq 0.0018$, and the power of stability varies from the values 0 to 0.225.

We can see in this figure the impact of the resonances on the power of stability of intervals close those resonances. The slope that envelops the data is the resonance 0:1. We observe a large gap for the resonance 1:2, and other gaps at resonances 1:3, 1:4 and 2:5. As expected, near the resonances, the KAM theorem does not imply much stability, since one has to consider very small values of γ to have a set $D_2(\gamma, \tau)$ that is not empty.
This image gives an idea of the KAM stability among the asteroid belt between Mars and Jupiter. We can re-scale this image looking this time at the semi-major axes, using the third Kepler's law. An example is shown in figure 5.6 for the previous values. One can relate very cautiously this figure to the number of asteroids between $2UA$ and $3.5UA$ in the solar system, shown in figure 5.7 (source:NASA [51]).

If some similarities appears, the KAM theorem does not imply anything on possible instabilities. Hence, one should not draw any conclusion from the two previous figures. As well, in the asteroid belt, we are not in the case of a three-body problem. Though, it is important to have an idea of the presence of Diophantine vectors in such a set.
To understand better the decay with τ of the power of stability, or more precisely of its approximation by the previous computation, we drew the figures for $\tau = 1.1$, $\tau = 1.5$, $\tau = 2$, $\tau = 3$,

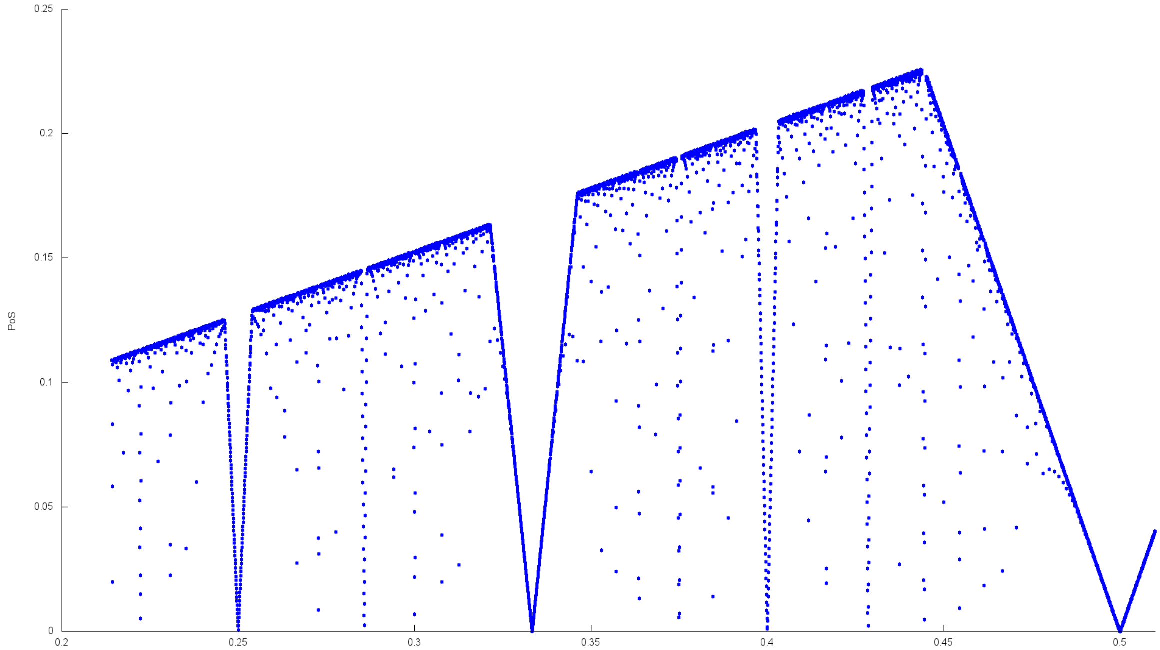

Figure 5.5: Power of Stability for $\tau = 2$ in the interval $[0.2, 0.51]$ for $N = 400$, $M = 10000$ and $m = 10000$, with an error $\delta_{m,N} < 0.0018$.

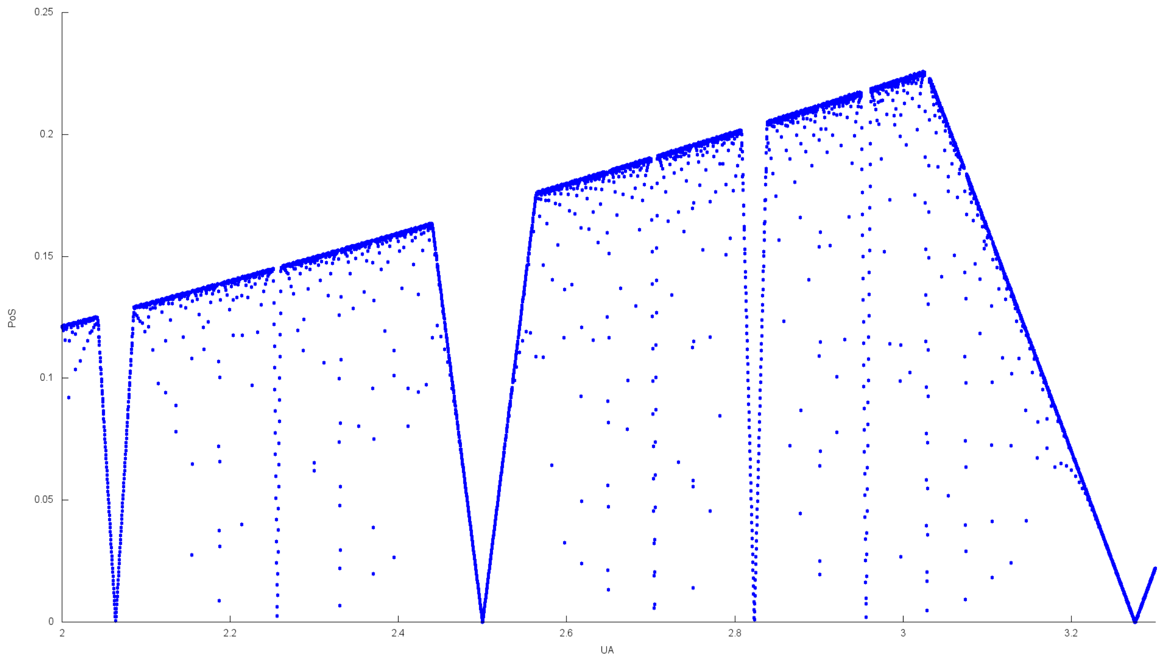

Figure 5.6: Rescaled Power of Stability for $\tau = 2$ and distances to the sun between $2UA$ and $3.3UA$, with $N = 400$, $M = 10000$ and $m = 10000$, and an error $\delta_{m,N} < 0.0018$.

and for the previous values of N, M, m. The error becomes very large for small values of τ, larger than the actual PoS.

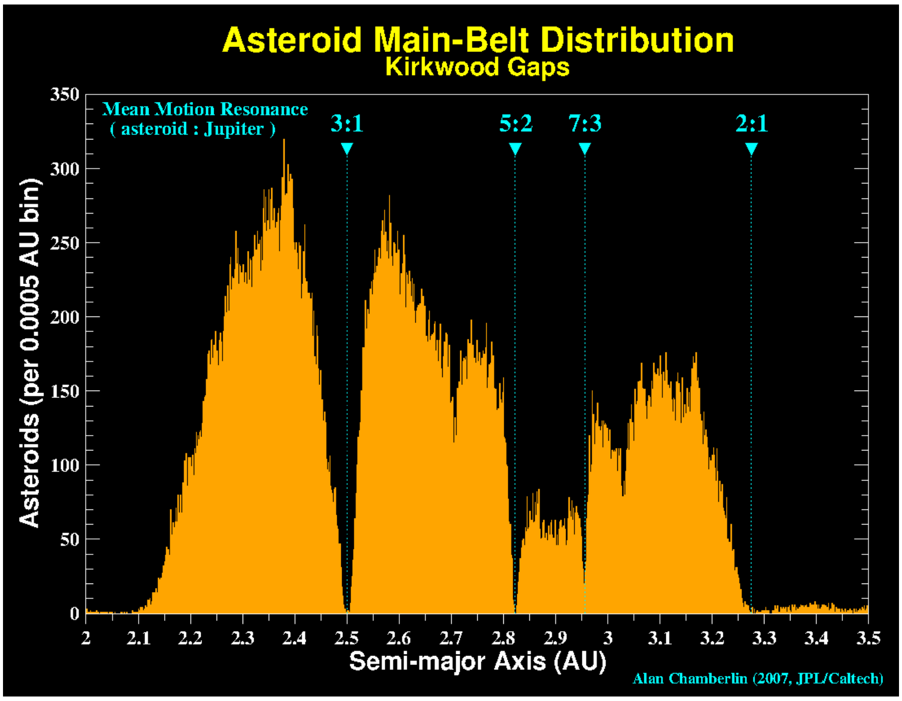

Figure 5.7: Number of asteroids with well-determined orbits with semi-major axis larger than $2UA$ and less than $3.5UA$ (source:NASA).

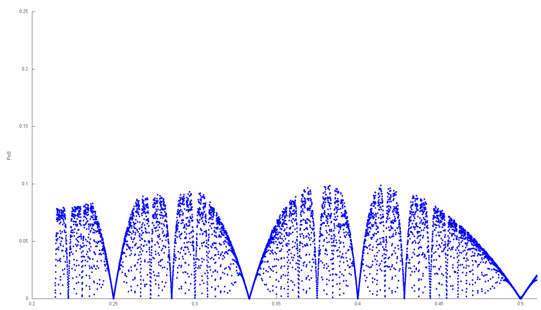

Figure 5.8: Power of Stability for $\tau = 1.1$, $N = 400$, $M = 10000$ and $m = 10000$.

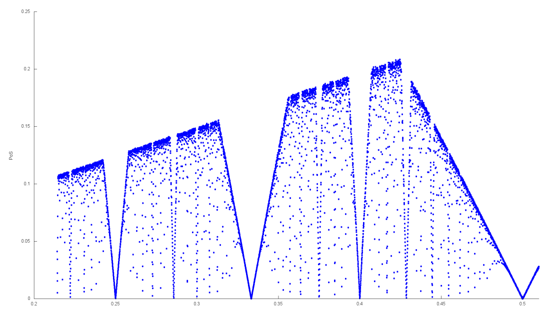

Figure 5.9: Power of Stability for $\tau = 1.5$, $N = 400$, $M = 10000$ and $m = 10000$.

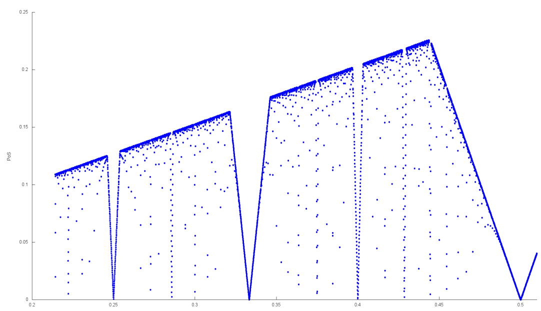

Figure 5.10: Power of Stability for $\tau = 2$, $N = 400$, $M = 10000$ and $m = 10000$.

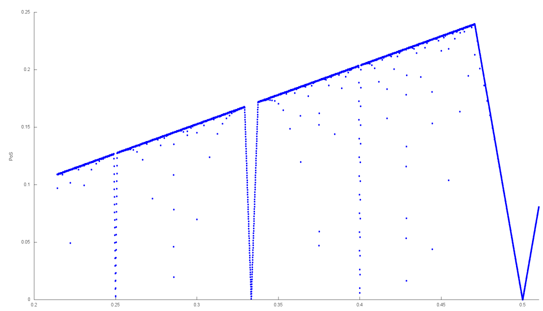

Figure 5.11: Power of Stability for $\tau = 3$, $N = 400$, $M = 10000$ and $m = 10000$.

Chapter 6

Computation and discussions

In this chapter, we apply the theorem of chapter 3 to the perturbation with the estimate of chapter 4. Because of the definition of the different variables and the complexity of the assumptions, we use a computer to verify them. However, since the computer has a finite precision, it can be necessary at some points to halve or double a constant because of the loss of precision we might have suffered. Let us give a short formulation of the result.

Theorem 6.1. *In the plane planetary three-body problem, if $m_1 \sim m_2 \sim 10^{-85}m_0$, there exists quasi-periodic motions, depending on three frequencies in the rotating reference frame, that is close to Keplerian motion.*

The constraints we require will be quantified more precisely along the computation. If KAM theorem can always be applied when decreasing the size of the perturbation P, the scheme we suggested relies a lot on the different steps we did, and on the norm of P as well. When releasing some constraints, as for instance letting m_2/m_1 go to 0, KAM theorem cannot be applied automatically anymore, and it needs to be discussed. Besides, when checking the hypotheses, we need to fix some values for the analyticity widths, and these values have a great impact on the possibility of applying the KAM theorem. Indeed, there is a competition between the size of the perturbation, depending on a lot of constants, and the size of ϵ for which the KAM theorem is valid. For instance, when letting the masses of the planets go to $10^{-200}m_0$, the initial choice of width that was working in the first case does not work anymore, one has to change the value of the analyticity width r' in the direction of the actions to make it work again.
For these reasons, we will be interested in studying only one system, with a fixed initial geometry, with fixed masses, and we will perform a change of variable such that the Hamiltonian is integrable. Close to these initial geometric values of the system, the KAM theorem applies.

6.1 Computation of the KAM theorem

We give here the different initial values to make the theorem work for a ratio of masses 10^{-85}. We give as well all the information that is necessary in order to understand the computation, as the analyticity widths, the size of the perturbation, the different necessary and sufficient conditions of KAM theorem, etc.

6.1.1 Initial Conditions

First, let us consider the geometry of the system. Regarding the semi-major axes, we consider the first planet to have a semi-major axis close to Jupiter's one, and the second planet will be considered a lot further:

$$\begin{aligned} a_1 &= 5.2\text{UA}, \\ a_2 &= 5.2 \times 10^{12}\text{UA}, \\ &\text{where } 1\text{UA} = 149597870700 \text{ m}. \end{aligned}$$

We will discuss the eccentricities of the system later. The masses are:

$$\begin{aligned} m_0 &= 2 \times 10^{30} \text{ kg}, \\ m_1 &= m_2 = 10^{-85} m_0, \\ G_{grav} &= 6.67408 \times 10^{-11}. \end{aligned}$$

We can then compute $M_1, M_2, \mu_1, \mu_2, \sigma_0, \sigma_1$, as well as H_{Kep}. First, we have $\Lambda_{0,1} = 2.04 \times 10^{-39} = 10^{-6}\Lambda_{0,2}$.
Considering the analyticity widths, we choose:

$$\begin{aligned} r_1 &= 1.35 \times 10^{-30}\Lambda_1 \sim 2.75 \times 10^{-69}, \\ \rho_{temp} &= \sqrt{2(\Lambda_1 - r_1)} \times 10^{-46} \sim 6.38 \times 10^{-66}, \\ r' &= 3 \times 10^{-219}. \end{aligned}$$

Finding values that work depends mostly on studying the ratio between the different terms of the perturbation and the condition of the KAM theorem.
Considering the analyticity width in the angles, we fix an initial value $t = 2$. With this value, we will be able to determine an analyticity width λ' using the implicit equation we derived when studying the complex Kepler's equation. As for now, only observe that we have $t \geq \lambda'$.
To define ρ', one had to take care of the change of variables from (ξ_i, η_i) to Cartesian coordinates. In this aim, define $m = \rho_{temp}^2/4$ and $M = 2m$. Then, using lemma 4.3, we define the value of ρ'_{temp} by the formula

$$\begin{aligned} \rho'_{temp} &= \max\left(\frac{r}{2\sqrt{2m}}\cosh s + \sqrt{2m}(\cosh s - 1), \frac{r}{2\sqrt{2M}}\cosh s + \sqrt{2M}(\cosh s - 1)\right), \\ \rho'_{temp} &\sim 2.50 \times 10^{-65}, \end{aligned}$$

using $s = t \geq \lambda'$. Now we can define

$$\begin{aligned} \rho_0 &\sim 2.55 \times 10^{-63}, \\ \rho'_0 &\sim 1.56 \times 10^{-63}, \\ \mu &\sim 3.83 \times 10^{-21} < \sqrt{\Lambda_1} \sim 4.51 \times 10^{-20}. \end{aligned}$$

Observe that there is only an order of magnitude between μ and $\sqrt{\Lambda_1}$. However, with the estimates on the perturbation made in chapter 1, we are far enough from a singularity of the perturbation and this value is not alarming.
This choice of m and M leads to values of real eccentricities verifying:

$$\begin{aligned} e_1 &\leq 1.42 \times 10^{-46}, \\ e_2 &\leq 1.42 \times 10^{-49}. \end{aligned}$$

As expected, they are very small. We can also find a value of λ' using the different definitions we have done. We have:

$$\begin{aligned} \lambda' &\sim 1.35234, \\ t_1 &= 2, \\ t_2 &\sim 1.35234, \end{aligned}$$

where t_1 and t_2 are the width in eccentric longitude.

6.1.2 Secular Hamiltonian and set of frequencies

In this part, we compute the initial secular Hamiltonian, its frequencies, and the analyticity width h' in frequencies.
The Kepler Hamiltonian is straightforward to compute. Regarding the secular part of the perturbation, we chose a very small ratio for the semi-major axes: 10^{-12}. Computing precisely the secular Hamiltonian has two main goals: finding the frequencies of the system, and determining h'. However, we do not compute every term of the secular Hamiltonian, but consider only the first term in the development in semi-major axes. This approximation relies on the fact that the ratio of the semi-major axes is chosen very small, and hence will have a very small impact on the determination of the frequencies (plus the fact that the precision of the machine is not infinite). An approximation could still change in a noticeable way the value of the analyticity width. Nevertheless, if the approximation is very small, we will divide or multiply by 2 the results each time we feel that the approximation would have possibly had any noticeable impact.
We compute first the coefficients $d_{i,j,k}$ of appendix A.2.1 to the first order. There is an order of magnitude of 24 when comparing the first term and the remainder of the expansion in power of the ratio of the semi-major axes (coming from the $(a_1/a_2)^2$ factor between consecutive terms). The value of the linear terms in I_3 and I_4 are:

$$\begin{aligned} d_{100} &\sim 1.26 \times 10^{-129}, \\ d_{010} &\sim 1.26 \times 10^{-135}, \\ d_{001} &\sim -3.16 \times 10^{-144}. \end{aligned}$$

Consider now the rotation variable υ: by definition, for $x = \frac{d_{100}-d_{010}}{d_{001}}$, it is of the form

$$\upsilon = -x + \sqrt{1+x^2} \sim -\frac{1}{x} \quad [\ |x| \to +\infty],$$

and therefore

$$\upsilon \sim -1/x = \frac{-d_{001}}{d_{100} - d_{010}} \sim 2.5 \times 10^{-15}.$$

Hence, we will do a further approximation to derive the terms $d'_{i,j,k}$. We will consider $\upsilon = 0$, and they will be equal to the terms $d_{i,j,k}$. This approximation relies again on the initial values we have chosen, and the change of frequencies would have been of order 10^{-20} times the frequencies without it.
With the coefficients $d'_{i,j,k}$, we can now derive the linear Hamiltonian $H_{0,1}$ explicitly, as well as the frequencies. It then remains to determine the analyticity width in frequencies. In the actual computation, one needs first to estimate the value of the transitional Hamiltonian that is required to compute $H_{0,2}$ and $H_{0,3}$. We will keep this calculation for later, though it is true that the terms involved in $H_{0,3} - H_{0,1}$ are again very small compared to the terms in $H_{0,1}$ (using Cauchy's inequality shows that the order of magnitude is more than 50). Proceeding as in the end of chapter 4, we can determine an approached value of h' by computing the largest term of the adjugate matrix, and considering the dominant matrix. In this computation, we do not hesitate to double or halve the results that appear, in order to remove the impact of the approximations. We can choose $h \sim 2.2 \times 10^{-173}$.

6.1.3 Size of the perturbation and KAM conditions

Using all the work we have done in chapter 1, we are able to determine the size of the perturbation in the set $\mathcal{D}_{\Lambda_0,r,\rho_0,\rho'_0,\lambda'}$:

$$\|H_{pert}\|_{\mathcal{D}_{\Lambda_0,r,\rho_0,\rho'_0,\lambda'}} \leq 4.26 \times 10^{-163}$$

Applying corollary 2.16, we get an estimate on the new perturbation, and we iterate it 2 more times. This is the first time we need to consider the Diophantine constant. When applying this corollary, we are in dimension $2n$ with $n = 2$. Hence the frequency vector is a vector (ω_1, ω_2). We choose it to be optimal in γ for $\tau = n$, where the optimality is define in 5.21. When using the computer, the numerical precision implies that we cannot choose our vector to be Diophantine. Though, in a small neighborhood of our initial values, where the application of the KAM theorem still works, such a vector exists. The ratio between these frequencies is 10^{-18}, and hence, by changing very slightly the initial values of the semi-major axes, one can be as close as wanted to this point. Hence, we choose $\gamma_2 = \min(\omega_1, \omega_2)/2 \sim 8.4 \times 10^{-27}$, the factor $1/2$ being there to absorb the approximations of the frequencies. γ_2 will be used only for these steps in dimension 4 ($n = 2$).
The size of the different transformations are:

$$\begin{aligned} \epsilon_2 &\leq 2.62 \times 10^{-200}, \\ \epsilon_3 &\leq 6.52 \times 10^{-270}, \\ \epsilon_4 &\leq 5.39 \times 10^{-406}. \end{aligned}$$

Now, applying the different proposition 4.8,4.9 and 4.10, we obtain respectively:

1. Remainder of the perturbation after 3 steps of corollary 2.16: $\epsilon_4 \leq 5.39 \times 10^{-406}$,
2. Remainder of the BNF: $\leq 3.49 \times 10^{-406}$,
3. Remainder of the part of the transitional Hamiltonian depending on the eccentricities: $\leq 1.24 \times 10^{-436}$,
4. Non-linear remainder: $\leq 2.23 \times 10^{-406}$.

Hence, we can let $\epsilon = 2 \times 10^{-405}$, and we have

$$\|H_{pert}\|_{r'/32, \lambda'/16} < \epsilon.$$

Let us now see the different conditions of the KAM theorem.
We will try to apply it for $\delta = 5 \times 10^6$, trying to make the three conditions of the KAM theorem of the same order. We have, for this value, $K = 953$. Observe besides that the first and the third terms appearing in the minimum of theorem 3.3 are linear with the Diophantine constant γ. Let us consider first that this constant is optimal in γ with $\tau = n = 4$. It corresponds to saying that $\gamma_4 = \min(\omega_1, \omega_2, \omega_3, \omega_4)/2$, and hence, we let $\gamma_4 \sim 6.31 \times 10^{-136}$.
The three minima have values:

$$\frac{\gamma_4 r' \sigma^\nu}{4^\nu C_0} \sim 6.44 \times 10^{-379}, \qquad \frac{h' r'}{2\delta} \sim 6.49 \times 10^{-399}, \qquad \frac{\gamma_4 r'}{2K^\nu \delta} \sim 2.41 \times 10^{-376}.$$

The first observation is that with the size of the perturbation we derived, we can apply the KAM theorem. The second observation, is that we do not need the Diophantine condition to be optimal in γ_4. Indeed, we can let $\gamma_4' = \gamma_4 \times 10^{-20}$, and then we have:

$$\frac{\gamma_4' r' \sigma^\nu}{4^\nu C_0} \sim 6.44 \times 10^{-399}, \qquad \frac{h' r'}{2\delta} \sim 6.49 \times 10^{-399}, \qquad \frac{\gamma_4' r'}{2K^\nu \delta} \sim 2.41 \times 10^{-396}.$$

The estimate on the norm of the perturbation not changing (since γ_2 is fixed), we can consider this value to be our final value for the Diophantine condition. We choose to let γ_4' have a very small value compared to γ_4, which was the optimal Diophantine constant γ, and hence to possibly work on a larger set of Diophantine vectors.
Note that we still have some "room" between the estimate of ϵ and the minimum of the previous values. The order of magnitude is 5, and therefore, it is clearly possible to change slightly the values of Λ_0 or of the m_i.

6.1.4 Size of the transformation

With the work done in chapter 3, we can derive some estimates on the size of the transformation that we just applied. Though, we are first interested on the new frequencies ω' of the system. We want to quantify the difference between the frequencies ω' corresponding to the quasi-periodic motion in the new variables, and the frequencies ω of the perturbed Hamiltonian. We will consider here that the initial frequencies are given by $\omega = H'_{0,1}(\Lambda_0, I_3, I_4)$, where $m < I_3, I_4 < M$. This frequency vector is entirely known, since we know explicitly $H_{0,1}$.
Along the scheme, we change three times of frequencies. The first and second times occur when "pushing" the perturbation further, or more precisely when we add $\bar{P}_2$ and $\bar{P}_3$ to $H_{0,1}$. Recall that we only add the part of these Hamiltonian that do not depend on the action I_3 and I_4. The third part comes directly from the KAM theorem, and a bound on its estimate is given in theorem 3.3.
Observe that the parts coming from $H_{0,3} - H_{0,1}$ are independent of the third and fourth actions. A simple Cauchy estimates gives a modification in the frequencies of order $\sim 10^{-93}$, which is very small compared to the values of ω_1 and ω_2, the latter being $\sim 1.68 \times 10^{-26}$.
The estimate in the KAM theorem for the change in frequencies $\varphi - Id$ on the set $\{\theta : |\Im\theta| < \lambda'/2\}$ is given by:

$$\|v(\omega)\| = \|\varphi - Id\| \leq 1.6 \times 10^{-5} h' \sim 1.39 \times 10^{-177}.$$

Hence, the frequencies ω' of the system in the new set of variables are:

$$\begin{cases} \omega'_1 = \omega_1 \pm 10^{-92} \sim 1.68 \times 10^{-8} \pm 10^{-132} \\ \omega'_2 = \omega_2 \pm 10^{-92} \sim 1.68 \times 10^{-26} \pm 10^{-132} \\ \omega'_3 = \omega_3 \pm 10^{-176} \sim 1.26 \times 10^{-129} \pm 10^{-180} \\ \omega'_4 = \omega_4 \pm 10^{-176} \sim 1.26 \times 10^{-135} \pm 10^{-180} \end{cases}$$

The change in frequencies is therefore very small. We can as well give the computation of the other norms appearing in the KAM theorem, so as to have an idea of the size of the transformations happening in the KAM theorem:

$$\begin{aligned} &\|\varphi - Id\|_L \leq 2.90 \times 10^9, \\ &\|W\Phi - \Phi_0\| \leq 4.48 \times 10^{-23}, \\ &\|W\Phi - \Phi_0\|_L \leq 1.16 \times 10^{167}. \end{aligned}$$

6.2 Discussion and improvements

In this section, we discuss the result and its dependency in the initial choices. The choices we made when constructing the final Hamiltonian have a great impact on the possibility of application of the KAM theorem. We choose to debate these choices, in order to make some points on possible improvements.

6.2.1 Dependency in the initial parameters

The initial parameters we choose suit the choices made in the computation. Though we can wonder what happens when trying to change them in a way or another, since all the parameters, the conditions on KAM theorem and the size of the different parts of the perturbation are intricately linked.

Decreasing the mass of the planets: When decreasing the mass of the planets, the size of the perturbation can become a problem for our construction. Indeed, the remainder of the BNF depends mostly on the size of the variable ρ_0 and μ, and does not decrease as fast as the other part of the perturbation. The constraints on the application of the KAM theorem on the other hand are getting more important. Without letting the eccentricities decrease, there exists some ratio between the mass of the planets and the sun at which we cannot expect the theorem to hold, using the construction we did. Hence, while letting this ratio decrease, one has to decrease the eccentricities as well so that we can still remove the perturbation.

Decreasing the mass m_2: When decreasing m_2, again the remainder of the BNF can cause a problem for our construction, and one needs to let ρ_{temp} (and therefore the eccentricities) decrease.

Decreasing the ratio of the semi-major axes: When decreasing a_1/a_2, the problem arises from the first transformation $\varphi^{\epsilon}_{X_1}$. Indeed, in this case, after applying this map, the remainder becomes larger than the previous perturbation: $||\tilde{H}_{pert} \circ \varphi^{\epsilon}_{X_1}|| \geq ||H_{pert}||$. The problem arises from the truncation of the Hamiltonian. When the two frequencies are too far, the value of the order of truncation K becomes close to 1: recall the inequality

$$K \leq \left(\frac{\gamma_2}{2r \left\| H''_{0,1} \right\|_{r,s}} \right)^{\frac{1}{3}},$$

the fact that γ_2 becomes very small then becomes a problem. When the frequencies are not close enough, one cannot expect in general to remove the same perturbation as when they are (we will discuss this problem in another paragraph). Hence, one has to make the ratio between the mass of the planets and the sun decrease at the same time.

Decreasing the eccentricities: When decreasing the eccentricities, we have to let as well the analyticity width r_1 decrease at the same time, in order to keep the variables I_3 and I_4 well-defined. Hence, the conditions on the KAM theorem become more restrictive. Though, the estimate ϵ_4 is independent of the analyticity width r_1 and do not decrease. Under some limit value, the conditions of the KAM theorem are not satisfied anymore because of the size of ϵ_4. It is nevertheless possible to deal with this problem, by iterating the corollary 2.16 a few more times, and hence to obtain a suitable value for the size of the perturbation.

Even though changing the initial conditions one by one makes it impossible to remove the perturbation with a change of variables, if we change the initial condition together, the result can still hold.
Observe as well that in a neighborhood of the initial values we considered, the result is true. Hence, it should be possible to change the initial frequencies in order to get close to a Diophantine vector verifying a sufficient condition to apply the scheme.

6.2.2 Possible improvements of the general results

We can now talk about the different improvements that can be made, in order to apply the theorem for a larger perturbation. Indeed, the number 10^{-85} for the ratio between the masses of the planets and the sun do not seem to be optimal at all. As well, the constraints on the semi-major axes and on the eccentricities need to be discussed. We will try to express the possible improvements in an intelligible way, although this discussion is made tougher because a lot of them are linked. Besides, there exists two different kinds of limits on the scheme we did: some are technical, as for example the value of the ratio of the semi-major axes, on which we rely to perform some approximations, and some are theoretical, and come from the theorems we are using.

Technical Improvements: Here, we talk mainly of the improvements that can be done with more computation in the formulas.

- **Limit on the semi-major axis:** The choice of ratio of the semi-major axes exists mainly for convenience. Indeed, using the formulas of the expansion of the secular Hamiltonian in power of the semi-major axes can become quickly cumbersome without any approximation. That is why we decided to approximate them by their first order. We also chose to let $\upsilon \sim 0$. These two approximations would not be possible in the case of close values of semi-major axes. Besides, the computation of the inverse map of the frequencies would require a lot more computation in this case. Yet, removing this assumption could lead to better results, while making calculation more difficult. Indeed, as we noticed in the previous discussion, having closer frequencies allows us to obtain a better result when applying corollary 2.16, as well as in the KAM theorem. If the ratio of the semi-major axes was close to one, we would have $\omega_1 \sim \omega_2$ and $\omega_3 \sim \omega_4$, which would lead to smaller values of ϵ_2, ϵ_3 and ϵ_4. However, the ratio ω_1/ω_3 would still be very large, and it should not be the first improvement one should try to perform.

- **Limit on the eccentricities:** The limit on the eccentricities comes mostly from the order of the BNF. We choose to put the secular Hamiltonian under BNF of order 4, and to consider every higher order as part of the perturbation. This choice obviously requires to have a very small eccentricity. It is possible to compute at a higher order, by iteration, the BNF of the secular Hamiltonian, which would lead to the possibility of considering much larger eccentricities. A possible way would be to compute each term of the perturbation at the wanted order, in order to reduce the analyticity loss while putting the perturbation under normal form. Besides, instead of considering a Diophantine condition, it could be more optimal to consider a non resonance condition of some order, so as to avoid the presence of a Diophantine constant that could be too small. Observe that in our scheme, we let the values ρ_0 and ρ_0' be of the same order. Indeed, with very small value of the eccentricities, it is not worthwhile to separate these values. However, when considering important eccentricities, it can become essential to consider a smaller value of ρ_0', so as to have r_1 as small as necessary. It remains important, while applying the BNF theorem, to keep track of the analyticity widths.

These remarks justify the fact that these restrictive constraints on the eccentricities and the ratio of the semi-major axes are not intrinsic in the plane planetary 3 body problem. With more computation, it is always possible to weaken those. The only limit that should exist, if we can compute the Hamiltonian at a very high order, is the limit of convergence of the series which bound the perturbation, which happens for high eccentricities and small ratio of semi-major axes.

Theoretical Improvements: We consider here the improvements that can be done in the different theorems, and in the general scheme of the result. We try to list the possible improvements by order of importance.

- **Several KAM steps by hand for the final Hamiltonian:** The greatest improvement is probably the possibility of doing one or several steps of corollary 2.16 with the final Hamiltonian, in dimension 4. Indeed, the conditions required by the KAM theorem are that the size of the perturbation is less than $\gamma_4 r \sigma^\nu / c$, where $c \sim 10^{18}$. In the corollary, this constant is $c \sim 10^8$ in dimension 4. Hence, it would be quite straightforward to gain a factor 10^{10} for the ratio of the masses.

- **A more suitable KAM theorem:** Our choice of KAM theorem was done by convenience. However, it asks the unperturbed Hamiltonian to be linear in the action. More precisely, it requires to add the non-linear terms to the perturbation. This requirement is, with the initial values, the most restrictive one concerning the value of r'. This kind of theorem is hence not very suitable to the plane planetary three-body problem. Indeed, the fact that the initial Hamiltonian is degenerate, and that we use a part of the small perturbation to remove the degeneracy, makes this theorem difficult to apply. Loosely speaking, the initial Hamiltonian is $H_{Kep}+\epsilon H_{pert}$, and after some transformation, we look at

$$H_{Kep} + \epsilon \bar{H}_{pert} + \epsilon^2 H_2.$$

Asking that $(H_{Kep} + \epsilon \bar{H}_{pert})^2 \cdot r'^2$ is of the size of $\epsilon^2 H_2$ requires that r' is of the size of the perturbation, and hence no more related to the actual size of Λ_1 or the eccentricities. This implies a huge and artificial loss of analyticity.
Quantitatively, the value of r' has two origins, the initial analyticity width related to the Λ_i, and the initial analyticity width coming from the value of ρ and ρ' (which are themselves given by to the size of the perturbation). In our case, the initial analyticity width related to the semi-major axes was $r_1 \sim 10^{-69}$). To make the non-linear part of the unperturbed Hamiltonian small enough, we have to choose $r' = 3 \times 10^{-219}$, which is a lot smaller than the first values of r_1. Yet, it is difficult to give an order of how much we could gain with a theorem not asking that the non-linear part must be put in the perturbation. Indeed, the numerous transformations we did make it difficult to easily track what would change. It might need some more transformation of the type given by corollary 2.16, but as well, one could consider larger eccentricities, which would lead to a possibly even better value of r'...A new theorem might have other requirements as well, and even if it is a good lead for a great improvement, we cannot pronounce ourselves on its order.

- **More generally, a KAM theorem adapted to low and high frequencies:** When performing a KAM step, we use the theorem of Rüssmann 2.10, that is optimal for a generic Hamiltonian. Though, as we previously observed, our Hamiltonian is specific in the way that its frequencies are not homogeneous, we have $\omega_1, \omega_2 \gg \omega_3, \omega_4$. And even more, in the case of a large ratio of Λ_1 and Λ_2: we have $\omega_1 \gg \omega_2$ and $\omega_3 \gg \omega_4$. In this case, for a small value of K, and a vector $k = (k_1, k_2, k_3, k_4) \in \mathbb{Z}^4$ such that $|k|_1 < K$, we have $\omega \cdot k \sim k_1\omega_1 \gg \gamma_4/|k|_1^\tau$. Hence, it is possible to improve the estimate of Rüssmann, by considering that the analyticity loss only takes place for high values of K, when the Fourier coefficient are getting extremely small. Such a theorem would require some precise study and estimates on the ratio between the frequencies, and also a more precise Diophantine condition. To a certain extent, we did this when we used Rüssmann theorem in the transformation $\varphi^\epsilon_{X_1}$, considering that our vector (ω_1, ω_2) satisfied an optimal Diophantine condition, and we applied Rüssmann theorem to the wanted Hamiltonian. One could precisely analyze the transformation for small K, and then, when K becomes large enough and that the $k_i\omega_i$ starts to be of the same order, estimate as usual the size of the new function. With such a theorem, one could improve the estimates on the size of the transformation we need to use, and therefore on the constants appearing in the KAM step (and even in the order of the exponents). However, it would be necessary to have a more specific Diophantine condition: for instance, we could ask for a frequency vector in dimension n to verify a Diophantine condition for $\tau = n$ and some constant γ_n, and then for each of its sub-vector of dimension $n' < n$ to verify a Diophantine condition for $\tau = n'$ and some constant $\gamma_{n,i}$ (the i referring to the precise sub-vector). Such a condition would allow a precise study of the first terms of the Fourier series of the Hamiltonian generating the transformation.

- **More than 3 steps of corollary 2.16:** For several reasons, this could improve the general result. Indeed, ϵ_4 was close to ϵ in the application of the KAM theorem. It is responsible for the fact that we cannot let, as an example, the eccentricities go to zero. If we judge that the loss of analyticity width, while performing another iteration, is small, then one could iterate the corollary several more times. The fact that we have a large Kepler problem Hamiltonian implies that this transformation works well at first (if the frequencies ω_1 and ω_2 are of the same order). One could then increase the ratio of the masses, and consider more terms in the perturbation. It implies a loss in the analyticity widths r_1 and $s_1 = \lambda'$, as well as a loss in the analyticity width in the frequencies.

- **BNF for the transitional Hamiltonian:** The transitional P_2, P_3 Hamiltonian appearing when performing a KAM step to the initial perturbation depends on the angle $g = g_1 - g_2$. Hence we chose to consider that the part depending on this angle is part of the perturbation. In the case we want to consider a larger perturbation, and that we need to apply more KAM steps, it is possible that the parts depending on the angle g are not small enough. In this case, it is possible to put the Hamiltonian $H_{0,i} + \bar{P}_{i+1}$ under normal form, so as to remove the dependency in the angle g of the unperturbed Hamiltonian $H_{0,i+1}$. A normal form up to some order n would lead to an estimate of these terms of the size $P_i(\rho_0 + \rho_0')^n/\mu^n$.

- **Computing the first Legendre polynomials:** We mentioned before some possible improvements on the estimate of the perturbation. The largest improvement that could be done was to calculate precisely the first terms of the Legendre polynomials, in order to have a better estimate when computing its analytic continuation. One can expect to gain a factor close to 3 by doing such a computation.

- **General optimization of the theorems:** Broadly speaking, we can say, without taking any risk, that trying further to improve any constant in the theorems we used could lead to a gain in the result. Improving any constant, trying to be as optimal as we can in every possible steps is an effort that could be rewarded by some orders of magnitude in the ratio of mass we are considering, and therefore will not be a vain effort.

This elements terminates the discussion about the possible improvements that can be done to improve the result. If some can be done quite straightforwardly, some requires an important labour to be set up.

Appendix A

Formulas of the Secular Hamiltonian

A.1 First terms of the series expansion in power of eccentricities

A.1.1 Before Integration

The formulas for the coefficients $b_{i,j,n}$ for $i+j \leq 4$ are trigonometric polynomials:

$$
\left\{
\begin{aligned}
b_{0,0,n} &= \frac{1}{4\pi^2}\int_{\mathbb{T}^2} P_n(\cos(v_2-v_1+g))dv_1dv_2 \\
b_{2,0,n} &= -\frac{1}{4\pi^2}\int_{\mathbb{T}^2}\left(n+\frac{3}{2}-(n+2)(n+3)\cos^2 v_1\right)P_n(\cos(v_2-v_1+g))dv_1dv_2 \\
b_{0,2,n} &= b_{2,0,n} \\
b_{1,1,n} &= -\frac{1}{4\pi^2}\int_{\mathbb{T}^2}(n+2)(n-1))\cos v_1\cos v_2 P_n(\cos(v_2-v_1+g))dv_1dv_2 \\
b_{4,0,n} &= \frac{1}{4\pi^2}\int_{\mathbb{T}^2}\left(\frac{(2n+1)(2n+3)}{8}-\frac{(n+2)(n+3)(2n+3)\cos^2 v_1}{4}+\right. \\
&\qquad \left.\frac{(n+2)(n+3)(n+4)(n+5)}{24}\cos^4 v_1\right)P_n(\cos(v_2-v_1+g))dv_1dv_2 \\
b_{0,4,n} &= \frac{1}{4\pi^2}\int_{\mathbb{T}^2}\left(\frac{(2n+1)(2n-1)}{8}+\frac{(n-2)(n-1)(2n-1)\cos^2 v_2}{4}+\right. \\
&\qquad \left.\frac{(n-1)(n-2)(n-3)(n-4)}{24}\cos^4 v_2\right)P_n(\cos(v_2-v_1+g))dv_1dv_2 \\
b_{2,2,n} &= b_{2,2,0,n}+b_{2,2,2,n}\cos^2 g = \frac{1}{4\pi^2}\int_{\mathbb{T}^2}\frac{1}{16}\left(n(n+1)+(n+2)(n+3)\cos(2v_1)\right)\times \\
&\qquad \left(n(n+1)+(n-2)(n-1)\cos(2v_2)\right)P_n(\cos(v_2-v_1+g))dv_1dv_2 \\
b_{3,1,n} &= \frac{1}{4\pi^2}\int_{\mathbb{T}^2}\left((n+2)(n+\frac{3}{2})\cos v_1-\frac{(n+2)(n+3)(n+4)}{6}\cos^3 v_1\right)\times \\
&\qquad (n-1)\cos v_2 P_n(\cos(v_2-v_1+g))dv_1dv_2 \\
b_{1,3,n} &= -\frac{1}{4\pi^2}\int_{\mathbb{T}^2}\left(n(n+7)\cos v_2-(n-2)(n-3)\cos^3 v_2\right)\times \\
&\qquad \frac{(n-1)(n+2)}{12}\cos v_1 P_n(\cos(v_2-v_1+g))dv_1dv_2
\end{aligned}
\right.
$$

A.1.2 After Integration

Call $L_p = \frac{1}{2^{4p}}\binom{2p}{p}^2$, after integration, for $p \geq 1$, these coefficients are:

$$\begin{cases} b_{0,0,2p} &= L_p, \\ b_{2,0,2p} &= \frac{p(2p+1)}{2}L_p, \\ b_{0,2,2p} &= \frac{p(2p+1)}{2}L_p, \\ b_{1,1,2p+1} &= -\frac{(2p+1)(2p+3)}{2}\frac{p}{p+1}L_p, \\ b_{4,0,2p} &= \frac{(p-1)p(2p-1)(2p+1)}{16}L_p, \\ b_{0,4,2p} &= \frac{p(p+1)(2p+1)(2p+3)}{16}L_p, \\ b_{2,2,0,2p} &= \frac{p(2p+1)(2p^2+p+3)}{8}L_p, \\ b_{2,2,2,2p} &= \frac{(p-1)p(2p+1)(2p+3)}{4}L_p, \\ b_{3,1,2p+1} &= -\frac{p(2p+1)^2(2p+3)}{8}\frac{p}{p+1}L_p, \\ b_{1,3,2p+1} &= -\frac{p(2p+1)(2p+3)^2}{8}L_p. \end{cases}$$

A.2 First terms of the series expansion in Poincaré coordinates

A.2.1 Full formulas of the coefficients

Calling

$$\mathcal{M}_n = G_{grav}^2 \sigma_n \frac{(m_0+m_1)^{3n+1}}{(m_0+m_1+m_2)^{n+1}} \frac{m_2^{2n+3}}{(m_0 m_1)^{2n-1}},$$

$$c_{i,j} = \sum_{n\geq 2} \mathcal{M}_n \frac{\Lambda_1^{2n}}{\Lambda_2^{2(n+1)}} \times b_{i,j,n},$$

the average of the Hamiltonian over the fast angle is:

$$\begin{aligned}\frac{1}{4\pi^2}\int_{\mathbb{T}^2} H_{pert}(\Lambda_1, \Lambda_2, &\lambda_1, \lambda_2, e_1, e_2, g_1, g_2) d\lambda_1 d\lambda_2 = \\ d_{0,0,0} + &d_{1,0,0}\Gamma_1 + d_{0,1,0}\Gamma_2 + d_{0,0,1}\Gamma_0 + \\ &d_{2,0,0}\Gamma_1^2 + d_{0,2,0}\Gamma_2^2 + d_{0,0,2}\Gamma_0^2 + \\ &d_{1,1,0}\Gamma_1\Gamma_2 + d_{1,0,1}\Gamma_1\Gamma_0 + d_{0,1,1}\Gamma_2\Gamma_0 + o((\Gamma)^2),\end{aligned}$$

with

$$\begin{cases} d_{0,0,0} = c_{0,0}, & & \\ d_{1,0,0} = \dfrac{2}{\Lambda_1} c_{2,0}, & d_{0,1,0} = \dfrac{2}{\Lambda_2} c_{0,2}, & d_{0,0,1} = \dfrac{2}{\sqrt{\Lambda_1\Lambda_2}} c_{1,1}, \\ d_{2,0,0} = \dfrac{1}{\Lambda_1^2}(4c_{4,0} - c_{2,0}), & d_{0,2,0} = \dfrac{1}{\Lambda_2^2}(4c_{0,4} - c_{0,2}), & d_{0,0,2} = \dfrac{4}{\Lambda_1\Lambda_2} c_{2,2,2}, \\ d_{1,1,0} = \dfrac{4}{\Lambda_1\Lambda_2} c_{2,2,0}, & d_{1,0,1} = \dfrac{1}{\Lambda_1^{\frac{3}{2}}\Lambda_2^{\frac{1}{2}}}(4c_{3,1} - c_{1,1}), & d_{0,1,1} = \dfrac{1}{\Lambda_1^{\frac{1}{2}}\Lambda_2^{\frac{3}{2}}}(4c_{1,3} - c_{1,1}). \end{cases}$$

A.2.2 Expansion at the second order in Λ_1/Λ_2

The expansion of the coefficient $d_{i,j}$ at the second order in the ratio of the Λ_i can be written as follows (we forget to write $o\left(\left(\frac{\Lambda_1}{\Lambda_2}\right)^8\right)$ for convenience):

$$
\begin{cases}
d_{0,0,0} = +\dfrac{1}{64\Lambda_2^2}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4 \left(16\mathcal{M}_2 + 9\ \mathcal{M}_4\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) \\
d_{1,0,0} = +\dfrac{1}{32\Lambda_2^3}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^3 \left(24\mathcal{M}_2 + 45\ \mathcal{M}_4\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) \\
d_{0,1,0} = +\dfrac{1}{32\Lambda_2^3}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4 \left(24\mathcal{M}_2 + 45\ \mathcal{M}_4\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) = \dfrac{\Lambda_1}{\Lambda_2}d_{1,0,0} \\
d_{0,0,1} = -\dfrac{1}{32\Lambda_2^3}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^{11/2} \left(60\mathcal{M}_3 + 105\mathcal{M}_5\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) \\
d_{2,0,0} = -\dfrac{3}{128\Lambda_2^4}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^2 \left(16\mathcal{M}_2 - 15\ \mathcal{M}_4\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) \\
d_{0,2,0} = +\dfrac{3}{128\Lambda_2^4}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4 \left(64\mathcal{M}_2 + 285\mathcal{M}_4\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) \\
d_{0,0,2} = +\dfrac{315}{128\Lambda_2^4}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^7 \left(4\ \mathcal{M}_4 + 15\ \mathcal{M}_6\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) \\
d_{1,1,0} = +\dfrac{9}{64\Lambda_2^4}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^3 \left(16\mathcal{M}_2 + 65\ \mathcal{M}_4\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) \\
d_{1,0,1} = -\dfrac{15}{64\Lambda_2^4}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^{9/2} \left(8\ \mathcal{M}_3 + 63\ \mathcal{M}_5\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right) \\
d_{0,1,1} = -\dfrac{15}{16\Lambda_2^4}\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^{11/2} \left(9\ \mathcal{M}_3 + 35\ \mathcal{M}_5\left(\dfrac{\Lambda_1}{\Lambda_2}\right)^4\right)
\end{cases}
$$

Appendix B

Analytical tools

B.1 Remainder of the truncated Fourier series

Let $\mathcal{A}_s$ be the set of functions defined on $\mathbb{T}^n$ that are bounded and analytic on the set $\mathbf{T}^n_s = \{\theta \in \mathbb{T}^n_{\mathbb{C}}, |\Im\theta| < s\}$.
Let $f \in \mathcal{A}_s$, for $\theta \in \mathbf{T}^n_s$, we can write

$$f(\theta) = \sum_{k \in \mathbb{Z}^n} f_k e^{\imath k \cdot \theta}.$$

For all $k \in \mathbb{Z}^n$, we have $|f_k| \leq |f|_s e^{-|k|s}$. Indeed, this result is straightforward using the fact that f is 2π-periodic in each variable, and analytic and bounded on its set of definition.
Let us consider the truncation of order $K \in \mathbb{N}$ of f:

$$T_K f = \sum_{k \leq K} f_k e^{\imath k \cdot \theta}.$$

Lemma B.1. *Let $s > 0$ and $\sigma < s$. If $f \in \mathcal{A}_s$, and $K\sigma \geq n-1$ then*

$$|f - T_K f|_{s-\sigma} \leq 4^n n! K^n e^{-K\sigma} |f|_s, \qquad 0 \leq \sigma \leq s$$

Proof. We have:

$$\begin{aligned} |f - T_K f|_{s-\sigma} &\leq \sum_{k \in \mathbb{Z}^n, |k|_1 > K} |f_k| \exp(|k|_1(s-\sigma)) \\ &\leq |f|_s \sum_{k \in \mathbb{Z}^n, |k|_1 > K} \exp(-|k|\sigma) \\ &\leq 4^n |f|_s \sum_{l \in \mathbb{N}, l > K} l^{n-1} \exp(-l\sigma), \end{aligned}$$

where we used the fact that the number of $k \in \mathbb{Z}^n$ such that $|k|_1 = l$ is less than $4^n l^{n-1}$. As for the last sum, since the general term is strictly decreasing, it can be bounded by the incomplete gamma function:

$$\begin{aligned}
\sum_{l\in\mathbb{N},l>K} l^{n-1}\exp(-l\sigma) &\leq \int_K^{\infty} x^{n-1}\exp(-x\sigma)dx \\
&\leq \frac{1}{\sigma^n}\int_{K\sigma}^{\infty} x^{n-1}\exp(-x)dx \\
&\leq \frac{(n-1)!}{\sigma^n}\exp(-K\sigma)\sum_{k=0}^{n-1}\frac{(K\sigma)^k}{k!} \\
&\leq \frac{(n-1)!}{\sigma^n}\exp(-K\sigma)n\times(K\sigma)^{n-1} \\
&\leq n!\frac{K^{n-1}}{\sigma}\exp(-K\sigma) \\
&\leq n!K^n\exp(-K\sigma).
\end{aligned}$$

Injecting this result in the previous inequation, the lemma is proved. □

B.2 Inversion of analytic map close to the identity

Recall that we defined the set O_h as the open complex neighborhood of radius h of the subset of frequencies $\Omega_{\gamma,\tau}$ for some $\gamma > 0$. Using as usual the sup-norm for maps and vectors, we will prove the following lemma for the inversion of the frequency vector:

Lemma B.2. *Let $f : O_h \subset \mathbb{C}^n \to \mathbb{C}^n$ be analytic, such that $|f - Id| \leq \delta \leq h/4$ on the set O_h. Then f has an analytic inverse g on $O_{h/4}$, and it satisfies:*

$$|g - Id|_{h/4},\ \frac{h}{4}|g' - Id| \leq \delta$$

Proof. Call $v = f - Id :\ O_h \to \mathbb{C}^n$.
Let $x \neq y$ be two points of $O_{h/2}$ such that $f(x) = f(y)$. We have

$$\begin{aligned}
|x-y| = |v(x) - v(y)| \leq |v'|_{h/2}|x-y| &\leq \frac{|v|_h}{h/2}|x-y| \\
&< |x-y|.
\end{aligned}$$

Hence, the functions v and f are one-to-one.
Moreover, for any given $y \in O_{h/4}$, the function

$$\begin{aligned}
\varphi : O_{h/2} &\to O_{h/2} \\
x &\mapsto y - v(x)
\end{aligned}$$

is a contraction. Hence, it has a unique fixed point. We can define for any point $y \in O_{h/4}$ such a pre-image by the function f, and this pre-image depends analytically on y.
The function f is therefore a biholomorphism from the set $O_{h/2}$ into a subset of O_h. Define the function $g = f^{-1}$ on the set $O_{h/4}$ onto the set $O_{h/2}$. Let $y \in O_{h/4}$, we have:

$$|g(y) - y| = |v(g(y))|,$$

hence the estimate on $|g - Id|_{h/4}$. As for the derivative:

$$\begin{aligned}
|g' - Id|_{h/4} \leq |f'^{-1} - Id|_{h/2} &\leq \frac{|v'|_{h/2}}{1 - |v'|_{h/2}} \\
&\leq \frac{\delta/(h/2)}{1 - \delta/(h/2)} \\
&\leq \frac{4\delta}{h}.
\end{aligned}$$

□

B.3 Semi-global inversion theorem

When looking at the complex Kepler's equation, we determined a set such that we have an analytic local diffeomorphism induced by the equation at each point. We require a theorem to ensure that we not have only a local diffeomorphism, but a semi-global diffeomorphism on the whole set. We will derive this theorem from a classical global theorem.

Recall the following definitions:

Definition B.3. *Lipeomorphism (or Bi-Lipschitz mapping)*: Let U be an open subset of $\mathbb{R}^n$, and $f : U \to f(U) \subset \mathbb{R}^n$ a one-to-one correspondence. f is called a *lipeomorphism* if f and f^{-1} are Lipschitz continuous.

Definition B.4. *Holomorphic diffeomorphism (or Biholomorphism:* Let U be an open subset of $\mathbb{C}^n$. A holomorphic map $f : U \to f(U)$ is called a *holomorphic diffeomorphism* if f is a one-to-one correspondence and $f^{-1} : U' \to U$ is holomorphic.

We will consider in the following three different framework for the theorem: the case of a Lipschitz function, the case of a $\mathcal{C}^k$ function with $k \geq 1$, and the case of a holomorphic function. To simplify the statements, consider the following definitions:

Definition B.5. Let U be an open subset of $\mathbb{C}^n$, and $f : U \to f(U) \subset \mathbb{R}^n$ a function. If $h, u : U \to \mathbb{R}^n$, with h a lipeomorphism (resp. a $\mathcal{C}^k$-diffeomorphism) and u a Lipschitz function (resp. $\mathcal{C}^k$-function) such that $f = h + u$, we call the couple (h, u) a lipeo-decomposition of f (resp. a $\mathcal{C}^k$-decomposition of f).
Let U be an open subset of $\mathbb{C}^n$, and $f : U \to f(U) \subset \mathbb{C}^n$ a function. If $h, u : U \to \mathbb{C}^n$, with h a biholomorphism and u a holomorphic function such that $f = h + u$, we call the couple (h, u) a holo-decomposition of f.

Observe that the decomposition is not unique.

Lipschitz case:

Recall the following classical global inversion theorem:

Theorem B.6. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ a function and (h, u) a lipeo-decomposition of f such that*

$$\operatorname{lip}(u \circ h^{-1}) < 1, \tag{B.1}$$

then $h + u$ is a lipeomorphism and

$$\operatorname{lip}(f^{-1}) \leq \frac{\operatorname{lip}(h^{-1})}{1 - \operatorname{lip}(u \circ h^{-1})}$$

We will need another theorem before stating our result. This theorem is dealing with the Lipschitz extension of a map Lipschitz on some subset of a Hilbert space, and was stated and proved in the case of an Euclidean space by Kirszbraun.

Theorem B.7. *(Kirszbraun theorem) Let H_1 and H_2 be Hilbert spaces. Let $A \subset H_1$, and $f : A \to H_2$ a K-Lipschitz map. Then there exists $F : H_1 \to H_2$ such that F is K-Lipschitz and $F_{|A} = f$.*

The original proof in the Euclidean case can be found in [34], a full outline of the proof can be found in the very pedagogical paper of Fremlin [23].
Now we can show the following corollary:

Corollary B.8. *Semi-global inversion theorem:*
Let $A \subset \mathbb{R}^n$, be a set, $f : A \to \mathbb{R}^n$ a map such that there exists a lipeo-decomposition (h, u) verifying

$$\mathrm{lip}(u \circ h^{-1}) < 1, \tag{B.2}$$

then $f : A \to f(A)$ is a lipeomorphism and

$$\mathrm{lip}(f^{-1}) \leq \frac{\mathrm{lip}(h^{-1})}{1 - \mathrm{lip}(u \circ h^{-1})}$$

With the two previous theorems, the proof is rather easy.

Proof. Let A and f satisfy the hypotheses of the corollary, call (h, u) the lipeo-decomposition such that $\mathrm{lip}(u \circ h^{-1}) < 1$.
Call $v = u \circ h^{-1}$ and $K = \mathrm{lip}\, v$, we have $f = (Id + v) \circ h$. Using Kirszbraun theorem B.7, there exists $V : \mathbb{R}^n \to \mathbb{R}^n$ such that V is K-Lipschitz, and $V_{|A} = v$. Since Id is a lipeomorphism with Lipschitz constant 1, the hypotheses of theorem B.6 are verified for the function $Id + V$. Hence, $Id + V : \mathbb{R}^n \to \mathbb{R}^n$ is a lipeomorphism and

$$\mathrm{lip}((Id + V)^{-1}) \leq \frac{1}{1 - \mathrm{lip}(u \circ h^{-1})} \tag{B.3}$$

It is therefore still the case for the restriction of $Id + V$ to A, and $Id + v : A \to (Id + v)(A)$ is a lipeomorphism.
Whence, $f : A \to f(A)$ is a lipeomorphism and

$$\mathrm{lip}(f^{-1}) = \mathrm{lip}(h^{-1} \circ (Id + u \circ h^{-1})^{-1}) \leq \frac{\mathrm{lip}\, h^{-1}}{1 - \mathrm{lip}(u \circ h^{-1})} \tag{B.4}$$

□

$\mathcal{C}^k$ case:

Consider now the case of a function that is $\mathcal{C}^k$ for $k \geq 1$. In addition to the global inversion theorem, we need to add a proposition on the regularity. Consider the following (and classical) proposition:

Proposition B.9. *(Regularity) Let $x \in \mathbb{R}^n$, and U neighborhood of x. Let $f : U \to \mathbb{R}^n$ be a lipeomorphism.*

- *If f is differentiable in x, and that $\det f'(x) \neq 0$, then h^{-1} is differentiable in $y = f(x)$ and $(f^{-1})'(y) = h'(x)^{-1}$*
- *If besides f is $\mathcal{C}^k$ with $k \geq 1$, then f^{-1} is also $\mathcal{C}^k$.*

Corollary B.10. *Semi-global inversion theorem $\mathcal{C}^k$:*
Let $k \geq 1$ be an integer, $A \subset \mathbb{R}^n$ be a bounded domain (i.e. a connected and open set), and $f : A \to \mathbb{R}^n$ a $\mathcal{C}^k$-function. Assume there exists a $\mathcal{C}^k$-decomposition (h, u) of f, such that for all $x \in A$, $\det f'(x) \neq 0$, and such that for all closed subset $B \subset A$, $\sup_B \|(u \circ h^{-1})'\| < 1$, then $f : A \to f(A)$ is a $\mathcal{C}^k$-diffeomorphism.

Remark: The hypothesis $\sup_B \|(u \circ h^{-1})'\| < 1$ for any closed subset $B \subset A$ implies that there is no singularity in the open set A, though it is possible to have one on the boundary of A, in other words, it does not prevent the case $\sup_A \|(u \circ h^{-1})'\| = 1$.

Proof. Let A be an bounded domain of $\mathbb{R}^n$ and f a function satisfying the hypotheses of the lemma, with (h, u) its $\mathcal{C}^k$-decomposition.
Let $B \subset A$ be a compact set, then h' and $(h^{-1})'$ are bounded on B, and therefore $h : B \to h(B)$ is a lipeomorphism. For the same reason, u is Lipschitz on B, and therefore $u \circ h^{-1}$ too. Calling $K = \max_B \|(u \circ h^{-1})'\|$, we have $K < 1$, $u \circ h^{-1}$ is therefore K-Lipschitz with $K < 1$. The hypothesis of corollary B.8 are satisfied, and therefore $f_{|B}$ is a lipeomorphism on its image.
We can now apply proposition B.9 to prove the regularity of the function in the set $\mathring{B}$, which is open, and therefore the function $f : \mathring{B} \to f(\mathring{B})$ is a $\mathcal{C}^k$-diffeomorphism.
Choosing a sequence of increasing closed subset of A, $(B_i)_{i \geq 1}$, such that their union is equal to A, we can apply the preceding scheme on each of these sets. Hence, since $\bigcup_{B \subset A,\ B \text{ closed}} \mathring{B} = A$, the function $f : A \to f(A)$ is a $\mathcal{C}^k$-diffeomorphism. □

Complex case:

Corollary B.11. *Holomorphic semi-global inversion theorem:*
Let $A \subset \mathbb{C}^n$ a bounded domain, and $f : A \to \mathbb{C}^n$ a holomorphic map. Assume that f there exists a holo-decomposition (h, u) of f such that for all $z \in A$, its real Jacobian evaluated at the point z is non-null, and such that for all closed subset $B \subset A$, we have $\sup_B \|(u \circ h^{-1})'\| < 1$, then $f : A \to f(A)$ is a biholomorphism.

Proof. The proof is straightforward. Indeed, identifying $\mathbb{C}$ and $\mathbb{R}^2$, the hypotheses of corollary B.10 are satisfied for $k \geq 1$, and the complex valued function can be seen as a diffeomorphism of $2n$ real variables. Besides, its inverse satisfies the Cauchy-Riemann equations (as in the case of the local inversion theorem for a complex valued function), and therefore, its inverse seen as a function of $\mathbb{C}^n$ to $\mathbb{C}^n$ is holomorphic. Hence the corollary. □

Particular case: We will be interested later in the case $n = 1$. Indeed, the hypothesis on the real Jacobian then becomes $f'(z) \neq 0$.

B.4 Classical formulas for analytic multivariate functions

In this section, we will recall the Taylor's theorem and the Cauchy's formula for multivariate functions (see [50] for more details on the latter). As well, we will write the Whitney extension theorem (for the demonstration, see [70]). First, we need some definitions.
Define the following notations for $\alpha \in \mathbb{N}^n$ and $x \in \mathbb{R}^n$ with $n > 0$:

$$|\alpha| = \alpha_1 + ... + \alpha_n$$
$$\alpha! = \alpha_1!...\alpha_n!$$
$$x^\alpha = x_1^{\alpha_1}...x_n^{\alpha_n}$$

Introduce as well for an analytic function f:

$$f^{(\alpha)} = \frac{\partial^{|\alpha|} f}{\partial_{x_1}^{\alpha_1}...\partial_{x_n}^{\alpha_n}}$$

B.4.1 Taylor expansion of analytic function

Theorem B.12. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function analytic at the point $a \in \mathbb{R}^n$. Then, for $k \geq 0$, there exists a function $R_k : \mathbb{R}^n \to \mathbb{R}$ such that:*

$$f(x) - \sum_{|\alpha| \leq k} \frac{f^{(\alpha)}}{\alpha!}(x - a)^\alpha = R_k(x) = o((x - a)^k).$$

Moreover, on a closed ball B around a, we have for $x \in B$:

$$R_k(x) = \sum_{|\beta|=k+1} R_{k,\beta}(x)(x-a)^\beta,$$

with the bound

$$\max_{x\in B} |R_{k,\beta}(x)| \leq \frac{1}{\beta!} \max_{|\alpha|=|\beta|} \left(\max_{x'\in B} f^{(\alpha)}(x') \right).$$

B.4.2 Cauchy Formula

Let $a \in \mathbb{C}^n$ and $\rho = (\rho_1, ..., \rho_n)$ with $\rho_i > 0$. Define the polydisc with center a and radius ρ:

$$P(a,\rho) = \{z \in \mathbb{C}^n, \text{ s.t. } |z_i - a_i| < \rho_i \text{ for } i \in [\![1,n]\!]\}$$

Proposition B.13 (Cauchy's formula)**.** *Let Ω be an open set in $\mathbb{C}^n$, f a function holomorphic on Ω, $a \in \Omega$ and let $\rho = (\rho_1, ..., \rho_n)$ with $\rho_i > 0$ be such that $\overline{P(a,\rho)} \subset \Omega$. Then, for $z \in P(a,\rho)$, we have*

$$f(z) = \frac{1}{(2\pi\imath)^n} \int_{|\zeta_1-a_1|=\rho_1} ... \int_{|\zeta_n-a_n|=\rho_n} \frac{f(\zeta_1,...,\zeta_n)}{(\zeta_1-z_1)...(\zeta_n-z_n)} d\zeta_1...d\zeta_n.$$

Corollary B.14 (Cauchy's inequality)**.** *If f is holomorphic on Ω and $\overline{P(a,\rho)} \subset \Omega$, we have*

$$f^{(\alpha)}(a) \leq \left(\sup_{|\zeta_i-a_i|=\rho_j} |f(\zeta)| \right) \alpha! \rho^{-\alpha}.$$

B.4.3 Whitney theorem

Define, for some function f defined on $\mathbb{R}^n$, some $m \in \mathbb{N}$, and some $\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq n$, the following functions:

$$f_\alpha(x) = \sum_{|k|\leq m-|alpha|} \frac{f_{k+\alpha}(x')}{k!}(x-x')^k + R_\alpha(x,x').$$

Now let A be a closed subset of $\mathbb{R}^n$. We will need some condition of smoothness in this set.

Definition B.15 ($\mathcal{C}^m$ in the Whitney sense)**.** Let f be a function defined in the set A and let m be a positive integer. f is said to be of class $\mathcal{C}^m$ in A in the Whitney sense if the functions f_α ($|\alpha| \leq m$) are defined in A and the remainders R_α are such that for any point x of A, any $\epsilon > 0$, there exists $\delta > 0$ such that if x' and x'' are any two points of $A \cap B(x,\delta)$ then

$$|R_\alpha(x',x'')| \leq |x'-x''|^{m-|\alpha|}\epsilon.$$

With this definition, we have the following statement:

Theorem B.16 (Whitney)**.** *Let A be a closed subset of $\mathbb{R}^n$ and let f be of class $\mathcal{C}^m$ (m finite or infinite) in the Whitney sense. Then there is a function F of class $\mathcal{C}^m$ (in the ordinary sense) in $\mathbb{R}^n$ such that*

(1) $F^{(\alpha)}(x) = f_\alpha(x)$ in A for $|\alpha| \leq m$,

(2) $F(x)$ is analytic in $\mathbb{R}^n$.

In particular, $f = F_{|A}$.

B.5 Solving the cohomological equation

We will prove here a lemma on the solution of the cohomological equation. We are not looking for an optimal statement with optimal exponents and constants such as the one Rüssmann derived in [64].
Recall that $\mathcal{A}_0^s$ is the set of function analytic on a polydisc of size s and of 0-average.

Lemma B.17. *Let $0 < \sigma < s \leq 1$, $\alpha \in D(\gamma, \tau)$ and $g \in \mathcal{A}_0^s$. There exists a unique function $f \in \mathcal{A}_0^{s-\sigma}$ such that*

$$\partial_\alpha f = g.$$

Besides, there exists a constant C_0 depending only on n and τ such that

$$|f|_{s-\sigma} \leq \frac{C_0}{\gamma \sigma^{n+\tau}} |g|_s.$$

Proof. We have, for $\theta \in \mathbb{T}_s^n$, the Fourier expansion of g:

$$g(\theta) = \sum_{k \in \mathbb{Z}^n \setminus \{0\}} g_k \exp(\imath k \cdot \theta$$

Formally, it is straightforward to see that the unique solution f of the cohomological equation is:

$$f(\theta) = \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{g_k}{\imath \alpha \cdot k} \exp(\imath k \cdot \theta.$$

Since α verifies a Diophantine condition, the coefficients

$$f_k = \frac{g_k}{\imath \alpha \cdot k}$$

are well-defined. The function g being analytic on the polydisc of size s, we have

$$|g_k| \leq |g|_s \exp(-|k|_1 s).$$

Hence, injecting in the formula of f and using the Diophantine condition, we find:

$$\begin{aligned} |f|_{s-\sigma} &\leq \frac{|g|_s}{\gamma} \sum_{k \in \mathbb{Z}^n \setminus \{0\}} |k|_1^\tau \exp(-|k|_1 \sigma) \\ &\leq \frac{4^n |g|_s}{\gamma} \sum_{l \geq 1} l^{n-1+\tau} \exp(-l\sigma) \\ &\leq \frac{4^n |g|_s}{\gamma \sigma^{n+\tau-1}} \sum_{l \geq 1} (l\sigma)^{n-1+\tau} \exp(-l\sigma) \end{aligned}$$

The last sum is roughly equivalent to

$$\frac{1}{\sigma} \int_0^{+\infty} x^{n+\tau} \exp(-x) dx,$$

where the integral depends only on n and τ, whence the lemma. □

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables.* Applied mathematics series. Dover Publications, 1964.

[2] J. C. Adams. Explanation of the observed irregularities in the motion of Uranus, on the hypothesis of disturbance by a more distant planet; with a determination of the mass, orbit, and position of the disturbing body. *Monthly Notices of the Royal Astronomical Society*, 7:149–152, 1846.

[3] G. B. Arfken and H. J. Weber. *Mathematical Methods For Physicists International Student Edition.* Elsevier Science, 2005.

[4] V. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. In A. B. Givental, B. A. Khesin, J. E. Marsden, A. N. Varchenko, V. A. Vassiliev, O. Y. Viro, and V. M. Zakalyukin, editors, *Collected Works: Representations of Functions, Celestial Mechanics and KAM Theory, 1957–1965*, pages 306–412. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[5] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1989. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein.

[6] 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. [Dynamical systems. III], Translated from the Russian original by E. Khukhro.

[7] D. Bambusi. An introduction to birkhoff normal form. `http://users.mat.unimi.it/users/bambusi/pedagogical.pdf`, 2014.

[8] B. Bates, M. Bunder, and K. Tognetti. Locating terms in the Stern-Brocot tree. *European J. Combin.*, 31(3):1020–1033, 2010.

[9] B. Bates, M. Bunder, and K. Tognetti. Child's addition in the stern–brocot tree. *European Journal of Combinatorics*, 33(2):148 – 167, 2012.

[10] A. Celletti and L. Chierchia. Rigorous estimates for a computer-assisted KAM theory. *J. Math. Phys.*, 28(9):2078–2086, 1987.

[11] A. Celletti and L. Chierchia. On the stability of realistic three-body problems. *Comm. Math. Phys.*, 186(2):413–449, 1997.

[12] A. Celletti and L. Chierchia. Kam stability for a three-body problem of the solar system. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 57(1):33–41, 2005.

[13] A. Celletti and L. Chierchia. KAM stability for a three-body problem of the solar system. *Z. Angew. Math. Phys.*, 57(1):33–41, 2006.

[14] A. Chenciner and J. Laskar. Groupe de travail sur la lecture de : Méthodes nouvelles de la Mécanique Céleste (H. Poincaré). Comptes rendus 18 et 25 février 1988, 24 mars 1988. Alain Chenciner : Intégration du problème de Kepler par la méthode de Hamilton-Jacobi : coordonnées "action-angles" de Delaunay; Jacques Laskar : Les variables de Poincaré et le développement de la fonction pertubatrice. *Notes Scientifiques et Techniques du Bureau des Longitudes, #S026. 1989, 58 pp.*, 26, 1989.

[15] L. Chierchia and Gabriella Pinzari. The planetary n-body problem: symplectic foliation, reductions and invariant tori. *Inventiones mathematicae*, 186(1):1–77, 2011.

[16] J. L. de Lagrange and M. J. A. Serret. *Oeuvres de Lagrange*. Number v. 6 in Oeuvres de Lagrange. Gauthier-Villars, 1873.

[17] P.S. de Laplace and Académie des sciences (France). *Œuvres complètes de Laplace: Mémoires divers. 1904-12*. Œuvres complètes de Laplace: publiées sous les auspices de l'Académie des sciences. Gauthier-Villars, 1904.

[18] J. Féjoz. Global secular dynamics in the planar three-body problem. *Celestial Mech. Dynam. Astronom.*, 84(2):159–195, 2002.

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

[20] 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.

[21] J. Féjoz. On action-angle coordinates and the Poincaré coordinates. *Regul. Chaotic Dyn.*, 18(6):703–718, 2013.

[22] J. Fejoz. The three- and the N-body problem. In Alessandra Celletti, editor, *Celestial Mechanics*, pages 1–30. Unesco-EOLSS, 2013.

[23] D. H. Fremlin. Kirszbraun's theorem. `http://www.essex.ac.uk/maths/people/fremlin/n11706.pdf`.

[24] J. Féjoz. *Introduction to KAM theory, with a view to celestial mechanics*. Radon Series on Computational and Applied Mathematics. de Gruyter, 2016.

[25] E. Ghys. Résonances et petits diviseurs. `http://perso.ens-lyon.fr/ghys/articles/resonancesdiviseurs.pdf`.

[26] A. Giorgilli. Notes on exponential stability of hamiltonian systems. `http://www.mat.unimi.it/users/antonio/ricerca/papers/pisa.pdf`, 2002.

[27] J. Hadamard. Henri poincaré, le problème des trois corps. `https://fr.wikisource.org/wiki/Henri_Poincar%C3%A9,_le_probl%C3%A8me_des_trois_corps`, 1914.

[28] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

[29] M. Hénon. Letter to Arnold. Letter, 1965.

[30] H. Hofer and E. Zehnder. *Symplectic invariants and Hamiltonian dynamics*. Modern Birkhäuser Classics. Birkhäuser Verlag, Basel, 2011. Reprint of the 1994 edition.

[31] W. H. Jefferys and J. Moser. Quasi-periodic solutions for the three-body problem. *Astronom. J.*, 71:568–578, 1966.

[32] T. Kappeler and J. Pöschel. *KdV & KAM*, volume 45 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 2003.

[33] A. Y. Khinchin. *Continued fractions*. Dover Publications, Inc., Mineola, NY, russian edition, 1997. With a preface by B. V. Gnedenko, Reprint of the 1964 translation.

[34] M. Kirszbraun. Über die zusammenziehende und lipschitzsche transformationen. *Fundamenta Mathematicae*, 22(1):77–108, 1934.

[35] A. N. Kolmogorov. On dynamical systems with an integral invariant on the torus. *Doklady Akad. Nauk SSSR (N.S.)*, 93:763–766, 1953.

[36] A. N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954.

[37] A. N. Kolmogorov. Preservation of conditionally periodic movements with small change in the Hamilton function. In *Stochastic behavior in classical and quantum Hamiltonian systems (Volta Memorial Conf., Como, 1977)*, volume 93 of *Lecture Notes in Phys.*, pages 51–56. Springer, Berlin-New York, 1979.

[38] P.-S. Laplace. *Essai philosophique sur les probabilités*. Bachelier, 1840.

[39] J. Laskar. Le Système solaire est-il stable ?, 2010. Séminaire Poincaré XIV.

[40] J. Laskar. Andoyer construction for Hill and Delaunay variables. *ArXiv e-prints*, September 2014.

[41] J. Laskar and P. Robutel. Stability of the planetary three-body problem. I. Expansion of the planetary Hamiltonian. *Celestial Mech. Dynam. Astronom.*, 62(3):193–217, 1995.

[42] U. J. J. Le Verrier. *Annales de l'observatoire impérial de Paris: Observations de Paris*. Annales de l'observatoire de Paris. Mallez-Bachelier, 1858.

[43] B. B. Lieberman. Existence of quasi-periodic solutions to the three-body problem. *Celestial mechanics*, 3(4):408–426, 1971.

[44] U. Locatelli and A. Giorgilli. Invariant tori in the secular motions of the three-body planetary systems. *Celestial Mech. Dynam. Astronom.*, 78(1-4):47–74 (2001), 2000. New developments in the dynamics of planetary systems (Badhofgastein, 2000).

[45] U. Locatelli and A. Giorgilli. Invariant tori in the Sun-Jupiter-Saturn system. *Discrete Contin. Dyn. Syst. Ser. B*, 7(2):377–398, 2007.

[46] P. Lochak. Canonical perturbation theory via simultaneous approximation. *Russian Mathematical Surveys*, 47(6):57, 1992.

[47] P. Lochak and A. I. Neishtadt. Estimates of stability time for nearly integrable systems with a quasiconvex hamiltonian. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2(4):495–499, 1992.

[48] M. Mignotte. *Approximation des nombres algébriques*. Département de Mathématique, Université de Paris-Sud, Orsay, 1977. Publications Mathématiques d'Orsay, No. 77-74.

[49] J. Moser. Convergent series expansions for quasi-periodic motions. *Mathematische Annalen*, 169:136–176, 1967.

[50] R. Narasimhan. *Several complex variables.* Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1995. Reprint of the 1971 original.

[51] NASA. Asteroid Main-Belt Distribution, June 2007. Online.

[52] N. N. Nekhoroshev. An exponential estimate of the time of stability of nearly-integrable hamiltonian systems. *Russian Mathematical Surveys*, 32(6):1–65, 1977.

[53] I. Newton, A. Motte, and J. Machin. *The Mathematical Principles of Natural Philosophy.* Number vol. 1 in The Mathematical Principles of Natural Philosophy. B. Motte, 1729.

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

[55] M. Niqui. Exact arithmetic on the Stern-Brocot tree. *J. Discrete Algorithms*, 5(2):356–379, 2007.

[56] M. S. Petrovskaya. Expansions of the negative powers of mutual distances between bodies. *Celestial mechanics*, 3(1):121–128, 1970.

[57] H. Poincaré. Mémoires et observations. Sur le développement de la fonction perturbatrice. *Bulletin Astronomique, Serie I*, 15:449–464, 1898.

[58] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste. Tome I.* Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics]. Librairie Scientifique et Technique Albert Blanchard, Paris, 1987.

[59] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste. Tome II.* Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics]. Librairie Scientifique et Technique Albert Blanchard, Paris, 1987.

[60] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste. Tome III.* Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics]. Librairie Scientifique et Technique Albert Blanchard, Paris, 1987.

[61] H. Poincaré. *Leçons de mécanique céleste: professées à la Sorbonne.* Gauthier-Villars (Paris), 1905-1910.

[62] J. Pöschel. A lecture on the classical KAM theorem. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 707–732. Amer. Math. Soc., Providence, RI, 2001.

[63] P. Robutel. Stability of the planetary three-body problem. II. KAM theory and existence of quasiperiodic motions. *Celestial Mech. Dynam. Astronom.*, 62(3):219–261, 1995.

[64] H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. volume 38, pages 598–624. Springer Berlin Heidelberg, Berlin, Heidelberg, 1975.

[65] C. L. Siegel and J. K. Moser. *Lectures on celestial mechanics.* Classics in mathematics. Springer, Paris, 1971.

[66] G. Silva. Sur les limites de convergence du développement de la fonction perturbatrice. *Bulletin Astronomique, Serie I*, 26:49–75, 1909.

[67] V. G. Sokolov. Convergence of the expansions of the disturbing functions in the planar three-body planetary problem. *Solar System Research*, 41(2):162–170, 2007.

[68] K. F. Sundman. Nouvelles recherches sur le problème des trois corps. volume t. 35 (1909), page 558. Helsingfors :Suomen Tiedeseura,, 1909.

[69] K. F. Sundman. Sur les Conditions Nécessaires et Suffisantes Pour la Convergence du Développement de la Fonction Perturbatrice Dans le Mouvement Plan. volume 58, pages 1–19. [S.l. :s.n.] ;, 1916.

[70] H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.*, 36(1):63–89, 1934.

[71] A. Wintner. *The Analytical Foundations of Celestial Mechanics.* Princeton Mathematical Series, v. 5. Princeton University Press, Princeton, N. J., 1941.

[72] J.-C. Yoccoz. An introduction to small divisors problems. In *From number theory to physics (Les Houches, 1989)*, pages 659–679. Springer, Berlin, 1992.

[73] W. H. Young. On the Theory of Functions of two Complex Variables. *Proc. London Math. Soc.*, S2-21(1):441, 1922.